

DISSERTATION

---

ZUR KONVERGENZ DISKRETER  
LEAST-SQUARES METHODEN AUF  
ÄQUIDISTANTEN STÜTZSTELLEN

---

RENÉ GOERTZ

$$LS_n^N[f]$$

$$\sup_{x \in [a, b]} |f^{(n+1)}(x)|$$

$$P_k^{\alpha, \beta}$$

$$\sum_{k=0}^n \hat{f}_k Q_k$$

$$Q_k(\cdot; \alpha, \beta, N)$$

$$x_\mu = -1 + \frac{2\mu}{N}$$

$$\min_{\varphi \in U}$$

$$\langle f, Q_k \rangle_\omega$$

$$(\varphi(x_\mu) - f(x_\mu))^2$$

$$\mathcal{BV}[a, b]$$



Cuvillier Verlag Göttingen  
Internationaler wissenschaftlicher Fachverlag



Zur Konvergenz diskreter Least-Squares Methoden auf äquidistanten Stützstellen





**Zur Konvergenz  
diskreter Least-Squares Methoden  
auf äquidistanten Stützstellen**

Von der  
Carl-Friedrich-Gauß-Fakultät  
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades eines

**Doktors der Naturwissenschaften (Dr. rer. nat.)**

genehmigte Dissertation

von  
René Goertz  
geboren am 22.05.1987  
in Wolfsburg

Eingereicht am: 04.01.2018  
Disputation am: 01.02.2018  
1. Referent: Prof. Dr. Thomas Sonar  
2. Referent: Prof. Dr. Klaus-Jürgen Förster

2018



Bibliografische Information der Deutschen Nationalbibliothek  
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der  
Deutschen Nationalbibliografie; detaillierte bibliografische Daten  
sind im Internet über <http://dnb.d-nb.de> abrufbar.

1. Aufl. - Göttingen: Cuvillier, 2018  
Zugl.: (TU) Braunschweig, Univ., Diss., 2018

© CUVILLIER VERLAG, Göttingen 2018  
Nonnenstieg 8, 37075 Göttingen  
Telefon: 0551-54724-0  
Telefax: 0551-54724-21  
[www.cuvillier.de](http://www.cuvillier.de)

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung  
des Verlages ist es nicht gestattet, das Buch oder Teile  
daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie)  
zu vervielfältigen.

1. Auflage, 2018

Gedruckt auf umweltfreundlichem, säurefreiem Papier  
aus nachhaltiger Forstwirtschaft.

ISBN 978-3-7369-9743-1  
eISBN 978-3-7369-8743-2



# Vorwort

Die vorliegende Arbeit ist während meiner Tätigkeit als wissenschaftlicher Mitarbeiter des Instituts für Mathematik und Angewandte Informatik der Universität Hildesheim in Kooperation mit dem Institut Computational Mathematics der Technischen Universität Braunschweig entstanden.

In Braunschweig gilt mein besonderer Dank meinem langjährigen mathematischen Lehrer und Doktorvater, Professor Dr. Thomas Sonar, für seine fachliche und persönliche Betreuung. Seine Begeisterung für Mathematik und seine wissenschaftliche Professionalität sind mir von Anfang an ein richtungsweisendes Vorbild gewesen.

In Hildesheim danke ich herzlich meinem Mentor und Institutsleiter, Professor Dr. Klaus-Jürgen Förster, für die umfangreiche Unterstützung. Die vielen anregenden Gespräche haben mir wichtige Impulse für meine Arbeit und zugleich einen anderen Blickwinkel auf die Thematik gegeben. Darüber hinaus bedanke ich mich für das stets vertrauensvolle und harmonische Arbeitsklima.

Ein weiterer Dank geht an meine Arbeitskollegen während meiner Lehrtätigkeit an der Universität Hildesheim, insbesondere an Dr. Karl-Heinz Schlote, für das freundschaftliche Arbeitsklima.

Außerdem möchte ich mich bei meinem Bruder Marcel Goertz und insbesondere bei meinem Braunschweiger Kollegen Dr. Philipp Öffner für eine detaillierte Durchsicht von Teilen einer früheren Version dieser Arbeit bedanken.

Meiner Familie, meinen Freunden und besonders meinen Eltern sowie meiner Lebensgefährtin danke ich sowohl für ihre menschliche Unterstützung als auch für ihr Verständnis, dass während meiner gesamten Promotionszeit die mathematische Forschung häufig im Vordergrund stand.





# Inhaltsverzeichnis

<b>Zusammenfassung</b>	<b>1</b>
<b>1 Einleitung</b>	<b>3</b>
<b>2 Vorbereitende Grundlagen</b>	<b>9</b>
2.1 Funktionenklassen . . . . .	9
2.1.1 Analytische Funktionen . . . . .	9
2.1.2 Funktionen von beschränkter Variation . . . . .	9
2.1.3 Stetigkeitsmodul . . . . .	11
2.2 Spezielle Funktionen . . . . .	12
2.2.1 Gammafunktion . . . . .	12
2.2.2 Verallgemeinerte hypergeometrische Funktion . . . . .	14
2.3 Lineare Operatoren . . . . .	16
2.4 Quadratur . . . . .	17
<b>3 Orthogonale Polynome</b>	<b>19</b>
3.1 Stetige orthogonale Polynome . . . . .	19
3.1.1 Definition und Eigenschaften . . . . .	19
3.1.2 Jacobi-Polynome . . . . .	21
3.2 Diskrete orthogonale Polynome . . . . .	24
3.2.1 Definition und Eigenschaften . . . . .	24
3.2.2 Hahn-Polynome . . . . .	26
<b>4 Approximation mit Polynomen</b>	<b>29</b>
4.1 Polynominterpolation . . . . .	29
4.1.1 Polynominterpolation in äquidistanten Stützstellen . . . . .	32
4.1.2 Polynominterpolation in Tschebyscheff-Stützstellen . . . . .	34
4.2 Approximation mit Bernsteinpolynomen . . . . .	37
4.3 Die Methode der kleinsten Quadrate . . . . .	39
4.3.1 Der kontinuierliche Fall . . . . .	39
4.3.2 Der diskrete Fall . . . . .	41
4.4 Polynome bester Approximation . . . . .	44
<b>5 Divergenz der Methode der kleinsten Quadrate</b>	<b>47</b>
<b>6 Punktweise Konvergenz der Methode der kleinsten Quadrate</b>	<b>51</b>
6.1 Hauptergebnis zur punktweisen Konvergenz . . . . .	51
6.2 Folgerungen aus dem Hauptergebnis . . . . .	74



<b>7</b>	<b>Gleichmäßige Konvergenz der Methode der kleinsten Quadrate</b>	<b>79</b>
7.1	Hauptergebnis zur gleichmäßigen Konvergenz . . . . .	79
7.2	Folgerungen aus dem Hauptergebnis . . . . .	85
7.3	Vergleich zum kontinuierlichen Fall . . . . .	90
7.4	Vergleich zur Polynominterpolation . . . . .	92
7.5	Vergleich zur Approximation mit Bernsteinpolynomen . . . . .	93
7.6	Vergleich zu anderer Stützstellenwahl . . . . .	94
7.7	Vergleich zum Polynom bester Approximation . . . . .	98
<b>8</b>	<b>Numerische Resultate</b>	<b>99</b>
8.1	Vergleich zu anderen Approximationsmethoden . . . . .	100
8.1.1	Vergleich zum kontinuierlichen Fall . . . . .	101
8.1.2	Vergleich zur Approximation mit Bernsteinpolynomen . . . . .	103
8.1.3	Vergleich zur Polynominterpolation . . . . .	105
8.2	Variation der Stützstellenanzahl . . . . .	107
8.2.1	Vergleich mit größerer Stützstellenanzahl . . . . .	108
8.2.2	Vergleich mit geringerer Stützstellenanzahl . . . . .	110
<b>9</b>	<b>Ausblick</b>	<b>115</b>
	<b>Literaturverzeichnis</b>	<b>121</b>

# Zusammenfassung

Wir betrachten die bekannte Methode der kleinsten Quadrate auf einem äquidistanten Gitter mit  $N + 1$  Stützstellen auf dem Intervall  $[-1, 1]$  mit dem Ziel der Approximation einer Funktion  $f \in \mathcal{C}[-1, 1]$  durch ein Polynom  $n$ -ten Grades. Wir untersuchen das folgende Problem:

Für welches Verhältnis  $N/n$  und welche Funktionen konvergiert die zugehörige Operatorfolge  $LS_n^N : \mathcal{C}[-1, 1] \rightarrow \mathcal{P}_n$  der Methode der kleinsten Quadrate? Dabei gehen wir sowohl auf punktweise als auch auf gleichmäßige Konvergenz ein.

Wir untersuchen die obige Fragestellung bezogen auf eine diskrete Gewichtung vom Jacobi-Typ. Dementsprechend analysieren wir zunächst die Beziehung zwischen den Jacobi-Polynomen  $P_k^{\alpha, \beta}$  und den Hahn-Polynomen  $Q_k(\cdot; \alpha, \beta, N)$ . Hierzu drücken wir den Operator  $LS_n^N$  durch eine abgebrochene Reihenentwicklung einer Funktion durch Hahn-Polynome aus.

Ohne zusätzliche Voraussetzungen an Funktionen  $f \in \mathcal{C}[-1, 1]$  kann keine gleichmäßige Konvergenz garantiert werden. Wir können jedoch für  $\alpha = \beta$  unter zusätzlichen Bedingungen an  $f$  und  $(N_n)_{n \in \mathbb{N}}$  Konvergenz garantieren und beweisen folgende Ergebnisse:

Die Reihenentwicklung  $\sum_{k=0}^n \hat{f}_k Q_k$  einer Funktion  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  durch Hahn-Polynome  $Q_k$  konvergiert für  $\alpha \geq 0$  punktweise, wenn die Reihenentwicklung  $\sum_{k=0}^n \hat{f}_k P_k$  der Funktion  $f$  durch Jacobi-Polynome  $P_k$  punktweise konvergiert und wenn  $n^{3+\alpha+\max\{1, \alpha\}}/N \rightarrow 0$  für  $n, N \rightarrow \infty$  gilt.

Hieraus folgt unter anderem folgendes Resultat:

Sei  $f \in \{g \in \mathcal{C}^1[-1, 1] : g' \in \mathcal{BV}[-1, 1]\}$  und sei  $(N_n)_n$  eine Folge von natürlichen Zahlen mit  $n^4/N_n \rightarrow 0$ . Dann konvergiert für  $\alpha = 0$  die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  für jedes  $x \in [-1, 1]$ .

Bezüglich der gleichmäßigen Konvergenz beweisen wir folgendes Resultat:

Für fest gewähltes  $\alpha \geq 0$  sei  $f \in \left\{ g \in \mathcal{C}^\infty[-1, 1] : \lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} |g^{(n)}(x)| \frac{n^{\alpha+1/2}}{2^n n!} = 0 \right\}$  und sei  $(N_n)_n$  eine Folge natürlicher Zahlen mit  $N_n \geq 2n(n+1)$ . Dann konvergiert die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  gleichmäßig auf  $[-1, 1]$ .



# 1 Einleitung

Vor über 200 Jahren haben unter anderem Legendre und Gauß begonnen mit der Methode der kleinsten Quadrate zu arbeiten (vgl., z.B., [57]). Seitdem wird diese Methode in vielen Gebieten der Mathematik verwendet und ist heutzutage ein Standardwerkzeug der Angewandten Mathematik (vgl., z.B., [1], [6], [14], [35], [71]). In dieser Arbeit untersuchen wir Approximationseigenschaften dieser Methode.

Die Methode der kleinsten Quadrate ist folgendermaßen definiert (vgl., z.B., [35, S. 59], [34, S. 217], [58, S. 291]):

Seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, N$ . Weiter sei  $\omega : [a, b] \rightarrow \mathbb{R}$  eine auf  $\{x_\mu\}_{\mu=0}^N$  positive Funktion, die sogenannte Gewichtsfunktion. Für ein  $n \leq N$  sei  $U$  ein Unterraum von  $\mathcal{C}[a, b]$  mit  $\dim U = n + 1$  auf  $\{x_\mu\}_{\mu=0}^N$ . Der Operator der Methode der kleinsten Quadrate  $LS_n^N : \mathcal{C}[a, b] \rightarrow U$  ist eindeutig definiert durch die Eigenschaft

$$\sum_{\mu=0}^N \left( LS_n^N[f](x_\mu) - f(x_\mu) \right)^2 \omega(x_\mu) = \min_{\varphi \in U} \sum_{\mu=0}^N (\varphi(x_\mu) - f(x_\mu))^2 \omega(x_\mu). \quad (1.1)$$

In dieser Arbeit untersuchen wir den folgenden Fall:

- $U = \mathcal{P}_n$  ist der Raum der Polynome vom Höchstgrad  $n$ ,
- $\{x_0, \dots, x_N\}$  ist ein äquidistantes Gitter mit  $N + 1$  Stützstellen auf dem Intervall  $[-1, 1]$ , d. h.  $x_\mu = -1 + 2\mu/N$  für  $\mu = 0, \dots, N$ .

Diese Situation tritt in der Praxis häufig auf: Polynome sind seit Jahrhunderten eine bewährte und intensiv untersuchte Funktionenklasse zur Approximation. Zudem lassen sie sich sehr effektiv auf Computern einsetzen, da für jede Berechnung nur die elementaren Operationen Addition und Multiplikation verwendet werden. Äquidistant erhobene Informationen liegen häufig vor, insbesondere bedingt durch Datenerhebung auf großen (oftmals mehrdimensionalen) äquidistanten Gittern.

Ohne Beschränkung der Allgemeinheit sei in dieser Arbeit das Intervall  $[a, b]$  auf das Standardintervall  $[-1, 1]$  normiert.

Wir untersuchen folgendes **Problem**:

Für welche Funktionen  $f \in K \subset \mathcal{C}[-1, 1]$  und welches Verhältnis  $N/n$  konvergiert die Folge  $(LS_n^N[f])$ ? Dabei betrachten wir sowohl punktweise als auch gleichmäßige Konvergenz.

Um dieses Problem zu untersuchen, drücken wir den Operator  $LS_n^N$  in Abschnitt 4.3 durch eine abgebrochene Reihenentwicklung einer Funktion durch Hahn-Polynome

$Q_k(\cdot; \alpha, \beta, N)$  aus. Die Hahn-Polynome  $Q_k \equiv Q_k(\cdot; \alpha, \beta, N)$  sind klassische diskrete orthogonale Polynome auf dem Intervall  $I = [0, N]$  vom Grad  $k$ . Sie sind orthogonal auf  $I$  bezüglich des Skalarprodukts

$$\langle f, g \rangle_\omega := \sum_{i=0}^N f(i)g(i)\omega(i), \quad (1.2)$$

wobei die Gewichtsfunktion  $\omega$  gegeben ist durch

$$\omega(x) := \binom{\alpha + x}{x} \binom{\beta + N - x}{N - x}. \quad (1.3)$$

Sie sind durch

$$\langle Q_k(\cdot; \alpha, \beta, N), Q_k(\cdot; \alpha, \beta, N) \rangle_\omega = \frac{(-1)^k (k + \alpha + \beta + 1)_{N+1} (\beta + 1)_k k!}{(2k + \alpha + \beta + 1)(\alpha + 1)_k (-N)_k N!} \quad (1.4)$$

normiert (vgl., z.B., [47, S. 204]). Weitere Eigenschaften der Hahn-Polynome diskutieren wir in Abschnitt 3.2 über diskrete orthogonale Polynome.

Es ist bekannt, dass der Operator der Methode der kleinsten Quadrate  $LS_n^N$  durch die Verwendung von Hahn-Polynomen folgendermaßen repräsentiert werden kann (vgl., z.B., [35, S. 62-63], [58, S. 270], [77, S. 218-232]):

$$LS_n^N[f] = \sum_{k=0}^n \frac{\langle f\left(\frac{2}{N}(\cdot) - 1\right), Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k\left(\frac{N}{2}(1 + \cdot)\right), \quad (1.5)$$

mit  $f \in \mathcal{C}[-1, 1]$ .

Wir merken zunächst in Kapitel 5 an, dass wir ohne eine zusätzliche Voraussetzung an die Funktionen  $f \in \mathcal{C}[-1, 1]$  keine gleichmäßige Konvergenz der Folge  $(LS_n^N[f])$  garantieren können.

Bezogen auf unser oben genanntes Problem ist es also vollkommen egal welches Verhältnis  $N/n$  wir zwischen der Anzahl der Stützstellen  $N + 1$  und dem Polynomgrad  $n$  wählen: Es gibt zu jedem Verhältnis  $N/n$  eine stetige Funktion  $f \in \mathcal{C}[-1, 1]$ , sodass die Methode der kleinsten Quadrate  $LS_n^N[f]$  nicht gleichmäßig gegen  $f$  konvergiert. Um die gleichmäßige Konvergenz garantieren zu können, müssen wir also die Funktionenklasse verkleinern. Dies untersuchen wir in Kapitel 7.

Wir beschäftigen uns zuvor in Kapitel 6 mit der punktweisen Konvergenz und stellen dazu einen Vergleich zum kontinuierlichen Fall an. Hierbei untersuchen wir die Beziehung zwischen den Jacobi-Polynomen  $P_k^{\alpha, \beta}$  und den Hahn-Polynomen  $Q_k(\cdot; \alpha, \beta, N)$ . Die Jacobi-Polynome behandeln wir genauer in Abschnitt 3.1 über stetige orthogonale Polynome. Die folgende Beziehung zwischen den Hahn-Polynomen  $Q_k$  und den Jacobi-Polynomen  $P_k \equiv P_k^{\alpha, \beta}$  ist bereits bekannt. Es gilt

$$\lim_{N \rightarrow \infty} (-1)^k \binom{k + \alpha}{k} Q_k\left(\frac{N}{2}(1 + x); \alpha, \beta, N\right) = P_k^{\beta, \alpha}(x), \quad (1.6)$$

für jedes  $x \in [-1, 1]$  (vgl., z.B., [60, S. 45]). Hier, wie auch in der gesamten Arbeit, sind bei asymptotischen Aussagen die Parameter  $\alpha$  und  $\beta$  fest gewählt. Die Jacobi-Polynome  $P_k^{\alpha, \beta}$  sind orthogonal auf dem Intervall  $I = [-1, 1]$  bezüglich des Skalarprodukts

$$(f, g)_\varrho := \int_{-1}^1 f(x)g(x)\varrho(x)dx, \quad (1.7)$$

wobei

$$\varrho(x) := (1-x)^\alpha(1+x)^\beta \quad (1.8)$$

die Gewichtsfunktion ist. Sie sind normiert durch

$$\left(P_k^{\alpha, \beta}, P_k^{\alpha, \beta}\right)_\varrho = \frac{2^{\alpha+\beta+1}\Gamma(k+\alpha+1)\Gamma(k+\beta+1)}{(2k+\alpha+\beta+1)k!\Gamma(k+\alpha+\beta+1)} \quad (1.9)$$

(vgl., z.B., [47, S. 217]). Aufgrund von Gleichung (1.6) können die Hahn-Polynome als das diskrete Analogon der Jacobi-Polynome gesehen werden. Die Entwicklung einer Funktion  $f$  in Jacobi-Polynomen ist gut untersucht und wird zur Approximation der Funktion  $f$  häufig verwendet (vgl., z.B., [3], [8], [10], [17], [42], [50], [64]). Analog zur Darstellung der abgebrochenen Reihenentwicklung durch Hahn-Polynome in Gleichung (1.5) hat die Entwicklung in Jacobi-Polynomen die Form:

$$LS_n[f] = \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k. \quad (1.10)$$

Bezogen auf diese beiden Reihenentwicklungen beweisen wir in Kapitel 6 folgende, neue Abschätzung:

Sei  $\alpha \geq 0$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2}\sqrt{(2\alpha+1)(2\alpha+2N+1)}. \quad (1.11)$$

Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  gilt

$$\begin{aligned} & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \right| + \mathcal{O} \left( \frac{n^{3+\alpha+\max\{1, \alpha\}}}{N} \right), \end{aligned} \quad (1.12)$$

mit  $n \leq n(\alpha, N)$ , für jedes  $x \in [-1, 1]$ .

Hier, wie auch bei allen anderen in dieser Arbeit von uns erzielten Approximationsresultaten, betrachten wir den wichtigen symmetrischen (sogenannten ultrasphärischen) Fall  $\alpha = \beta$ . Aus (1.12) erhalten wir bezüglich der punktweisen Konvergenz folgendes Ergebnis:

Die Reihenentwicklung  $\sum_{k=0}^n \hat{f}_k Q_k$  einer Funktion  $f$  durch Hahn-Polynome  $Q_k$  konvergiert punktweise, wenn die Reihenentwicklung  $\sum_{k=0}^n \hat{f}_k P_k$  der Funktion  $f$  durch Jacobi-Polynome  $P_k$  punktweise konvergiert und wenn  $n^{3+\alpha+\max\{1,\alpha\}}/N \rightarrow 0$  für  $n, N \rightarrow \infty$  gilt. Als Beispiel für die vielfältigen Anwendungsmöglichkeiten sei hier das folgende Ergebnis für den wichtigen Spezialfall  $\alpha = 0$  genannt:

Sei  $f \in K := \{g \in \mathcal{C}^1[-1, 1] : g' \in \mathcal{BV}[-1, 1]\}$  und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge von natürlichen Zahlen mit

$$\lim_{n \rightarrow \infty} \frac{n^4}{N_n} = 0. \quad (1.13)$$

Dann konvergiert die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  mit  $\alpha = 0$  für jedes  $x \in [-1, 1]$ .

Es gibt viele weitere Anwendungen der obigen Abschätzung bei denen wir die punktweise Konvergenz erhalten. Denn für etliche Klassen von Funktionen wurde in den letzten Jahrzehnten die Konvergenz der zugehörigen Reihenentwicklungen in Jacobi-Polynomen gezeigt. Dies ist in Kapitel 6 ausführlich dargestellt.

In diesem engen Zusammenhang zwischen der Entwicklung einer Funktion nach Jacobi-Polynomen, vgl. (1.10), und der Entwicklung nach Hahn-Polynomen, vgl. (1.5), liegt die für mich grundlegende Motivation zu meinen hier vorgestellten, von Thomas Sonar und Tom Koornwinder angeregten, Untersuchungen: Die Entwicklung nach Jacobi-Polynomen ist seit Jahrzehnten eine bewährte Methode zur Modellierung insbesondere techno-mathematischer Fragestellungen, da hier viele physikalisch-technische Eigenschaften bei der Modellierung mit erfasst werden können. Allerdings müssen zur Ermittlung der Koeffizienten in (1.10) Integrale ausgewertet werden. Dies geschieht üblicherweise durch Diskretisierung mit Hilfe quadraturtheoretischer Methoden. Da hier also bei der Integralapproximation sowieso diskretisiert werden muss, ist die Frage naheliegend, ob man auf dem direkten Weg (also ohne Berechnung von Integralen) mit Hilfe diskreter Orthogonalpolynome gleichwertige Ergebnisse erzielen kann. In dieser Hinsicht untersuchen wir in Kapitel 7 die Methode der kleinsten Quadrate auf gleichmäßige Konvergenz. Das Hauptergebnis lautet:

Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (1.14)$$

Weiter sei

$$D_{n,N} := \frac{2^{n+1} \Gamma(n + 2\alpha + 2) \Gamma(n + \alpha + 2) N!}{(n + 1)! \Gamma(2n + 2\alpha + 3) \Gamma(\alpha + 1) N^{n+1} (N - n - 1)!}. \quad (1.15)$$

Dann gilt für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$  mit  $n + 1 \leq n(\alpha, N)$

$$\sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2} (1 + x) \right) \right| \leq D_{n,N} \sup_{x \in [-1, 1]} |f^{(n+1)}(x)|. \quad (1.16)$$

Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $D_{n,N}$  in Ungleichung (1.16) durch keine kleinere Zahl ersetzt werden kann.

Hiermit eröffnet sich nun eine direkte Vergleichsmöglichkeit zwischen dem diskreten Ansatz und der klassischen (stetigen) Vorgehensweise durch Betrachtung des Maximalfehlers („worst case“) in den Funktionenklassen

$$\mathcal{K}_{n+1} := \left\{ f \in \mathcal{C}^{n+1}[-1, 1] : \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \leq 1 \right\}. \quad (1.17)$$

Dieser Maximalfehler ist gemäß (1.16) der Wert  $D_{n,N}$  und dieser ist überraschenderweise (zumindest für mich) für  $n + 1 \leq \frac{1}{2} - \alpha + \frac{1}{2}\sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}$  für jedes so gewählte Verhältnis  $N/n$  kleiner als im entsprechenden klassischen Fall, man vergleiche hierzu Abschnitt 7.3.

Als Beispiel für die vielfältigen weiteren Anwendungsmöglichkeiten des obigen Hauptergebnisses aus Kapitel 7 sei hier das folgende Ergebnis genannt:

Sei  $\alpha \geq 0$ , sei

$$f \in K := \left\{ g \in \mathcal{C}^\infty[-1, 1] : \lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} |g^{(n)}(x)| \frac{n^{\alpha + \frac{1}{2}}}{2^n n!} = 0 \right\} \quad (1.18)$$

und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge mit  $N_n \geq 2n(n + 1)$ . Dann konvergiert die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  gleichmäßig auf dem Intervall  $[-1, 1]$ .

Wir erhalten aus dem obigen Resultat weitere Abschätzungen, welche wir in Kapitel 7 darstellen und diskutieren. Anschließend vergleichen wir unsere Approximationsresultate mit entsprechenden Resultaten bekannter Approximationsmethoden, welche wir in Kapitel 4 ausführlich darstellen.

In Kapitel 8 präsentieren wir beispielhaft numerische Resultate zur Methode der kleinsten Quadrate. Zunächst vergleichen wir numerisch den diskreten mit dem kontinuierlichen Fall der Methode der kleinsten Quadrate. Außerdem vergleichen wir die Methode im diskreten Fall mit anderen in Kapitel 4 vorgestellten Approximationsmethoden. Weiterhin untersuchen wir, basierend auf unseren Ergebnissen in den vorherigen Kapiteln, numerisch für ausgewählte Testfunktionen bei welchem Verhältnis  $N/n$  der Maximalfehler bei wachsenden  $n$  und  $N$  ansteigt.

Im letzten Kapitel 9 geben wir einen Ausblick auf eine mögliche Anwendung der Methode der kleinsten Quadrate im diskreten Fall bei Spektralen-Differenzen-Verfahren zur Lösung hyperbolischer Erhaltungsgleichungen, die seit vielen Jahren in der Braunschweiger Forschungsgruppe von Thomas Sonar, Martina Wirz, Philipp Öffner und weiteren Mitgliedern höchst erfolgreich untersucht werden.





## 2 Vorbereitende Grundlagen

In diesem Kapitel stellen wir vorbereitende Grundlagen zusammen, welche wir für die weiteren Kapitel dieser Arbeit benötigen.

Wir beginnen mit der Einführung einiger wichtiger Funktionenklassen, wie beispielsweise die Klasse der Funktionen von beschränkter Variation.

Insbesondere beschäftigen wir uns mit den sogenannten „Speziellen Funktionen“, deren Studium ein eigenes mathematisches Teilgebiet ist. Wir stellen insbesondere Eigenschaften der Gammafunktion und der verallgemeinerten hypergeometrischen Funktion bereit.

Im Anschluss behandeln wir in Hinsicht auf die Kapitel 4 und 5 die Theorie der linearen Operatoren.

Abschließend stellen wir ausgewählte Grundlagen der Quadraturtheorie zusammen, die in Kapitel 6 benötigt werden.

### 2.1 Funktionenklassen

#### 2.1.1 Analytische Funktionen

Im Folgenden definieren wir die Klasse der analytischen Funktionen, die aus der Funktionentheorie wohlbekannt sind. Wir kommen später in dieser Arbeit auf diese Funktionenklasse zurück.

**Definition 2.1** (vgl., z.B., [33, S. 45]). *Sei  $D \subset \mathbb{C}$  offen. Weiter sei eine Funktion  $f : D \rightarrow \mathbb{C}$  gegeben. Dann heißt  $f$  **analytisch**, falls  $f$  in jedem Punkt  $a \in D$  komplex differenzierbar ist. Weiterhin heißt  $f$  im Punkt  $a$  **analytisch**, falls es eine Umgebung  $U \subset D$  von  $a \in U$  gibt, sodass  $f$  auf  $U$  analytisch ist.*

**Satz 2.2** (vgl., z.B., [33, S. 106]). *Sei  $D \subset \mathbb{C}$  offen. Weiter sei eine analytische Funktion  $f : D \rightarrow \mathbb{C}$  gegeben. Dann ist  $f$  auf jeder in  $D$  enthaltenen Kreisscheibe  $\mathcal{B}_R(a) := \{z \in \mathbb{C} : |z - a| < R\} \subset D$  in eine Potenzreihe entwickelbar und es gilt*

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (z - a)^n, \text{ für alle } z \in \mathcal{B}_R(a). \quad (2.1)$$

#### 2.1.2 Funktionen von beschränkter Variation

In diesem Unterabschnitt beschäftigen wir uns mit der Klasse der Funktionen von beschränkter Variation. Diese Klasse spielt besonders in Kapitel 6, in welchem wir

uns mit der punktweisen Konvergenz der Methode der kleinsten Quadrate auseinandersetzen, eine wichtige Rolle. Um diese Klasse zu definieren, benötigen wir zunächst die folgende Definition der totalen Variation:

**Definition 2.3** (vgl., z.B., [41, S. 493]). Die **totale Variation** einer reellwertigen Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist definiert durch

$$\mathcal{V}_a^b[f] = \sup_{P \in \mathcal{P}} \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)|, \quad (2.2)$$

wobei das Supremum auf der Menge

$$\mathcal{P} = \{P = \{x_0, \dots, x_n\} : a = x_0 < x_1 < \dots < x_{n-1} < x_n = b, n \in \mathbb{N}\} \quad (2.3)$$

über alle Partitionen des Intervalls  $[a, b]$  gebildet wird. Der Kürze halber schreiben wir  $\mathcal{V}[f] \equiv \mathcal{V}_{-1}^1[f]$ .

Wir können nun die Klasse von Funktionen von beschränkter Variation definieren.

**Definition 2.4** (vgl., z.B., [41, S. 493]). Eine reellwertige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist von **beschränkter Variation**, wenn ihre totale Variation endlich ist. Wir definieren die Klasse von Funktionen von beschränkter Variation auf dem Intervall  $[a, b]$  durch

$$\mathcal{BV}[a, b] = \{f : [a, b] \rightarrow \mathbb{R} : \exists C > 0 : \mathcal{V}_a^b[f] < C\}. \quad (2.4)$$

Es stellt sich die Frage, wie man Funktionen von beschränkter Variation charakterisieren kann. Zudem ist es interessant, welche gängigen Funktionenklassen Teil der Klasse von Funktionen von beschränkter Variation sind und welche nicht. Wir fassen einige wichtige Eigenschaften zusammen (vgl., z.B., [41, S. 496-499]):

- Die Menge  $\mathcal{BV}[a, b]$  enthält alle Treppenfunktionen sowie alle monotonen oder Lipschitz-stetigen Funktionen. Insbesondere enthält  $\mathcal{BV}[a, b]$  also auch alle stetig differenzierbaren Funktionen  $f \in \mathcal{C}^1[a, b]$  mit beschränkter Ableitung  $f'$  auf  $[a, b]$ .
- Es existieren stetige Funktionen  $f \in \mathcal{C}[a, b]$ , die nicht von beschränkter Variation sind. Ein Beispiel dafür ist die bekannte Funktion  $f : [0, 1] \rightarrow \mathbb{R}$ , die durch

$$f(x) := \begin{cases} x \cos\left(\frac{\pi}{x}\right) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0 \end{cases} \quad (2.5)$$

definiert ist.

- Seien  $f \in \mathcal{BV}[a, b]$  und  $c \in (a, b)$  gegeben. Dann gilt

$$\mathcal{V}_a^b[f] = \mathcal{V}_a^c[f] + \mathcal{V}_c^b[f]. \quad (2.6)$$

Funktionen von beschränkter Variation können folgendermaßen charakterisiert werden:

**Satz 2.5** (vgl., z.B., [41, S. 498]). *Eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist genau dann von beschränkter Variation, wenn sie als Differenz zweier monoton steigender Funktionen dargestellt werden kann:*

$$f \in \mathcal{BV}[a, b] \iff \exists f_1, f_2 : [a, b] \rightarrow \mathbb{R} \text{ monoton steigend: } f = f_1 - f_2. \quad (2.7)$$

*Insbesondere gilt für stetige Funktionen, dass sie genau dann von beschränkter Variation sind, wenn sie sich als Differenz zweier stetiger und monoton steigender Funktionen darstellen lassen.*

Daraus lässt sich unmittelbar für eine Funktion von beschränkter Variation  $f \in \mathcal{BV}[a, b]$  folgern (vgl., z.B., [41, S. 498]):

- Für alle  $x_0 \in [a, b]$  existieren die links- und rechtsseitigen Grenzwerte

$$\lim_{x \rightarrow x_0^+} f(x) \text{ und } \lim_{x \rightarrow x_0^-} f(x). \quad (2.8)$$

- Es existieren höchstens abzählbar viele Stellen  $\{x_i\}_{i \in I} \subset [a, b]$  an denen  $f$  unstetig ist.
- $f$  ist Riemann-integrierbar.

Unser Hauptergebnis Theorem 6.25 aus Kapitel 6 beweist eine Fehlerabschätzung für Funktionen der Klasse  $\mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$ . Mit Hilfe der soeben zusammengestellten Eigenschaften lassen sich leicht Funktionen dieser Klasse angeben, auf die wir Theorem 6.25 anwenden können.

### 2.1.3 Stetigkeitsmodul

In diesem Unterabschnitt betrachten wir das Stetigkeitsmodul einer stetigen Funktion. Das Stetigkeitsmodul wurde erstmals von H. Lebesgue im Jahr 1910 eingeführt (vgl. [52]). Wir benötigen es in Abschnitt 6.2, in welchem wir das Hauptergebnis zur punktweisen Konvergenz der Methode der kleinsten Quadrate genauer diskutieren. Mit Hilfe des Stetigkeitsmoduls lässt sich die Klasse der stetigen Funktionen differenzierter betrachten.

**Definition 2.6** (vgl., z.B., [63, S. 116]). *Für  $f \in \mathcal{C}[a, b]$  heißt die durch*

$$\omega(f, [a, b], \delta) = \sup \{|f(x) - f(t)| : x, t \in [a, b] \text{ und } |x - t| \leq \delta\}, \quad (2.9)$$

*definierte Funktion  $\omega(f, [a, b], \cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  **Stetigkeitsmodul** von  $f$  auf  $[a, b]$ .*

Mit Hilfe des Stetigkeitsmoduls lassen sich gleichmäßig stetige Funktionen wie folgt charakterisieren:

**Satz 2.7** (vgl., z.B., [63, S. 117]). *Eine Funktion  $f$  ist im Intervall  $[a, b]$  genau dann gleichmäßig stetig, wenn*

$$\lim_{\delta \rightarrow 0} \omega(f, [a, b], \delta) = 0. \quad (2.10)$$

Wir fassen im Folgenden einige Eigenschaften des Stetigkeitsmoduls einer Funktion  $f \in \mathcal{C}[a, b]$  zusammen (vgl., z.B., [70, S. 179]):

- Die Funktion  $\omega(f, [a, b], \cdot)$  ist monoton steigend, d. h. für alle  $\delta_1 \leq \delta_2$  gilt

$$\omega(f, [a, b], \delta_1) \leq \omega(f, [a, b], \delta_2). \quad (2.11)$$

- Die Funktion  $\omega(f, [a, b], \cdot)$  ist subadditiv, d. h. für alle  $\delta_1, \delta_2 \geq 0$  gilt

$$\omega(f, [a, b], \delta_1 + \delta_2) \leq \omega(f, [a, b], \delta_1) + \omega(f, [a, b], \delta_2). \quad (2.12)$$

- Für alle  $m \in \mathbb{N}_0$  gilt

$$\omega(f, [a, b], m\delta) \leq m\omega(f, [a, b], \delta). \quad (2.13)$$

- Für alle  $\lambda \geq 0$  gilt

$$\omega(f, [a, b], \lambda\delta) \leq (\lambda + 1)\omega(f, [a, b], \delta). \quad (2.14)$$

- Die Funktion  $\omega(f, [a, b], \cdot)$  ist stetig.

## 2.2 Spezielle Funktionen

In den folgenden Unterabschnitten stellen wir für die Beweise unserer Theoreme wichtige Eigenschaften der Gammafunktion und der verallgemeinerten hypergeometrischen Funktion bereit. Diese gehören zu den sogenannten „Speziellen Funktionen“, die nicht nur in der Analysis, sondern auch in vielen Anwendungsbereichen eine wichtige Rolle spielen.

### 2.2.1 Gammafunktion

Die Gammafunktion, im Folgenden mit  $\Gamma$ -Funktion abgekürzt, ist eine „Spezielle Funktion“, von der wir im Folgenden wesentliche Eigenschaften zusammenfassen. Die  $\Gamma$ -Funktion ist auf der rechten Halbebene der komplexen Zahlen definiert und wurde erstmals von L. Euler 1729/30 eingeführt:

**Definition 2.8** (vgl., z.B., [33, S. 194]). *Die Funktion*

$$\Gamma : \{z \in \mathbb{C} : \operatorname{Re}(z) > 0\} \rightarrow \mathbb{C}, \quad (2.15)$$

*definiert durch*

$$\Gamma(z) := \int_0^{\infty} t^{z-1} e^{-t} dt \quad (2.16)$$

*heißt  $\Gamma$ -Funktion. Dabei ist  $t^{z-1} := e^{(z-1)\log(t)}$ , mit  $\log(t) \in \mathbb{R}$ .*

**Anmerkung 2.9** (vgl., z.B., [33, S. 195]). *Das sogenannte Gammaintegral*

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt \quad (2.17)$$

konvergiert in der Halbebene  $\{z \in \mathbb{C} : \operatorname{Re}(z) > 0\}$  absolut, d. h.  $\Gamma(z)$  ist für jedes  $z \in \mathbb{C}$  mit  $\operatorname{Re}(z) > 0$  wohldefiniert. Insbesondere stellt die  $\Gamma$ -Funktion in der rechten Halbebene eine analytische Funktion dar.

Außerdem sind von der  $\Gamma$ -Funktion viele weitere interessante Eigenschaften bekannt. Einige für uns hier wichtige sind im Folgenden aufgeführt (vgl., z.B., [33, S. 194–209]):

- Die  $\Gamma$ -Funktion lässt sich auf  $\mathbb{C} \setminus \{0, -1, -2, -3, \dots\}$  eindeutig analytisch fortsetzen.
- Die  $\Gamma$ -Funktion genügt der Funktionalgleichung

$$\Gamma(z+1) = z\Gamma(z) \text{ für alle } z \in \mathbb{C} \setminus \{0, -1, -2, -3, \dots\}. \quad (2.18)$$

- Man kann die  $\Gamma$ -Funktion als Erweiterung der Fakultät auffassen, denn es gilt

$$\Gamma(n+1) = n! \text{ für alle } n \in \mathbb{N}_0. \quad (2.19)$$

- (*Gauß'sche Produktentwicklung*) Es gilt für alle  $z \in \mathbb{C}$

$$\frac{1}{\Gamma(z)} = \lim_{n \rightarrow \infty} \frac{n^{-z}}{n!} z(z+1)(z+2) \cdots (z+n). \quad (2.20)$$

- Die  $\Gamma$ -Funktion besitzt keine Nullstellen.
- (*Ergänzungssatz*, L. Euler, 1749) Für alle  $z \in \mathbb{C} \setminus \mathbb{Z}$  gilt

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}. \quad (2.21)$$

Darüber hinaus genügt die  $\Gamma$ -Funktion der folgenden asymptotischen Eigenschaft, welche wir sehr häufig verwenden:

**Lemma 2.10** (vgl., z.B., [2, S. 257]). *Für  $a, b > 0$  gilt*

$$N^{b-a} \frac{\Gamma(N+a)}{\Gamma(N+b)} = 1 + \frac{(a-b)(a+b-1)}{2N} + \mathcal{O}\left(\frac{1}{N^2}\right). \quad (2.22)$$

Im Folgenden definieren wir weitere Funktionen, die sich aus Termen von Gammafunktionen zusammensetzen und die im Verlauf dieser Arbeit noch benötigt werden. Zunächst definieren wir das Pochhammer-Symbol, welches in der Definition der verallgemeinerten hypergeometrischen Funktion verwendet wird.

**Definition 2.11** (vgl., z.B., [47, S. 4]). Für beliebige  $z \in \mathbb{C}$  setzen wir

$$(z)_0 := 1 \text{ und } (z)_k := \prod_{i=1}^k (z + i - 1) \text{ für alle } k \in \mathbb{N}. \quad (2.23)$$

Die für alle  $k \in \mathbb{N}_0$  definierten  $(z)_k$  heißen **Pochhammer-Symbole**.

**Anmerkung 2.12.** Aufgrund der Funktionalgleichung (2.18) der  $\Gamma$ -Funktion lässt sich das Pochhammer-Symbol auch folgendermaßen darstellen (vgl., z.B., [2, S. 256]):

$$(z)_k = \frac{\Gamma(z + k)}{\Gamma(z)} \text{ für alle } k \in \mathbb{N}. \quad (2.24)$$

Weiterhin gilt der interessante Spezialfall (vgl., z.B., [47, S. 4])

$$(1)_k = k! \text{ für alle } k \in \mathbb{N}_0. \quad (2.25)$$

Nun definieren wir den verallgemeinerten Binomialkoeffizienten. Diesen benötigen wir unter anderem bei der Definition der Hahn-Polynome.

**Definition 2.13** (vgl., z.B., [47, S. 4]). Für beliebige  $\alpha \in \mathbb{C} \setminus \{-1, -2, -3, \dots\}$  und  $\beta \in \mathbb{C}$  ist der **Binomialkoeffizient**  $\binom{\alpha}{\beta}$  definiert durch

$$\binom{\alpha}{\beta} := \frac{\Gamma(\alpha + 1)}{\Gamma(\beta + 1)\Gamma(\alpha - \beta + 1)}. \quad (2.26)$$

**Anmerkung 2.14.** Insbesondere gilt für  $n, k \in \mathbb{N}_0$  mit  $n \geq k$  der bekannte Spezialfall (vgl., z.B., [47, S. 4])

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}, \quad (2.27)$$

der unmittelbar aus der Beziehung (2.19) folgt.

**Anmerkung 2.15.** Für  $n \in \mathbb{N}_0$  und für  $\alpha > -1$  gilt die asymptotische Eigenschaft

$$\binom{n + \alpha}{n} = \mathcal{O}(n^\alpha), \quad (2.28)$$

die unmittelbar aus Gleichung (2.22) folgt.

## 2.2.2 Verallgemeinerte hypergeometrische Funktion

In diesem Unterabschnitt betrachten wir die verallgemeinerte hypergeometrische Funktion. Wir benötigen sie später im Zusammenhang mit orthogonalen Polynomen.

**Definition 2.16** (vgl., z.B., [47, S. 5]). Für beliebige  $r, s \in \mathbb{N}_0$  definieren wir die Funktion  ${}_rF_s$  durch

$${}_rF_s \left( \begin{matrix} a_1, \dots, a_r \\ b_1, \dots, b_s \end{matrix}; z \right) := \sum_{k=0}^{\infty} \frac{(a_1)_k \cdot \dots \cdot (a_r)_k}{(b_1)_k \cdot \dots \cdot (b_s)_k} \frac{z^k}{k!}. \quad (2.29)$$

Die Funktion  ${}_rF_s$  heißt **verallgemeinerte hypergeometrische Funktion**. Die zugehörige Reihe

$$\sum_{k=0}^{\infty} \frac{(a_1)_k \cdot \dots \cdot (a_r)_k}{(b_1)_k \cdot \dots \cdot (b_s)_k} \frac{z^k}{k!} \quad (2.30)$$

heißt **hypergeometrische Reihe**.

**Anmerkung 2.17.** Die Parameter  $a_i, b_j$  sollten so gewählt werden, dass die einzelnen Faktoren in den Nennern der Terme der hypergeometrischen Reihe von null verschieden sind. Ansonsten divergiert die Reihe. Beispielsweise sind in der Reihe

$${}_1F_1 \left( \begin{matrix} 1 \\ -1 \end{matrix}; z \right) = \sum_{k=0}^{\infty} \frac{(1)_k}{(-1)_k} \frac{z^k}{k!} = \sum_{k=0}^{\infty} \frac{z^k}{(-1)_k} \quad (2.31)$$

die Nenner  $(-1)_k$  für alle  $k \geq 2$  null. Denn es gilt

$$(-1)_k = \prod_{i=1}^k (-1 + i - 1) = -1 \cdot 0 \cdot 1 \cdot \dots \cdot (k - 2) = 0. \quad (2.32)$$

Ist andererseits ein Parameter  $a_i$  des Zählers gegeben durch  $a_i = -n$ , wobei  $n \in \mathbb{N}_0$ , so ist die hypergeometrische Funktion  ${}_rF_s$  ein Polynom  $n$ -ten Grades in  $z$  (vgl., z.B., [47, S. 5]). Denn dann gilt für alle  $k > n$

$$(-n)_k = -n \cdot (-n + 1) \cdot \dots \cdot (-n + (n + 1) - 1) \cdot \dots \cdot (-n + k - 1) = 0 \quad (2.33)$$

und man erhält

$${}_rF_s \left( \begin{matrix} a_1, \dots, a_r \\ b_1, \dots, b_s \end{matrix}; z \right) = \sum_{k=0}^n \frac{(a_1)_k \cdot \dots \cdot (a_r)_k}{(b_1)_k \cdot \dots \cdot (b_s)_k} \frac{z^k}{k!}. \quad (2.34)$$

Trifft weder der eine noch der andere oben genannte Fall zu, so ist der Konvergenzradius  $R$  der hypergeometrischen Reihe gegeben durch

$$R = \begin{cases} \infty & \text{für } r < s + 1, \\ 1 & \text{für } r = s + 1, \\ 0 & \text{für } r > s + 1 \end{cases} \quad (2.35)$$

(vgl., z.B., [47, S. 5]).

Der Begriff „hypergeometrisch“ ist zum ersten Mal von J. Wallis in seiner Arbeit „Arithmetica Infinitorum“ im Jahr 1655 verwendet worden (vgl. [72, S. 1]). Er untersuchte dabei Verallgemeinerungen der gewöhnlichen geometrischen Reihe

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + q^3 + \dots \quad (2.36)$$

Inbesondere beschäftigte er sich dabei genauer mit der Reihe

$$\sum_{k=0}^{\infty} (a)_k = 1 + a + a(a+1) + a(a+1)(a+2) + \dots \quad (2.37)$$

(vgl. [72, S. 1]). In den darauf folgenden 150 Jahren haben sich viele weitere Mathematiker, darunter Euler und Gauß, mit ähnlichen Reihen auseinandergesetzt. Zentraler Gegenstand war dabei die Gauß Reihe, die auch als gewöhnliche hypergeometrische Reihe bezeichnet wird. Sie ist ein Spezialfall der verallgemeinerten hypergeometrischen Funktion und ist durch

$${}_2F_1[a, b; c; z] := \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!} = {}_2F_1\left(\begin{matrix} a, b \\ c \end{matrix}; z\right) \quad (2.38)$$

definiert (vgl., z.B., [72, S. 1]).

## 2.3 Lineare Operatoren

In diesem Abschnitt stellen wir einige grundlegende Aussagen über lineare Operatoren in normierten Räumen bereit, die wir zur Untersuchung der Divergenz der Methode der kleinsten Quadrate in Kapitel 5 benötigen.

**Definition 2.18** (vgl., z.B., [77, S. 45]). *Seien  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$  normierte Vektorräume über  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Eine Abbildung  $T : X \rightarrow Y$  heißt **linearer Operator**, falls für alle  $x, y \in X$  und  $\lambda \in K$  die Eigenschaften*

- $T(\lambda x) = \lambda T(x)$ ,
- $T(x + y) = T(x) + T(y)$

gelten.

Eine fundamentale Klassifizierung von linearen Operatoren ist die Unterscheidung zwischen beschränkten und unbeschränkten linearen Operatoren. Die Beschränktheit ist dabei äquivalent zur Stetigkeit der Operatoren (vgl., z.B., [77, S. 46]).

**Definition 2.19** (vgl., z.B., [77, S. 46]). *Seien  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$  normierte Vektorräume über  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Ein linearer Operator  $T : X \rightarrow Y$  heißt **beschränkt**, falls es eine Konstante  $M \geq 0$  gibt, sodass für alle  $x \in X$  die Ungleichung*

$$\|T(x)\|_Y \leq M \|x\|_X \quad (2.39)$$

erfüllt ist. Mit  $L(X, Y)$  wird der folgende Raum bezeichnet:

$$L(X, Y) := \{T : X \rightarrow Y : T \text{ ist linear und beschränkt}\}. \quad (2.40)$$

Der Raum  $L(X, Y)$  der linearen beschränkten Operatoren ist in folgendem Sinn selbst ein normierter Vektorraum:

**Satz 2.20** (vgl., z.B., [77, S. 47]). *Seien  $(X, \|\cdot\|_X)$  und  $(Y, \|\cdot\|_Y)$  normierte Vektorräume über  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Dann wird durch*

$$\|T\| := \sup_{\|x\|_X \leq 1} \|T(x)\|_Y, \quad (2.41)$$

für alle  $T \in L(X, Y)$ , eine Operatornorm

$$\|\cdot\| : L(X, Y) \rightarrow \mathbb{R} \quad (2.42)$$

definiert.

Die Norm  $\|T\|$  eines Operators  $T \in L(X, Y)$  entspricht also der kleinstmöglichen Konstante  $M \geq 0$  aus Definition 2.19, d. h. es gilt

$$\|T\| = \inf \{M \geq 0 : \|T(x)\|_Y \leq M \|x\|_X \quad \forall x \in X\} \quad (2.43)$$

(vgl., z.B., [77, S. 46]).

Des Weiteren benötigen wir den Begriff der Projektion, welchen wir im Folgenden definieren:

**Definition 2.21** (vgl., z.B., [77, S. 161]). *Sei  $(X, \|\cdot\|_X)$  ein normierter Vektorraum über  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Weiter sei  $U$  ein Untervektorraum von  $X$ . Eine Abbildung  $P : X \rightarrow U$  heißt **Projektion**, oder auch **Projektionsoperator**, falls*

$$P \circ P = P \text{ und } P(U) = U. \quad (2.44)$$

Für Kapitel 4 stellen wir abschließend den Begriff des positiven Operators auf dem Raum der stetigen Funktionen  $\mathcal{C}[a, b]$  bereit:

**Definition 2.22** (vgl., z.B., [63, S. 249]). *Ein linearer Operator  $T : \mathcal{C}[a, b] \rightarrow \mathcal{C}[c, d]$  heißt **positiv**, falls aus  $f \geq 0$  stets  $T(f) \geq 0$  folgt.*

## 2.4 Quadratur

In diesem Abschnitt stellen wir in Bezug auf Kapitel 6 ausgewählte Grundlagen der Quadraturtheorie bereit. Quadraturverfahren gehören zu den wichtigsten numerischen Grundlagen-Methoden. Unter anderem sucht und analysiert man in der Quadraturtheorie Verfahren, die zur Auswertung des Integrals

$$\int_a^b f(x) dx \quad (2.45)$$

über einer stetigen Funktion  $f \in \mathcal{C}[a, b]$  dienen (vgl., z.B., [18, S. 9]). Denn nur selten ist es möglich zu einer Funktion  $f \in \mathcal{C}[a, b]$  eine Stammfunktion  $F$  explizit zu bestimmen, mit deren Hilfe sich das obige Integral einfach berechnen ließe.

**Definition 2.23** (vgl., z.B., [18, S. 12-13]). Sei ein Intervall  $[a, b] \subset \mathbb{R}$  und ein  $n \in \mathbb{N}$  gegeben. Weiter seien Werte  $x_k \in [a, b]$  und  $a_k \in \mathbb{R}$ , für  $k = 1, \dots, n$  gegeben. Dann heißt ein lineares Funktional  $Q_n : \mathcal{C}[a, b] \rightarrow \mathbb{R}$ , welches durch

$$Q_n[f] = \sum_{k=1}^n a_k f(x_k), \quad (2.46)$$

für alle  $f \in \mathcal{C}[a, b]$  definiert ist, **Quadraturformel**  $n$ -ter Ordnung auf dem Grundintervall  $[a, b]$ . Die  $a_k$  heißen **Gewichte** der Quadraturformel und die  $x_k$  heißen **Stützstellen** der Quadraturformel. Eine Folge von Quadraturformeln wachsender Ordnung zum gleichen Grundintervall heißt **Quadraturverfahren**.

Quadraturverfahren werden vor allem nach den Kriterien Konvergenz, Genauigkeit, Einfachheit, Zugänglichkeit der Restabschätzung, Fehlerfortpflanzung und Stützstellenwahl untersucht (vgl., z.B., [18, S. 15-16]). Wir gehen im Folgenden näher auf das in Kapitel 6 benötigte, sogenannte Trapezverfahren ein.

**Definition 2.24** (vgl., z.B., [18, S. 20]). Das durch

$$Q_{n+1}^{Tr}[f] := \frac{b-a}{n} \left[ \frac{1}{2}f(a) + \sum_{k=2}^n f\left(a + (k-1)\frac{b-a}{n}\right) + \frac{1}{2}f(b) \right] \quad (2.47)$$

definierte Quadraturverfahren, heißt **Trapezverfahren**.

**Anmerkung 2.25.** Das Trapezverfahren  $Q_{n+1}^{Tr}$  basiert auf der sogenannten **elementaren Trapezregel**

$$Q_2^{Tr}[f] := \frac{b-a}{2} [f(a) + f(b)] \quad (2.48)$$

für das Intervall  $[a, b]$ . Transformation dieser elementaren Trapezregel auf die Intervalle

$$\left[ a + (k-1)\frac{b-a}{n}, a + k\frac{b-a}{n} \right] \quad (2.49)$$

für  $k = 1, \dots, n$  und anschließende Addition ergibt das Trapezverfahren  $Q_{n+1}^{Tr}$ .

Man findet in der Literatur zahlreiche Fehlerabschätzungen zum Trapezverfahren, siehe dazu beispielsweise [21]. Wir benötigen insbesondere in Kapitel 6 die folgende Abschätzung:

**Satz 2.26** (vgl., z.B., [21, S. 218, (7.14)]). Für jedes  $f \in \mathcal{C}[a, b] \cap \mathcal{BV}[a, b]$  und alle  $n \in \mathbb{N}$  gilt

$$\left| \int_a^b f(x) dx - Q_{n+1}^{Tr}[f] \right| \leq \frac{b-a}{2n} \mathcal{V}[f]. \quad (2.50)$$

## 3 Orthogonale Polynome

In diesem Kapitel betrachten wir orthogonale Polynome und ihre Eigenschaften. Orthogonale Polynome sind seit Ende des 19. Jahrhunderts vor allem in der Analysis ein wichtiger Bestandteil, sowohl vieler tieflyingender theoretischer Ergebnisse wie auch zahlreicher praktischer Anwendungen. Sie weisen Verbindungen zu trigonometrischen, hypergeometrischen, Bessel und elliptischen Funktionen auf. Darüber hinaus spielen sie in den Bereichen Approximation und Quadratur häufig eine sehr entscheidende Rolle. Zudem sind sie in der Theorie der Differential- und Integralgleichungen ein unverzichtbares Hilfsmittel. Für weitere Anwendungsgebiete sei etwa auf [47], [60], [73] verwiesen. Für uns ist im Wesentlichen der Zusammenhang zur Methode der kleinsten Quadrate von Bedeutung. Dabei stellen wir den Operator der Methode der kleinsten Quadrate durch eine abgebrochene Reihenentwicklung in orthogonalen Polynomen dar. Hierbei beschäftigen wir uns insbesondere mit den Jacobi-Polynomen und den Hahn-Polynomen. Wir unterscheiden im Folgenden zwischen stetigen orthogonalen Polynomen und diskreten orthogonalen Polynomen, in Abhängigkeit von dem jeweiligen inneren Produkt.

### 3.1 Stetige orthogonale Polynome

#### 3.1.1 Definition und Eigenschaften

**Definition 3.1** (vgl., z.B., [47, S. 1]). *Sei ein Intervall  $(a, b)$  mit  $a, b \in \mathbb{R} \cup \{\infty\}$  gegeben. Weiter sei eine stetige, stückweise stetige oder Riemann-integrierbare Funktion  $\omega : (a, b) \rightarrow (0, \infty)$  gegeben, sodass*

$$0 < \int_a^b x^{2n} \omega(x) dx < \infty \text{ für alle } n \in \mathbb{N}_0 \quad (3.1)$$

*gilt. Ein System  $\{p_n\}_{n \in \mathbb{N}_0}$  von Polynomen, wobei für jedes  $n \in \mathbb{N}_0$  das Polynom  $p_n$  den Polynomgrad  $n$  hat, heißt **orthogonal** auf dem Intervall  $(a, b)$  bezüglich  $\omega$ , falls*

$$\int_a^b p_n(x) p_m(x) \omega(x) dx = 0 \text{ für alle } n, m \in \mathbb{N}_0, \text{ mit } n \neq m \quad (3.2)$$

*gilt. Die Funktion  $\omega$  heißt dann **Gewichtsfunktion**.*

**Anmerkung 3.2.** *Bei Systemen  $\{p_n\}_{n \in \mathbb{N}_0}$ , die die Voraussetzungen aus Definition 3.1 erfüllen, sprechen wir, zur Unterscheidung von den später definierten diskreten orthogonalen Polynomen, auch von **stetigen orthogonalen Polynomen** (vgl., z.B., [47, S. 2]).*

**Anmerkung 3.3.** In diesem Zusammenhang sprechen wir auch von Orthogonalität bezüglich des durch

$$(f, g) := \int_a^b f(x)g(x)\omega(x)dx \quad (3.3)$$

definierten Skalarprodukts.

Durch die Gleichung (3.2) werden die jeweiligen Systeme von orthogonalen Polynomen bis auf konstante Faktoren festgelegt (vgl., z.B., [47, S. 2]).

Wir wollen in der folgenden Anmerkung einige für uns wichtige Eigenschaften von stetigen orthogonalen Polynomen zusammenfassen.

**Anmerkung 3.4.** Sei ein System  $\{p_n\}_{n \in \mathbb{N}_0}$  aus orthogonalen Polynomen gegeben, welches orthogonal auf dem Intervall  $(a, b)$  bezüglich des Skalarprodukts

$$(f, g) := \int_a^b f(x)g(x)\omega(x)dx \quad (3.4)$$

ist. Dann gelten folgende Eigenschaften (vgl., z.B., [60, S. 11-15]):

- Jedes Polynom  $q$   $n$ -ten Grades lässt sich eindeutig darstellen durch

$$q = \sum_{k=0}^n \frac{(q, p_k)}{(p_k, p_k)} p_k. \quad (3.5)$$

- Die orthogonalen Polynome  $p_n$  erfüllen eine Rekursionsgleichung der Form

$$xp_n(x) = \alpha_n p_{n+1}(x) + \beta_n p_n(x) + \gamma_n p_{n-1}(x), \quad (3.6)$$

wobei  $\alpha_n$ ,  $\beta_n$  und  $\gamma_n$  Konstanten sind.

- Die Nullstellen der Polynome  $p_n$  sind alle einfach und liegen im Intervall  $(a, b)$ .

Von besonderer Bedeutung sind orthogonale Polynome, die eine Gleichung vom hypergeometrischen Typ erfüllen. Dies ist eine gewöhnliche Differentialgleichung, welche wir im Folgenden definieren:

**Definition 3.5** (vgl., z.B., [60, S. 2]). Eine Differentialgleichung der Form

$$\sigma y'' + \tau y' + \lambda y = 0, \quad (3.7)$$

wobei  $\sigma$  ein Polynom höchstens zweiten Grades,  $\tau$  ein Polynom höchstens ersten Grades und  $\lambda$  eine Konstante ist, heißt **Gleichung vom hypergeometrischen Typ**. Lösungen einer solchen Gleichung heißen **Funktionen vom hypergeometrischen Typ**. Sind diese Lösungen Polynome, so heißen diese **Polynome vom hypergeometrischen Typ**.

Viele Probleme der Angewandten Mathematik sowie der Mathematischen Physik führen zu Gleichungen vom hypergeometrischen Typ. Gleichung (3.7) lässt sich durch Multiplikation einer Funktion  $\varrho$  in folgende selbstadjungierte Form bringen:

$$(\sigma \varrho y)' + \lambda \varrho y = 0, \quad (3.8)$$

wobei  $\varrho$  der Differentialgleichung

$$(\sigma \varrho)' = \tau \varrho \quad (3.9)$$

genügen muss. Die einfachsten Lösungen dieser Gleichungen sind Polynome. Wir betrachten die Lösungen auf einem Intervall  $(a, b)$  mit  $a, b \in \mathbb{R} \cup \{\infty\}$ . Erfüllt die Funktion  $\varrho$  die Randbedingungen

$$\lim_{x \rightarrow a} \sigma(x) \varrho(x) x^k = 0 \text{ und } \lim_{x \rightarrow b} \sigma(x) \varrho(x) x^k = 0 \text{ für alle } k \in \mathbb{N}_0, \quad (3.10)$$

so sind die zugehörigen Polynome vom hypergeometrischen Typ orthogonal auf dem Intervall  $(a, b)$  bezüglich  $\varrho$ . Sie heißen **klassische orthogonale Polynome** (vgl., z.B., [60, S. 7]). Dies sind die Jacobi-Polynome, Hermite-Polynome und Laguerre-Polynome. Sie sind, bis auf lineare Transformationen des Arguments  $x$ , die einzigen Klassen von orthogonalen Polynomen, welche eine Gleichung vom hypergeometrischen Typ lösen (vgl., z.B., [60, S. 8-9]). Wir interessieren uns besonders für die Jacobi-Polynome und behandeln diese in Unterabschnitt 3.1.2. Bevor wir jedoch zu den Jacobi-Polynomen kommen, betrachten wir in der folgenden Anmerkung eine explizite Formel zur Bestimmung von Polynomen vom hypergeometrischen Typ:

**Anmerkung 3.6.** Die Polynome vom hypergeometrischen Typ erfüllen die sogenannte **Rodrigues-Formel**. Für ein Polynom  $p_n$  vom hypergeometrischen Typ  $n$ -ten Grades gilt

$$p_n(x) = \frac{B_n}{\varrho(x)} \frac{d^n}{dx^n} [\sigma^n(x) \varrho(x)], \quad (3.11)$$

wobei  $B_n$  konstant ist (vgl., z.B., [60, S. 4]).

### 3.1.2 Jacobi-Polynome

Die Jacobi-Polynome  $P_n \equiv P_n^{\alpha, \beta}$  sind klassische orthogonale Polynome auf dem Intervall  $I = [-1, 1]$  vom Grad  $n$ . Sie können mit Hilfe der hypergeometrischen Funktion folgendermaßen definiert werden:

**Definition 3.7** (vgl., z.B., [47, S. 216]). Seien  $\alpha, \beta > -1$ . Dann heißen die durch

$$\begin{aligned} P_n^{\alpha, \beta}(x) &= \frac{(\alpha + 1)_n}{n!} {}_2F_1 \left( \begin{matrix} -n, n + \alpha + \beta + 1 \\ \alpha + 1 \end{matrix}; \frac{1-x}{2} \right) \\ &= \frac{(\alpha + 1)_n}{n!} \sum_{k=0}^n \frac{(-n)_k (n + \alpha + \beta + 1)_k (1-x)^k}{(\alpha + 1)_k 2^k k!}, \end{aligned} \quad (3.12)$$

für alle  $n \in \mathbb{N}_0$  definierten Polynome  $P_n \equiv P_n^{\alpha, \beta}$ , **Jacobi-Polynome**.

Die Jacobi-Polynome  $P_n^{\alpha,\beta}$  sind orthogonal auf dem Intervall  $I = [-1, 1]$  bezüglich des Skalarprodukts

$$(f, g)_\varrho := \int_{-1}^1 f(x)g(x)\varrho(x)dx, \quad (3.13)$$

wobei

$$\varrho(x) := (1-x)^\alpha(1+x)^\beta \quad (3.14)$$

die Gewichtsfunktion ist. Sie sind normiert durch

$$\left(P_n^{\alpha,\beta}, P_n^{\alpha,\beta}\right)_\varrho = \frac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)n!\Gamma(n+\alpha+\beta+1)} \quad (3.15)$$

(vgl., z.B., [47, S. 217]).

Wir fassen im Folgenden einige Eigenschaften der Jacobi-Polynome zusammen (vgl., z.B., [47, S. 217-218]):

- Die Jacobi-Polynome erfüllen die folgende Rekursionsgleichung

$$xP_n^{\alpha,\beta}(x) = \alpha_n P_{n+1}^{\alpha,\beta}(x) + \beta_n P_n^{\alpha,\beta}(x) + \gamma_n P_{n-1}^{\alpha,\beta}(x), \quad \text{für } x \in [-1, 1] \quad (3.16)$$

mit den Konstanten

$$\alpha_n = \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}, \quad (3.17)$$

$$\beta_n = \frac{\beta^2 - \alpha^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)}, \quad (3.18)$$

$$\gamma_n = \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)}. \quad (3.19)$$

- Die Jacobi-Polynome  $y(x) = P_n^{\alpha,\beta}(x)$  sind Lösungen folgender Gleichung vom hypergeometrischen Typ

$$(1-x^2)y''(x) + [\beta - \alpha - (\alpha + \beta + 2)x]y'(x) + n(n + \alpha + \beta + 1)y(x) = 0. \quad (3.20)$$

- Die Jacobi-Polynome erfüllen folgende Rodrigues-Formel

$$(1-x)^\alpha(1+x)^\beta P_n^{\alpha,\beta}(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n} \left[ (1-x)^{n+\alpha}(1+x)^{n+\beta} \right]. \quad (3.21)$$

- Die Jacobi-Polynome genügen der Gleichung

$$\frac{d}{dx} P_n^{\alpha,\beta}(x) = \frac{n+\alpha+\beta+1}{2} P_{n-1}^{\alpha+1,\beta+1}(x). \quad (3.22)$$

Darüber hinaus sind die Jacobi-Polynome auf dem Intervall  $[-1, 1]$  folgendermaßen beschränkt:

**Satz 3.8** (vgl., z.B., [2, S. 786]). *Seien  $\alpha, \beta > -1$  mit  $\max\{\alpha, \beta\} \geq -\frac{1}{2}$ . Dann gilt*

$$\max_{x \in [-1, 1]} |P_n^{\alpha, \beta}(x)| = \binom{n + \max\{\alpha, \beta\}}{n}. \quad (3.23)$$

Wichtige Spezialfälle der Jacobi-Polynome sind für  $\alpha = \beta$  die Gegenbauer-Polynome, die auch ultrasphärische Polynome genannt werden, für  $\alpha = \beta = -\frac{1}{2}$  die Tschebyscheff-Polynome erster Art, für  $\alpha = \beta = \frac{1}{2}$  die Tschebyscheff-Polynome zweiter Art und für  $\alpha = \beta = 0$  die Legendre-Polynome. Aus historischen Gründen werden hier andere Normierungen verwendet:

**Definition 3.9** (vgl., z.B., [73, S. 59-60]). • *Sei  $\lambda > -\frac{1}{2}$ ,  $\lambda \neq 0$ . Dann heißen die durch*

$$P_n^\lambda(x) := \frac{\Gamma\left(\lambda + \frac{1}{2}\right) \Gamma(n + 2\lambda)}{\Gamma(2\lambda) \Gamma\left(n + \lambda + \frac{1}{2}\right)} P_n^{\lambda - \frac{1}{2}, \lambda - \frac{1}{2}}(x), \quad (3.24)$$

für alle  $n \in \mathbb{N}_0$  definierten Polynome  $P_n^\lambda$ , **Gegenbauer-Polynome**.

- *Die durch*

$$T_n(x) := \frac{P_n^{-\frac{1}{2}, -\frac{1}{2}}(x)}{P_n^{-\frac{1}{2}, -\frac{1}{2}}(1)}, \quad (3.25)$$

für alle  $n \in \mathbb{N}_0$  definierten Polynome  $T_n$ , heißen **Tschebyscheff-Polynome erster Art**.

- *Die durch*

$$U_n(x) := (n + 1) \frac{P_n^{\frac{1}{2}, \frac{1}{2}}(x)}{P_n^{\frac{1}{2}, \frac{1}{2}}(1)}, \quad (3.26)$$

für alle  $n \in \mathbb{N}_0$  definierten Polynome  $U_n$ , heißen **Tschebyscheff-Polynome zweiter Art**.

- *Die durch*

$$P_n(x) := P_n^{0, 0}(x), \quad (3.27)$$

für alle  $n \in \mathbb{N}_0$  definierten Polynome  $P_n$ , heißen **Legendre-Polynome**.



## 3.2 Diskrete orthogonale Polynome

### 3.2.1 Definition und Eigenschaften

**Definition 3.10** (vgl., z.B., [47, S. 1]). Sei  $X \subset \mathbb{R}$  eine nichtleere, höchstens abzählbare Teilmenge der reellen Zahlen. Falls  $X$  endlich ist, setze  $N := |X| - 1$ . Weiter seien positive Werte  $\omega_x > 0$  für alle  $x \in X$  gegeben, sodass

$$0 < \sum_{x \in X} x^{2n} \omega_x < \infty \text{ für alle } n \in \mathbb{N}_0 \quad (3.28)$$

gilt.

- Falls  $X$  endlich ist, heißt ein System  $\{p_n\}_{n=0}^N$  von Polynomen, wobei für jedes  $n = 0, \dots, N$  das Polynom  $p_n$  den Polynomgrad  $n$  hat, **orthogonal** auf  $X$  bezüglich der  $\omega_x$ ,  $x \in X$ , falls

$$\sum_{x \in X} p_n(x)p_m(x)\omega_x = 0 \text{ für alle } n, m = 0, \dots, N, \text{ mit } n \neq m \quad (3.29)$$

gilt.

- Falls  $X$  nicht endlich ist, heißt ein System  $\{p_n\}_{n \in \mathbb{N}_0}$  von Polynomen, wobei für jedes  $n \in \mathbb{N}_0$  das Polynom  $p_n$  den Polynomgrad  $n$  hat, **orthogonal** auf  $X$  bezüglich der  $\omega_x$ ,  $x \in X$ , falls

$$\sum_{x \in X} p_n(x)p_m(x)\omega_x = 0 \text{ für alle } n, m \in \mathbb{N}_0, \text{ mit } n \neq m \quad (3.30)$$

gilt.

Die Werte  $\omega_x$  heißen dann **Gewichte**.

**Anmerkung 3.11.** Bei Systemen  $\{p_n\}_{n \in \mathbb{N}_0}$  beziehungsweise  $\{p_n\}_{n=0}^N$ , die die Voraussetzungen aus Definition 3.10 erfüllen, sprechen wir, zur Unterscheidung von den schon definierten stetigen orthogonalen Polynomen, auch von **diskreten orthogonalen Polynomen** (vgl., z.B., [47, S. 2]).

**Anmerkung 3.12.** In diesem Zusammenhang sprechen wir auch von Orthogonalität bezüglich des durch

$$\langle f, g \rangle := \sum_{x \in X} f(x)g(x)\omega_x \quad (3.31)$$

definierten Skalarprodukts.

Durch die Gleichungen (3.29) und (3.30) werden die jeweiligen Systeme von orthogonalen Polynomen bis auf konstante Faktoren festgelegt (vgl., z.B., [47, S. 2]).

Wir wollen in der folgenden Anmerkung einige für uns wichtige Eigenschaften von diskreten orthogonalen Polynomen zusammenfassen, die analog auch schon die stetigen orthogonalen Polynome erfüllt haben.

**Anmerkung 3.13.** Sei  $X \subset \mathbb{R}$  eine nichtleere, höchstens abzählbare Teilmenge der reellen Zahlen. Falls  $X$  endlich ist, setze  $N := |X| - 1$  und es sei  $\mathcal{N} = \{0, 1, \dots, N\}$ , ansonsten sei  $\mathcal{N} = \mathbb{N}_0$ . Weiter sei ein System  $\{p_n\}_{n \in \mathcal{N}}$  aus orthogonalen Polynomen gegeben, welches orthogonal auf  $X$  bezüglich des Skalarprodukts

$$\langle f, g \rangle := \sum_{x \in X} f(x)g(x)\omega_x \quad (3.32)$$

ist. Dann gelten folgende Eigenschaften (vgl., z.B., [24], [79]):

- Jedes Polynom  $q$   $n$ -ten Grades, mit  $n \leq N$  lässt sich eindeutig darstellen durch

$$q = \sum_{k=0}^n \frac{\langle q, p_k \rangle}{\langle p_k, p_k \rangle} p_k. \quad (3.33)$$

- Besteht  $X$  aus isolierten Punkten, erfüllen die orthogonalen Polynome  $p_n$  eine Rekursionsgleichung der Form

$$xp_n(x) = \alpha_n p_{n+1}(x) + \beta_n p_n(x) + \gamma_n p_{n-1}(x), \quad (3.34)$$

wobei  $\alpha_n$ ,  $\beta_n$  und  $\gamma_n$  Konstanten sind.

- Die Nullstellen der Polynome  $p_n$  sind alle einfach und liegen im Intervall  $[\inf X, \max X]$ .

Wir beschäftigen uns nun mit diskreten orthogonalen Polynomen auf einem äquidistanten Gitter, d. h. wir betrachten den Fall  $X \subset \mathbb{Z}$ . In diesem Fall lassen sich viele wichtige Eigenschaften der stetigen orthogonalen Polynome übertragen. Zunächst benötigen wir die Differenzenoperatoren:

$$\Delta f(x) = f(x+1) - f(x), \quad (3.35)$$

$$\nabla f(x) = f(x) - f(x-1). \quad (3.36)$$

Damit können wir eine diskrete Formulierung der hypergeometrischen Gleichung (3.7) definieren:

**Definition 3.14** (vgl., z.B., [60, S. 19]). Eine Differenzgleichung der Form

$$\sigma(x)\Delta\nabla y(x) + \tau(x)\Delta y(x) + \lambda y(x) = 0, \quad (3.37)$$

wobei  $\sigma$  ein Polynom höchstens zweiten Grades,  $\tau$  ein Polynom höchstens ersten Grades und  $\lambda$  eine Konstante ist, heißt **Differenzgleichung vom hypergeometrischen Typ**.

Mit den Identitäten

$$\sigma(x) \equiv \tilde{\sigma}(x) - \frac{1}{2}\tilde{\tau}(x) \text{ und } \tau(x) \equiv \tilde{\tau}(x) \quad (3.38)$$

ist Gleichung (3.37) eine Diskretisierung der Gleichung vom hypergeometrischen Typ

$$\tilde{\sigma}y'' + \tilde{\tau}y' + \lambda y = 0, \quad (3.39)$$

(vgl., z.B., [47, S. 19]).

Wie schon im kontinuierlichen Fall lässt sich auch hier Gleichung (3.37) durch Multiplikation einer Funktion  $\varrho$  in folgende selbstadjungierte Form bringen:

$$\Delta [\sigma(x)\varrho(x)\nabla y(x)] + \lambda\varrho(x)y(x) = 0, \quad (3.40)$$

wobei  $\varrho$  der Differenzengleichung

$$\Delta [\sigma(x)\varrho(x)] = \tau(x)\varrho(x) \quad (3.41)$$

genügen muss. Die einfachsten Lösungen dieser Gleichungen sind Polynome. Wir betrachten die Lösungen auf einem Intervall  $(a, b)$  mit  $a, b \in \mathbb{Z} \cup \{\infty\}$ . Erfüllt die Funktion  $\varrho$  die Randbedingungen

$$\lim_{x \rightarrow a} \sigma(x)\varrho(x)x^k = 0 \text{ und } \lim_{x \rightarrow b} \sigma(x)\varrho(x)x^k = 0 \text{ für alle } k \in \mathbb{N}_0, \quad (3.42)$$

so sind die zugehörigen Polynome, die Gleichung (3.37) lösen, orthogonal auf dem Intervall  $(a, b - 1)$  bezüglich der  $\varrho(i)$ ,  $a \leq i \leq b - 1$ . Sie heißen **klassische diskrete orthogonale Polynome** (vgl., z.B., [60, S. 27]). Dies sind die Hahn-Polynome, Meixner-Polynome, Kravchuk-Polynome und Charlier-Polynome (vgl., z.B., [60, S. 30-39]). Wir interessieren uns besonders für die Hahn-Polynome und behandeln diese in Unterabschnitt 3.2.2. Sie lassen sich als Diskretisierung der Jacobi-Polynome auffassen. Bevor wir jedoch zu den Hahn-Polynomen kommen, betrachten wir in der folgenden Anmerkung die diskrete Rodrigues-Formel:

**Anmerkung 3.15.** *Polynome, die Gleichung (3.37) lösen, erfüllen die sogenannte **diskrete Rodrigues-Formel**. Für ein solches Polynom  $p_n$   $n$ -ten Grades gilt*

$$p_n(x) = \frac{B_n}{\varrho(x)} \Delta^n \left[ \varrho(x) \prod_{k=0}^{n-1} \sigma(x - k) \right], \quad (3.43)$$

wobei  $B_n$  konstant ist (vgl., z.B., [60, S. 24]).

### 3.2.2 Hahn-Polynome

Die Hahn-Polynome  $Q_n \equiv Q_n(\cdot; \alpha, \beta, N)$  sind klassische diskrete orthogonale Polynome auf dem Intervall  $I = [0, N]$  vom Grad  $n$ . Sie können mit Hilfe der hypergeometrischen Funktion folgendermaßen definiert werden:

**Definition 3.16** (vgl., z.B., [47, S. 204]). *Seien  $\alpha, \beta > -1$  und  $N \in \mathbb{N}_0$ . Dann heißen die durch*

$$\begin{aligned} Q_n(x; \alpha, \beta, N) &= {}_3F_2 \left( \begin{matrix} -n, & n + \alpha + \beta + 1, & -x \\ & \alpha + 1, & -N \end{matrix}; 1 \right) \\ &= \sum_{k=0}^n \frac{(-n)_k (n + \alpha + \beta + 1)_k (-x)_k}{(\alpha + 1)_k (-N)_k} \frac{1}{k!}, \end{aligned} \quad (3.44)$$

für alle  $n = 0, \dots, N$  definierten Polynome  $Q_n \equiv Q_n(\cdot; \alpha, \beta, N)$ , **Hahn-Polynome**.

Die Hahn-Polynome sind orthogonal auf  $I$  bezüglich des Skalarprodukts

$$\langle f, g \rangle_\omega := \sum_{i=0}^N f(i)g(i)\omega(i), \quad (3.45)$$

wobei die Gewichtsfunktion  $\omega$  gegeben ist durch

$$\omega(x) := \binom{\alpha + x}{x} \binom{\beta + N - x}{N - x}. \quad (3.46)$$

Sie sind durch

$$\langle Q_n(\cdot; \alpha, \beta, N), Q_n(\cdot; \alpha, \beta, N) \rangle_\omega = \frac{(-1)^n (n + \alpha + \beta + 1)_{N+1} (\beta + 1)_n n!}{(2n + \alpha + \beta + 1) (\alpha + 1)_n (-N)_n N!} \quad (3.47)$$

normiert (vgl., z.B., [47, S. 204]).

Wir fassen im Folgenden einige Eigenschaften der Hahn-Polynome zusammen (vgl., z.B., [47, S. 204-206]):

- Die Hahn-Polynome erfüllen die folgende Rekursionsgleichung

$$-xQ_n(x) = A_n Q_{n+1}(x) - (A_n + C_n) Q_n(x) + C_n Q_{n-1}(x), \quad \text{für } x \in [0, N] \quad (3.48)$$

mit den Konstanten

$$A_n = \frac{(n + \alpha + \beta + 1)(n + \alpha + 1)(N - n)}{(2n + \alpha + \beta + 1)(2n + \alpha + \beta + 2)}, \quad (3.49)$$

$$C_n = \frac{n(n + \alpha + \beta + N + 1)(n + \beta)}{(2n + \alpha + \beta)(2n + \alpha + \beta + 1)}. \quad (3.50)$$

- Die Hahn-Polynome  $y(x) = Q_n(x)$  sind Lösungen folgender Differenzgleichung vom hypergeometrischen Typ

$$n(n + \alpha + \beta + 1)y(x) = B(x)y(x + 1) - [B(x) + D(x)]y(x) + D(x)y(x - 1), \quad (3.51)$$

mit

$$B(x) = (x + \alpha + 1)(x - N), \quad (3.52)$$

$$D(x) = x(x - \beta - N - 1). \quad (3.53)$$

- Die Hahn-Polynome erfüllen folgende diskrete Rodrigues-Formel

$$\omega(x)Q_n(x; \alpha, \beta, N) = \frac{(-1)^n (\beta + 1)_n}{(-N)_n} \nabla^n \left[ \binom{\alpha + n + x}{x} \binom{\beta + N - x}{N - n - x} \right]. \quad (3.54)$$

- Die Hahn-Polynome genügen der Gleichung

$$\Delta Q_n(x; \alpha, \beta, N) = -\frac{n(n + \alpha + \beta + 1)}{(\alpha + 1)N} Q_{n-1}(x; \alpha + 1, \beta + 1, N - 1). \quad (3.55)$$

Die Hahn-Polynome  $Q_n(\cdot; \alpha, \beta, N)$  können als Diskretisierung der Jacobi-Polynome  $P_n^{\alpha, \beta}$  verstanden werden. Denn es gilt für fest gewähltes  $n$  folgende Grenzwertbeziehung zwischen den Hahn-Polynomen und den Jacobi-Polynomen:

$$\lim_{N \rightarrow \infty} (-1)^n \binom{n + \alpha}{n} Q_n\left(\frac{N}{2}(1 + x); \alpha, \beta, N\right) = P_n^{\beta, \alpha}(x), \quad (3.56)$$

für jedes  $x \in [-1, 1]$  (vgl., z.B., [60, S. 45]).

## 4 Approximation mit Polynomen

In diesem Kapitel beschäftigen wir uns mit der Approximation von Funktionen mit Polynomen. Wir definieren für jedes  $n \in \mathbb{N}_0$  die Menge der Polynome höchstens  $n$ -ten Grades durch

$$\mathcal{P}_n := \left\{ p \equiv p(x) = \sum_{\mu=0}^n a_\mu x^\mu : a_\mu \in \mathbb{R}, \mu = 0, \dots, n \right\}. \quad (4.1)$$

Polynome lassen sich nur durch die elementaren Operationen Addition und Multiplikation sehr einfach berechnen. Sie werden daher seit langem zur Approximation verwendet. Eine Begründung, Funktionen durch Polynome anzunähern, ist der folgende bekannte **Weierstraßsche Approximationssatz**:

**Satz 4.1** (vgl., z.B., [63, S. 87]). *Sei eine stetige Funktion  $f \in \mathcal{C}[a, b]$  gegeben. Dann existiert zu jedem  $\varepsilon > 0$  ein Polynom  $p$ , sodass für alle  $x \in [a, b]$*

$$|f(x) - p(x)| < \varepsilon \quad (4.2)$$

*gilt.*

Dieser Satz geht auf die Arbeit [76] aus dem Jahr 1885 von K. Weierstraß zurück.

In den folgenden Abschnitten stellen wir einige häufig verwendete Approximationsmethoden vor, die auf Polynomen beruhen. Wir fassen grundlegende Eigenschaften dieser Methoden zusammen und stellen die jeweiligen Vor- und Nachteile gegenüber. Dies ermöglicht uns, dass wir später in Kapitel 7 die Approximationseigenschaften der vorgestellten Methoden (Polynominterpolation, Approximation mit Bernsteinpolynomen, Polynome bester Approximation), mit unseren Ergebnissen zur Methode der kleinsten Quadrate vergleichen können.

### 4.1 Polynominterpolation

In diesem Abschnitt stellen wir alle für uns wichtigen Grundbegriffe und Aussagen über die Interpolation mit Polynomen zusammen. Das zugehörige Interpolationsproblem ist folgendermaßen definiert:

**Definition 4.2** (vgl., z.B., [69, S. 51]). *Sei eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  gegeben. Weiter seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, n$ . Gesucht ist ein Polynom  $p_n \in \mathcal{P}_n$ , sodass  $p_n(x_\mu) = f(x_\mu)$  für jedes  $\mu = 0, \dots, n$  gilt.*

Hier stellen sich die Fragen nach der Existenz und Eindeutigkeit, welche sich mit dem folgenden bekannten Satz sofort beantworten lassen:

**Satz 4.3** (vgl., z.B., [69, S. 52]). *Sei eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  gegeben. Weiter seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, n$ . Dann existiert ein eindeutig bestimmtes Polynom  $p_n \in \mathcal{P}_n$ , sodass  $p_n(x_\mu) = f(x_\mu)$  für jedes  $\mu = 0, \dots, n$  gilt.*

**Definition 4.4** (vgl., z.B., [69, S. 52]). *Sei eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  gegeben. Weiter seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, n$ . Dann heißt das eindeutig bestimmte Polynom  $p_n \in \mathcal{P}_n$  aus Satz 4.3, **Interpolationspolynom** zu  $\{(x_\mu, f(x_\mu))\}_{\mu=0}^n$ .*

Sind eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  sowie paarweise verschiedene Stützstellen  $x_\mu \in [a, b]$  gegeben, so lässt sich das zugehörige Interpolationspolynom

$$p_n(x) = \sum_{\mu=0}^n a_\mu x^\mu \quad (4.3)$$

durch Lösen des Gleichungssystems

$$p_n(x_\mu) = f(x_\mu), \quad \mu = 0, \dots, n \quad (4.4)$$

bestimmen (vgl., z.B., [69, S. 52]). (4.4) ist ein lineares Gleichungssystem, welches sich mit Hilfe der Vandermonde-Matrix durch

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{pmatrix} \quad (4.5)$$

darstellen lässt. Unter der Voraussetzung paarweiser verschiedener Stützstellen  $x_\mu$ , hat die Vandermonde-Matrix vollen Rang und (4.5) lässt sich eindeutig lösen (vgl., z.B., [69, S. 28,52]).

Weitere bekannte Möglichkeiten das Interpolationspolynom zu bestimmen, sind die Lagrange-Interpolation und die Newton-Interpolation (vgl., z.B., [71, S. 95-98]). Im Folgenden beschreiben wir kurz die Lagrange-Interpolation:

Dazu seien für jedes  $\nu = 0, \dots, n$  die sogenannten **Lagrange-Polynome** definiert:

$$L_{\nu,n}(x) = \prod_{\substack{\mu=0 \\ \mu \neq \nu}}^n \frac{x - x_\mu}{x_\nu - x_\mu}. \quad (4.6)$$

Es gilt für alle  $\mu, \nu = 0, \dots, n$

$$L_{\nu,n}(x_\mu) = \delta_{\mu\nu} = \begin{cases} 1 & \text{für } \mu = \nu, \\ 0 & \text{für } \mu \neq \nu. \end{cases} \quad (4.7)$$

Das Lagrange-Polynom  $L_{\nu,n}$  verschwindet also für  $\mu \neq \nu$  an den Stützstellen  $x_\mu$  und hat an der Stelle  $x_\nu$  den Wert Eins. Dadurch ergibt sich folgende Darstellung für das Interpolationspolynom (vgl., z.B., [71, S. 95]):

$$p_n(x) = \sum_{\mu=0}^n f(x_\mu) L_{\mu,n}(x). \quad (4.8)$$

Diese Darstellung des Interpolationspolynoms wird häufig für theoretische Untersuchungen verwendet. Für große Werte von  $n$  ist diese Methode numerisch instabil. Zudem hat die Lagrange-Interpolation den Nachteil, dass bei Hinzunahme weiterer Stützstellen die Berechnung in wesentlichen Teilen wiederholt werden muss. In der Praxis wird daher meist die Newton-Interpolation verwendet. Diese Berechnungsmethode basiert auf der Grundidee, bei Hinzunahme einer weiteren Stützstelle, einen entsprechenden Korrekturterm zum vorher berechneten Interpolationspolynom zu addieren (vgl., z.B., [35, S. 93-97], [69, S. 54-59], [71, S. 94-98]).

Der folgende Satz von G. Faber aus dem Jahr 1914 beschäftigt sich mit der Frage der Approximierbarkeit stetiger Funktionen durch Interpolationspolynome:

**Satz 4.5** (vgl., z.B., [48, S. 277]). *Zu jeder Folge  $(X_n)_{n \in \mathbb{N}_0}$ , bestehend aus beliebigen Verteilungen von Stützstellen*

$$X_n = \{x_{\mu,n} \in [a, b] : \mu = 0, 1, \dots, n\}, \quad (4.9)$$

*existiert eine stetige Funktion  $f \in \mathcal{C}[a, b]$ , sodass die zugehörige Folge von Interpolationspolynomen  $p_n \in \mathcal{P}_n$  auf dem Intervall  $[a, b]$  nicht gleichmäßig gegen  $f$  konvergiert.*

Es existiert also keine Verteilung von Stützstellen gemäß (4.9), sodass für alle stetigen Funktionen die gleichmäßige Konvergenz der zugehörigen Folge von Interpolationspolynomen garantiert werden kann.

Wir sind besonders an Fehlerabschätzungen interessiert, um sie mit unseren neuen Abschätzungen bei der Methode der kleinsten Quadrate in Kapitel 7 vergleichen zu können. Dazu betrachten wir die folgende bekannte Fehlerdarstellung:

**Satz 4.6** (vgl., z.B., [35, S. 78]). *Sei eine Funktion  $f \in \mathcal{C}^{n+1}[a, b]$  gegeben. Weiter seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, n$  und sei  $p_n \in \mathcal{P}_n$  das zugehörige Interpolationspolynom. Dann existiert für jedes  $x \in [a, b]$  ein  $\xi_x \in [a, b]$ , sodass*

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{\mu=0}^n (x - x_\mu). \quad (4.10)$$

Entscheidend für die Qualität der Approximation in der Klasse der  $(n+1)$ -mal stetig differenzierbaren Funktionen ist also die Wahl der Stützstellen, wie man dem Term

$$\prod_{\mu=0}^n (x - x_\mu) \quad (4.11)$$

aus Gleichung (4.10) entnimmt.

Insbesondere lässt sich Gleichung (4.10) für die Funktion  $f(x) = x^{n+1}$  anwenden. Mit  $f^{(n+1)}(x) = (n+1)!$  ergibt sich in diesem Fall

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{\mu=0}^n (x - x_\mu) = \prod_{\mu=0}^n (x - x_\mu) \quad (4.12)$$

für jedes  $x \in [a, b]$ . Man erhält daraus unmittelbar die Unverbesserbarkeit der folgenden, aus Gleichung (4.10) resultierenden Abschätzung:

**Korollar 4.7** (vgl., z.B., [35, S. 78]). *Sei eine Funktion  $f \in \mathcal{C}^{n+1}[a, b]$  gegeben. Weiter seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, n$  und sei  $p_n \in \mathcal{P}_n$  das zugehörige Interpolationspolynom. Des Weiteren sei*

$$I_n := \frac{1}{(n+1)!} \sup_{x \in [a, b]} \left| \prod_{\mu=0}^n (x - x_\mu) \right|. \quad (4.13)$$

Dann gilt

$$\sup_{x \in [a, b]} |f(x) - p_n(x)| \leq I_n \sup_{x \in [a, b]} |f^{(n+1)}(x)|. \quad (4.14)$$

*Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $I_n$  in Ungleichung (4.14) durch keine kleinere Zahl ersetzt werden kann.*

Die Frage, mit welcher Stützstellenwahl der Wert  $I_n$  minimiert werden kann, wird in Unterabschnitt 4.1.2 bei der Untersuchung der Interpolation in den Nullstellen der Tschebyscheff-Polynome behandelt.

### 4.1.1 Polynominterpolation in äquidistanten Stützstellen

In diesem Unterabschnitt beschäftigen wir uns mit der Polynominterpolation unter Verwendung von  $n+1$  äquidistanten Stützstellen. Diese sind im Intervall  $[a, b]$  gegeben durch die Menge

$$X_n^A = \left\{ x_\mu = a + \mu \frac{b-a}{n} : \mu = 0, 1, \dots, n \right\}. \quad (4.15)$$

Die Abstände zwischen zwei benachbarten Stützstellen sind also immer gleich. Dies macht sie für Rechnungen besonders handlich und damit sind sie auch für Anwendungen sehr interessant. Leider treten jedoch bei der Polynominterpolation in äquidistanten Stützstellen Probleme auf, die so erheblich sind, dass die Polynominterpolation in diesen Stützstellen als Approximationsmethode für sehr große Stützstellenanzahlen kaum verwendbar ist. Welche Probleme auftreten wollen wir nun in Hinblick auf unsere späteren Untersuchungen diskutieren.

Als einer der Ersten ist C. Runge im Jahr 1901 für äquidistante Stützstellen der Frage nachgegangen, wie sich die Interpolationspolynome  $p_n$  vom Grad  $n$  verhalten, wenn  $n$  ansteigt (vgl. [68]). Er hat festgestellt, dass die Interpolationspolynome in der Nähe der Intervallgrenzen sehr stark oszillieren können. Dieses Verhalten wurde nach ihm benannt und wird als *Runges Phänomen* bezeichnet. Dazu betrachten wir die, in diesem Zusammenhang ebenfalls nach ihm benannte, *Runge-Funktion*

$$f(x) := \frac{1}{1 + 25x^2}, \quad x \in [-1, 1] \tag{4.16}$$

(vgl., z.B., [30]). Also gilt  $f \in C^\infty[-1, 1]$  und zudem lässt sich diese Funktion auf eine in  $\mathbb{C}$  offene Menge  $D \subset \mathbb{C}$ , mit  $[-1, 1] \subset D$  und  $\frac{-i}{5}, \frac{i}{5} \notin D$  analytisch durch

$$\tilde{f}(z) := \frac{1}{1 + 25z^2}, \quad z \in D \tag{4.17}$$

fortsetzen. Trotzdem konvergiert die Folge  $(p_n)_n$  der zugehörigen Interpolationspolynome in äquidistanten Stützstellen nicht gleichmäßig gegen  $f$ , d. h. es gilt

$$\limsup_{n \rightarrow \infty} \sup_{x \in [-1, 1]} |p_n(x) - f(x)| \neq 0 \tag{4.18}$$

(vgl., z.B., [30]). Genauer gilt Folgendes (vgl., z.B., [35, S. 84]):

$$\lim_{n \rightarrow \infty} |p_n(x) - f(x)| = \begin{cases} 0 & \text{für } |x| < 0.7266, \\ \infty & \text{für } |x| > 0.7267. \end{cases} \tag{4.19}$$

In Abbildung 4.1 sind die Runge-Funktion  $f$  sowie die zugehörigen Interpolationspolynome in äquidistanten Stützstellen für  $n = 10$  und  $n = 20$  dargestellt.

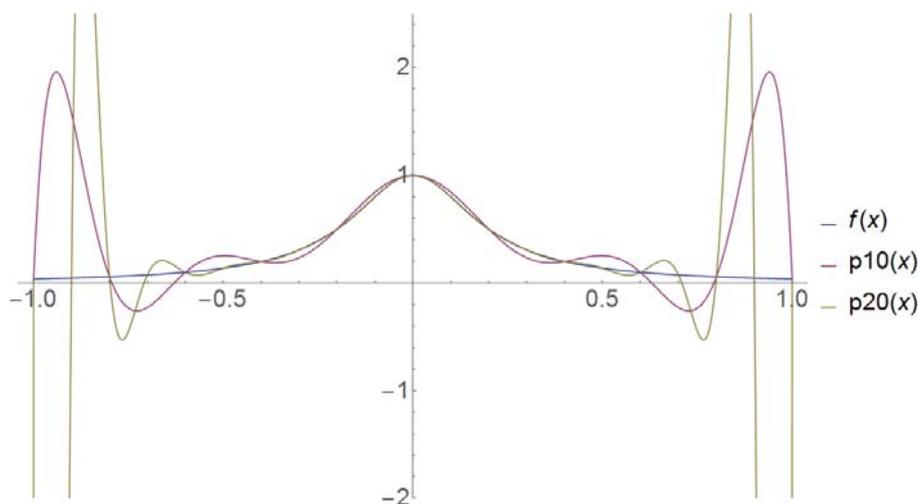


Abbildung 4.1: Die Runge-Funktion - Interpolation in äquidistanten Stützstellen.

Dieses Beispiel zeigt uns, dass man selbst für die Klasse der analytischen Funktionen im Allgemeinen keine gleichmäßige Konvergenz bei der Polynominterpolation in äquidistanten Stützstellen garantieren kann.

Ebenfalls ungünstig ist das Ergebnis bei der Betragsfunktion

$$g(x) := |x|, \quad x \in [-1, 1]. \quad (4.20)$$

Im Jahr 1912 hat S. Bernstein gezeigt, dass die zugehörige Folge von Interpolationspolynomen  $p_n$  in äquidistanten Stützstellen in allen Punkten  $x \in (-1, 1)$ ,  $x \neq 0$  divergiert. Genauer zeigte er, dass für jedes  $x \in (-1, 1)$ ,  $x \neq 0$  gilt:

$$\lim_{n \rightarrow \infty} |p_n(x) - g(x)| = \infty \quad (4.21)$$

(vgl. [13]). In Abbildung 4.2 sind die Betragsfunktion  $g$  sowie die zugehörigen Interpolationspolynome in äquidistanten Stützstellen für  $n = 10$  und  $n = 20$  dargestellt.

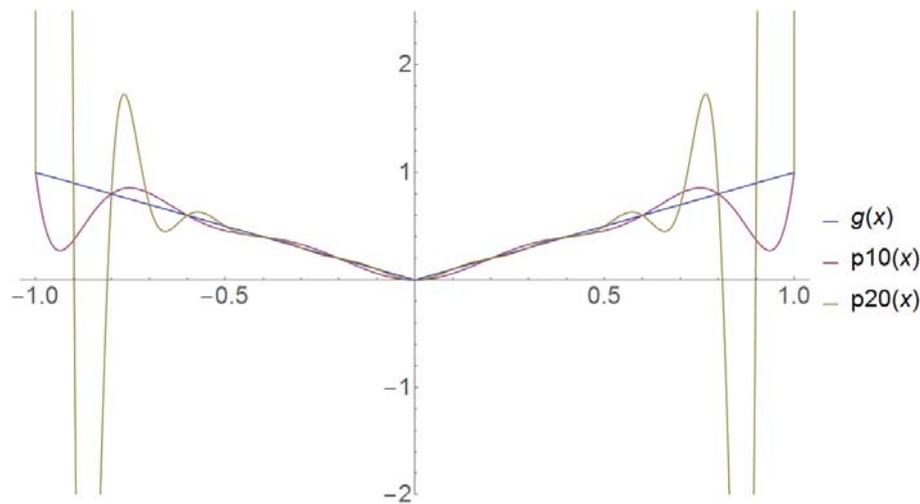


Abbildung 4.2: Die Betragsfunktion - Interpolation in äquidistanten Stützstellen.

### 4.1.2 Polynominterpolation in Tschebyscheff-Stützstellen

In diesem Unterabschnitt beschäftigen wir uns mit der Polynominterpolation in Tschebyscheff-Stützstellen:

**Anmerkung 4.8** (vgl., z.B., [35, S. 87]). Für jedes  $n \in \mathbb{N}$  sind die Nullstellen der Tschebyscheff-Polynome erster Art

$$T_n(x) = \frac{P_n^{-\frac{1}{2}, -\frac{1}{2}}(x)}{P_n^{-\frac{1}{2}, -\frac{1}{2}}(1)} \quad (4.22)$$

aus Definition 3.9, gegeben durch die Menge

$$X_n^T = \left\{ x_k^{(n)} = \cos\left(\frac{2k-1}{2n}\pi\right) : k = 1, 2, \dots, n \right\}. \quad (4.23)$$

Die Werte  $x_k^{(n)} = \cos\left(\frac{2k-1}{2n}\pi\right)$ ,  $k = 1, 2, \dots, n$  heißen **Tschebyscheff-Stützstellen**. Weiterhin erfüllen die Tschebyscheff-Polynome die folgende Rekursionsgleichung

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \text{ für } n \in \mathbb{N}, \quad (4.24)$$

mit

$$T_0(x) = 1 \text{ und } T_1(x) = x. \quad (4.25)$$

Bevor wir wie angekündigt zur Fehlerabschätzung (4.14) kommen, betrachten wir zunächst das folgende von S. Bernstein im Jahr 1912 bewiesene Resultat. Dabei verwenden wir das Stetigkeitsmodul  $\omega$  gemäß Definition 2.6.

**Satz 4.9** (vgl. [13]). *Sei eine Funktion  $f \in \mathcal{C}[-1, 1]$  mit der Eigenschaft*

$$\lim_{\delta \rightarrow 0} \omega(f, [-1, 1], \delta) \ln(\delta) = 0 \quad (4.26)$$

gegeben. Weiter seien für alle  $n \in \mathbb{N}_0$  die Tschebyscheff-Stützstellen durch  $x_k^{(n+1)} = \cos\left(\frac{2k-1}{2n+2}\pi\right)$  für  $k = 1, 2, \dots, n+1$  sowie die zugehörigen Interpolationspolynome  $p_n \in \mathcal{P}_n$  gegeben. Dann konvergiert die Folge der Interpolationspolynome  $(p_n)_n$  auf dem Intervall  $[-1, 1]$  gleichmäßig gegen  $f$ .

Die Eigenschaft (4.26) wird auch als **Dini-Lipschitz-Bedingung** bezeichnet. Zum Vergleich: Offensichtlich gilt für eine Lipschitz-stetige Funktion  $f : [-1, 1] \rightarrow \mathbb{R}$ , mit der Lipschitz-Konstanten

$$L := \sup_{x, t \in [-1, 1]} \frac{|f(x) - f(t)|}{|x - t|} \quad (4.27)$$

(vgl., z.B., [71, S. 186]), die Abschätzung  $\omega(f, [-1, 1], \delta) \leq L\delta$ . Also ist für jede Lipschitz-stetige Funktion  $f$  und jede Nullfolge  $(\delta_n)_n$ ,  $\delta_n > 0$  die Folge

$$\left(\omega(f, [-1, 1], \delta_n) \frac{1}{\delta_n}\right)_n \quad (4.28)$$

beschränkt, womit  $f$  die Bedingung (4.26) erfüllt.

Insbesondere genügen offensichtlich die Runge-Funktion  $f(x) = \frac{1}{1+25x^2}$  und die Betragsfunktion  $g(x) = |x|$  der Bedingung (4.26), da sie Lipschitz-stetig sind. Damit konvergieren die jeweiligen Folgen der Interpolationspolynome nach Satz 4.9 gleichmäßig gegen  $f$  beziehungsweise  $g$ . Im Fall von äquidistanten Stützstellen sind die Folgen der Interpolationspolynome für die Funktionen  $f$  und  $g$  in einzelnen Teilintervallen divergent, siehe Unterabschnitt 4.1.1.

Wie schon oben angedeutet, betrachten wir nun die Ungleichung (4.14) beziehungsweise den Wert  $I_n$  aus (4.13). Zunächst lassen sich nach Anmerkung 4.8 die Tschebyscheff-Polynome  $T_n$  für alle  $n \in \mathbb{N}$  durch

$$T_n(x) = 2^{n-1} \prod_{k=1}^n (x - x_k^{(n)}) \quad (4.29)$$

darstellen. Insbesondere minimieren die Tschebyscheff-Stützstellen den Wert  $I_n$  aus (4.13) im Fall  $a = -1$  und  $b = 1$ :

**Satz 4.10** (vgl., z.B., [35, S. 88]). Seien  $x_\mu \in [-1, 1]$  paarweise verschiedene Stützstellen für  $\mu = 1, \dots, n$ . Weiter seien für  $k = 1, 2, \dots, n$  die Tschebyscheff-Stützstellen gegeben durch  $x_k^{(n)} = \cos\left(\frac{2k-1}{2n}\pi\right)$ . Dann gilt für jedes  $n \in \mathbb{N}$

$$\sup_{x \in [-1, 1]} \left| \prod_{\mu=1}^n (x - x_\mu) \right| \geq \sup_{x \in [-1, 1]} \left| \prod_{k=1}^n (x - x_k^{(n)}) \right| = \frac{1}{2^{n-1}} \sup_{x \in [-1, 1]} |T_n(x)| = \frac{1}{2^{n-1}}. \quad (4.30)$$

Die Gleichheit in (4.30) gilt nur für  $x_k = x_k^{(n)}$  für  $k = 1, \dots, n$ .

In Abbildung 4.3 ist das Tschebyscheff-Polynom  $T_{10}$  dargestellt.

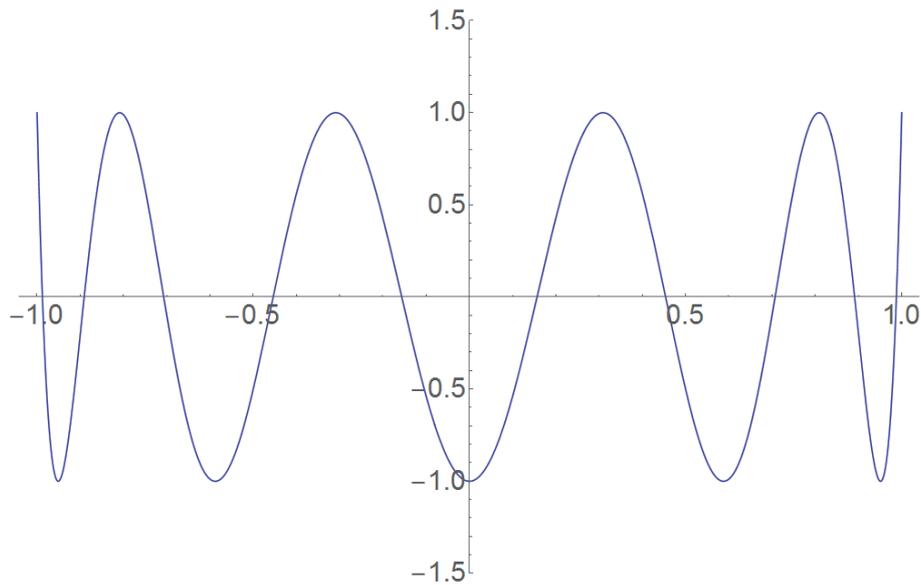


Abbildung 4.3: Das Tschebyscheff-Polynom  $T_{10}$ .

Damit sind die Tschebyscheff-Stützstellen  $x_k^{(n)}$  im Fall  $a = -1$  und  $b = 1$  gemäß unverbesserbarer Abschätzung (4.14) die optimalen Stützstellen bei der Polynominterpolation in der Klasse der  $(n + 1)$ -mal stetig differenzierbaren Funktionen. Ist  $[a, b] \neq [-1, 1]$ , so muss nur eine lineare Transformation zwischen den Intervallen  $[-1, 1]$  und  $[a, b]$  durchgeführt werden. Man erhält aus Korollar 4.7 und Satz 4.10 folgende bekannte Fehlerabschätzung:

**Korollar 4.11** (vgl., z.B., [35, S. 90]). Sei eine Funktion  $f \in \mathcal{C}^{n+1}[-1, 1]$  gegeben. Weiter seien für  $k = 1, 2, \dots, n + 1$  die Tschebyscheff-Stützstellen gegeben durch  $x_k^{(n+1)} = \cos\left(\frac{2k-1}{2n+2}\pi\right)$  und sei  $p_n \in \mathcal{P}_n$  das zugehörige Interpolationspolynom. Dann gilt für jedes  $n \in \mathbb{N}$

$$\sup_{x \in [-1, 1]} |f(x) - p_n(x)| \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{1}{(n+1)!2^n}. \quad (4.31)$$

Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $\frac{1}{(n+1)!2^n}$  in Ungleichung (4.31) durch keine kleinere Zahl ersetzt werden kann.

Die hier dargestellten Ergebnisse zur Interpolation in Tschebyscheff-Stützstellen sind ein erster Hinweis, warum diese Wahl von Stützstellen häufig in der Praxis angewendet wird. Allerdings setzt dies voraus, dass Messdaten gemäß dieser Stützstellenverteilung erhoben werden (können).

Zur Einordnung unserer Approximationsergebnisse bei der Methode der kleinsten Quadrate aus Kapitel 7 vergleichen wir diese Ergebnisse mit der Fehlerabschätzung aus dem vorherigen Korollar 4.11 in Abschnitt 7.4.

## 4.2 Approximation mit Bernsteinpolynomen

In diesem Abschnitt betrachten wir die Approximation mit Bernsteinpolynomen. Dies ist eine Approximationsmethode, die auf äquidistanten Stützstellen basiert. Im vorherigen Abschnitt 4.1 haben wir festgestellt, dass äquidistante Stützstellen bei der Polynominterpolation zu erheblichen Problemen führen können. Diese Probleme treten bei der Approximation mit Bernsteinpolynomen zwar nicht mehr auf, jedoch ist die Konvergenzordnung gering. Gewisse formerhaltende Eigenschaften sprechen allerdings für die Approximation mit Bernsteinpolynomen.

Wir vergleichen unsere Abschätzungen bei der Methode der kleinsten Quadrate mit Ergebnissen bei der Approximation mit Bernsteinpolynomen in Abschnitt 7.5.

**Definition 4.12** (vgl., z.B., [63, S. 247-248]). *Zu jedem  $n \in \mathbb{N}$  sei der Operator*

$$B_n : \mathcal{C}[0, 1] \rightarrow \mathcal{P}_n, \quad (4.32)$$

*definiert durch*

$$B_n[f](x) := \sum_{\nu=0}^n f\left(\frac{\nu}{n}\right) \binom{n}{\nu} x^\nu (1-x)^{n-\nu}, \quad f \in \mathcal{C}[0, 1], \quad x \in [0, 1]. \quad (4.33)$$

$B_n$  heißt **Bernstein-Operator**. Das zu einem  $f \in \mathcal{C}[0, 1]$  definierte Polynom  $B_n[f]$  heißt  **$n$ -tes Bernsteinpolynom der Funktion  $f$** .

Im Jahr 1912 hat S. Bernstein folgendes Resultat bezüglich der Approximation stetiger Funktionen gezeigt:

**Satz 4.13** (vgl. [12]). *Sei eine Funktion  $f \in \mathcal{C}[0, 1]$  gegeben. Dann gilt*

$$\lim_{n \rightarrow \infty} \sup_{x \in [0, 1]} |f(x) - B_n[f](x)| = 0. \quad (4.34)$$

Die Folge der Bernsteinpolynome  $B_n[f]$  der Funktion  $f$  konvergiert damit auf dem Intervall  $[0, 1]$  gleichmäßig gegen  $f$ . Insbesondere folgt aus diesem Resultat sofort der Weierstraßsche Approximationssatz, siehe Satz 4.1.

Über die Approximationsgüte lässt sich folgende Aussage treffen:

**Satz 4.14** (vgl., z.B., [54, S. 22]). *Sei eine Funktion  $f \in \mathcal{C}^2[0, 1]$  gegeben. Dann gilt für jedes  $x \in [0, 1]$*

$$\lim_{n \rightarrow \infty} n(f(x) - B_n[f](x)) = -\frac{x(1-x)}{2} f''(x). \quad (4.35)$$

Wenn insbesondere für ein  $x \in [0, 1]$  die zweite Ableitung  $f''(x) \neq 0$  nicht verschwindet, dann ist der Fehler  $f(x) - B_n[f](x)$  exakt von der Ordnung  $n^{-1}$ . Ist also  $f$  eine zweimal stetig differenzierbare Funktion mit beschränkter zweiter Ableitung und  $f$  kein Polynom vom Höchstgrad 1, so ist der Maximalfehler

$$\sup_{x \in [0, 1]} |f(x) - B_n[f](x)| \quad (4.36)$$

exakt von der Ordnung  $n^{-1}$ .

Die Bernsteinpolynome erfüllen viele formerhaltende Eigenschaften, weshalb sie für einige Anwendungen besonders interessant sind. In der folgenden Anmerkung fassen wir einige von ihnen zusammen:

**Anmerkung 4.15.** Die Bernstein-Operatoren  $B_n$  sind positive Operatoren nach Definition 2.22. D. h. es gelten die Eigenschaften (vgl., z.B., [63, S. 248-249]):

- *Linearität:*  $B_n(\lambda f + \mu g) = \lambda B_n[f] + \mu B_n[g]$ ,  $\lambda, \mu \in \mathbb{R}$ ,  $f, g \in \mathcal{C}[0, 1]$ .
- *Positivität:* Aus  $f \geq 0$  folgt stets  $B_n[f] \geq 0$ .

Es gilt für zwei Konstanten  $m, M \in \mathbb{R}$  die folgende Eigenschaft (vgl., z.B., [63, S. 249]):

$$m \leq f \leq M \Rightarrow m \leq B_n[f] \leq M. \quad (4.37)$$

Weiter gilt am Rand für alle  $f \in \mathcal{C}[0, 1]$  (vgl., z.B., [63, S. 248]):

$$B_n[f](0) = f(0) \text{ und } B_n[f](1) = f(1). \quad (4.38)$$

Ist  $f \in \mathcal{C}[0, 1]$  eine konvexe Funktion, d. h. ist

$$\lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2) \quad (4.39)$$

für alle  $\lambda, x_1, x_2 \in [0, 1]$ , so gelten (vgl., z.B., [63, S. 260]):

- $B_n[f](x) \geq f(x)$  für alle  $x \in [0, 1]$  und für alle  $n \geq 1$ .
- $B_{n-1}[f](x) \geq B_n[f](x)$  für alle  $x \in [0, 1]$  und für alle  $n \geq 2$ .

Für ein  $f \in \mathcal{C}[0, 1]$  ist die totale Variation von  $B_n[f]$  höchstens so groß wie die totale Variation von  $f$ , d. h. es gilt (vgl., z.B., [63, S. 281]):

$$\mathcal{V}_0^1[B_n[f]] \leq \mathcal{V}_0^1[f]. \quad (4.40)$$

Ist ein  $f \in \mathcal{C}[0, 1]$  monoton steigend beziehungsweise monoton fallend, so ist  $B_n[f]$  ebenfalls monoton steigend beziehungsweise monoton fallend (vgl., z.B., [63, S. 287]).

Die großen Vorteile der Bernsteinpolynome liegen in ihrer Einfachheit der Darstellung und Berechnung, in den vielen formerhaltenden Eigenschaften sowie in der Verwendung von äquidistanten Stützstellen. Nachteilig ist, dass der Maximalfehler für  $f \notin \mathcal{P}_1$  nur von der Ordnung  $n^{-1}$  ist.

Zur Einordnung unserer Approximationsergebnisse bei der Methode der kleinsten Quadrate aus Kapitel 7 vergleichen wir diese Ergebnisse mit der Fehlerabschätzung aus dem Satz 4.14 in Abschnitt 7.5.

## 4.3 Die Methode der kleinsten Quadrate

In diesem Abschnitt betrachten wir die Methode der kleinsten Quadrate. Schon vor über 200 Jahren haben unter anderem Legendre und Gauß begonnen mit der Methode der kleinsten Quadrate zu arbeiten (vgl., z.B., [57]). Seitdem wird diese Methode in vielen Gebieten der Mathematik sowie in anderen Wissenschaften verwendet (vgl., z.B., [1], [6], [14], [35], [71]). In den Kapiteln 5, 6 und 7 untersuchen wir die Approximationseigenschaften dieser Methode. In diesem Abschnitt stellen wir alle für uns wichtigen Grundbegriffe und Aussagen über die Methode der kleinsten Quadrate zusammen.

Bei der Methode der kleinsten Quadrate handelt es sich um die bestmögliche Approximation bezüglich einer Norm  $\|\cdot\|$ , die durch ein Skalarprodukt  $\langle \cdot, \cdot \rangle$  induziert ist, d. h. es ist  $\|f\| := \langle f, f \rangle^{\frac{1}{2}}$ .

Wir beschränken uns auf den Vektorraum der stetigen Funktionen  $C[a, b]$  und unterscheiden zwischen einem kontinuierlichen und einem diskreten Maß.

### 4.3.1 Der kontinuierliche Fall

Wir betrachten folgende Problemstellung (vgl., z.B., [35, S. 55-59], [34, S. 217], [58, S. 291]):

Sei  $U$  ein Unterraum von  $C[a, b]$  mit  $\dim U = n + 1$ . Weiter sei  $\omega : [a, b] \rightarrow \mathbb{R}$  eine nichtnegative, fast überall positive, integrierbare Funktion, die sogenannte **Gewichtsfunktion**. Zu einer Funktion  $f \in C[a, b]$  ist eine Funktion  $\hat{\varphi} \in U$  gesucht, die für alle  $\varphi \in U$  die Ungleichung

$$\int_a^b (\hat{\varphi}(x) - f(x))^2 \omega(x) dx \leq \int_a^b (\varphi(x) - f(x))^2 \omega(x) dx \quad (4.41)$$

erfüllt. Betrachtet man das folgende Skalarprodukt

$$(f, g)_\omega := \int_a^b f(x)g(x)\omega(x)dx, \quad (4.42)$$

so lässt sich durch dessen induzierte Norm

$$\|f\|_{\omega, C} := \left( \int_a^b |f(x)|^2 \omega(x) dx \right)^{\frac{1}{2}}, \quad (4.43)$$

die Ungleichung (4.41) darstellen durch

$$\|\hat{\varphi} - f\|_{\omega, C} \leq \|\varphi - f\|_{\omega, C}. \quad (4.44)$$

Zunächst können wir bezüglich der Fragen zur Existenz und Eindeutigkeit eines solchen  $\hat{\varphi} \in U$ , folgende bekannte Antwort liefern:

**Satz 4.16** (vgl., z.B., [35, S. 62]). *Seien eine stetige Funktion  $f \in \mathcal{C}[a, b]$  und ein Unterraum  $U$  von  $\mathcal{C}[a, b]$  mit  $\dim U = n + 1$  gegeben. Weiter sei  $\omega : [a, b] \rightarrow \mathbb{R}$  eine integrierbare, nichtnegative und fast überall positive Gewichtsfunktion. Dann existiert ein eindeutig bestimmtes  $\hat{\varphi} \in U$ , welches für alle  $\varphi \in U$  die Ungleichung (4.41) erfüllt.*

Nach Satz 4.16 sind die Existenz und Eindeutigkeit eines  $\hat{\varphi} \in U$  gesichert, welches der Ungleichung (4.41) genügt. Damit lässt sich folgendermaßen der Operator der Methode der kleinsten Quadrate im kontinuierlichen Fall definieren:

**Definition 4.17** (vgl., z.B., [35, S. 62]). *Sei  $U$  ein Unterraum von  $\mathcal{C}[a, b]$  mit  $\dim U = n + 1$ . Weiter sei  $\omega : [a, b] \rightarrow \mathbb{R}$  eine nichtnegative, fast überall positive, integrierbare Gewichtsfunktion. Der **Operator der Methode der kleinsten Quadrate**  $LS_n : \mathcal{C}[a, b] \rightarrow U$  ist definiert durch die Eigenschaft*

$$\int_a^b (LS_n[f](x) - f(x))^2 \omega(x) = \min_{\varphi \in U} \int_a^b (\varphi(x) - f(x))^2 \omega(x). \quad (4.45)$$

**Anmerkung 4.18.** *Der Operator  $LS_n$  ist nach Definition 2.21 eine Projektion. Denn für jedes  $f \in U$  gilt die Ungleichung*

$$0 = \int_a^b (f(x) - f(x))^2 \omega(x) dx \leq \int_a^b (\varphi(x) - f(x))^2 \omega(x) dx \quad (4.46)$$

für alle  $\varphi \in U$ . Nach Satz 4.16 ist diese Lösung eindeutig, womit  $LS_n[f] = f$  für jedes  $f \in U$  gilt.

Um zu einer Funktion  $f \in \mathcal{C}[a, b]$  die entsprechende Funktion  $LS_n[f] \in U$  zu bestimmen, benötigt man eine Basis  $\{\varphi_\nu\}_{\nu=0}^n$  von  $U$ .  $LS_n[f]$  besitzt dann eine eindeutige Darstellung

$$LS_n[f] = \sum_{\nu=0}^n a_\nu \varphi_\nu, \quad (4.47)$$

mit entsprechenden Koeffizienten  $a_0, \dots, a_n$ .

Sind eine Funktion  $f \in \mathcal{C}[a, b]$  sowie eine Basis  $\{\varphi_\nu\}_{\nu=0}^n$  von  $U$  gegeben, so lässt sich die zugehörige Lösung  $LS_n[f]$  der Methode der kleinsten Quadrate durch Lösen des linearen Gleichungssystems

$$\sum_{\nu=0}^n a_\nu (\varphi_\mu, \varphi_\nu)_\omega = (f, \varphi_\mu)_\omega, \quad \mu = 0, \dots, n \quad (4.48)$$

bestimmen (vgl., z.B., [35, S. 61]). Dieses Gleichungssystem besteht aus den sogenannten **Normalgleichungen**. Dessen eindeutige Lösung  $\{a_\nu\}_{\nu=0}^n$  ergibt die zugehörige Lösung

$$LS_n[f] = \sum_{\nu=0}^n a_\nu \varphi_\nu \quad (4.49)$$

der Methode der kleinsten Quadrate.

Beim Lösen der Normalgleichungen können folgende Probleme auftreten:

Zum einen sind die Koeffizienten  $a_0, \dots, a_n$  alle abhängig von  $n$ , sodass man bei der Erhöhung von  $n$  das entsprechende Gleichungssystem vollständig neu lösen muss. Zum anderen können bei der Berechnung der Lösung erhebliche numerische Probleme auftreten, wenn beispielsweise die zugehörige Matrix  $\left((\varphi_\mu, \varphi_\nu)_\omega\right)_{\mu, \nu=0, \dots, n}$  schlecht konditioniert ist (vgl., z.B., [35, S. 62-63]).

Es lässt sich jedoch Abhilfe schaffen, wenn das System  $\{\varphi_\nu\}_{\nu=0}^n$  orthogonal auf  $[a, b]$  bezüglich des Skalarprodukts  $(f, g)_\omega$  gemäß (4.42) ist, d. h. wenn

$$(\varphi_\mu, \varphi_\nu)_\omega = 0 \text{ für alle } \mu \neq \nu \text{ und } (\varphi_\mu, \varphi_\mu)_\omega > 0 \text{ für alle } \mu = 0, \dots, n \quad (4.50)$$

gilt. Dann vereinfacht sich die Matrix  $\left((\varphi_\mu, \varphi_\nu)_\omega\right)_{\mu, \nu=0, \dots, n}$  zu einer Diagonalmatrix und der Operator der Methode der kleinsten Quadrate lässt sich darstellen durch

$$LS_n[f] = \sum_{\nu=0}^n \frac{(f, \varphi_\nu)_\omega}{(\varphi_\nu, \varphi_\nu)_\omega} \varphi_\nu \quad (4.51)$$

(vgl., z.B., [35, S. 63]).

In dieser Arbeit betrachten wir folgenden, in der Praxis wichtigen, Standardfall

- $[a, b] = [-1, 1]$ ,
- $U = \mathcal{P}_n$  ist der Raum der Polynome vom Höchstgrad  $n$ ,
- $\omega(x) := \varrho(x) = (1-x)^\alpha(1+x)^\beta$  mit  $\alpha, \beta > -1$ .

Das Orthogonalsystem besteht dann aus den in Unterabschnitt 3.1.2 untersuchten Jacobi-Polynomen  $P_n^{\alpha, \beta}$ . In diesem Fall hat der Operator der Methode der kleinsten Quadrate die Gestalt

$$LS_n[f] = \sum_{k=0}^n \frac{\left(P_k^{\alpha, \beta}, f\right)_\varrho}{\left(P_k^{\alpha, \beta}, P_k^{\alpha, \beta}\right)_\varrho} P_k^{\alpha, \beta}. \quad (4.52)$$

Die Approximationseigenschaften dieser Methode untersuchen wir ausführlich in den folgenden Kapiteln 5, 6 und 7. Sie dient uns vor allem zum Vergleich mit der Methode im diskreten Fall, welche wir im folgenden Unterabschnitt 4.3.2 diskutieren.

### 4.3.2 Der diskrete Fall

Die Vorgehensweise in diesem Unterabschnitt ist häufig formal analog mit der Vorgehensweise in Unterabschnitt 4.3.1. Zum Komfort der Leserin und des Lesers stellen wir trotzdem die einzelnen Definitionen und Gleichungen für spätere Kapitel explizit bereit.

Wir betrachten folgende Problemstellung (vgl., z.B., [35, S. 55-59], [34, S. 217], [58, S. 291]):

Seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, N$ . Weiter sei  $\omega : [a, b] \rightarrow \mathbb{R}$  eine auf  $\{x_\mu\}_{\mu=0}^N$  positive Funktion, die sogenannte **Gewichtsfunktion**. Für ein  $n \leq N$  sei  $U$  ein Unterraum von  $\mathcal{C}[a, b]$  mit  $\dim U = n + 1$  auf  $\{x_\mu\}_{\mu=0}^N$ . D. h. jede Basis  $\{\varphi_\nu\}_{\nu=0}^n$  von  $U$  genügt folgender Eigenschaft:

$$\sum_{\nu=0}^n c_\nu \varphi_\nu(x_\mu) = 0, \quad \mu = 0, \dots, N \Rightarrow c_0 = \dots = c_n = 0. \quad (4.53)$$

Zu einer Funktion  $f \in \mathcal{C}[a, b]$  ist eine Funktion  $\hat{\varphi} \in U$  gesucht, die für alle  $\varphi \in U$  die Ungleichung

$$\sum_{\mu=0}^N (\hat{\varphi}(x_\mu) - f(x_\mu))^2 \omega(x_\mu) \leq \sum_{\mu=0}^N (\varphi(x_\mu) - f(x_\mu))^2 \omega(x_\mu) \quad (4.54)$$

erfüllt. Betrachtet man das folgende Skalarprodukt

$$\langle f, g \rangle_\omega := \sum_{\mu=0}^N f(x_\mu) g(x_\mu) \omega(x_\mu), \quad (4.55)$$

so lässt sich durch dessen induzierte Norm

$$\|f\|_{\omega, D} := \left( \sum_{\mu=0}^N |f(x_\mu)|^2 \omega(x_\mu) \right)^{\frac{1}{2}}, \quad (4.56)$$

die Ungleichung (4.54) darstellen durch

$$\|\hat{\varphi} - f\|_{\omega, D} \leq \|\varphi - f\|_{\omega, D}. \quad (4.57)$$

Zunächst präsentieren wir bezüglich der Fragen zur Existenz und Eindeutigkeit eines solchen  $\hat{\varphi} \in U$ , ähnlich wie im kontinuierlichen Fall, folgende bekannte Ergebnisse:

**Satz 4.19** (vgl., z.B., [35, S. 62]). *Sei eine stetige Funktion  $f \in \mathcal{C}[a, b]$  gegeben. Seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, N$ . Weiter sei  $\omega : [a, b] \rightarrow \mathbb{R}$  eine auf  $\{x_\mu\}_{\mu=0}^N$  positive Gewichtsfunktion. Für ein  $n \leq N$  sei  $U$  ein Unterraum von  $\mathcal{C}[a, b]$  mit  $\dim U = n + 1$  auf  $\{x_\mu\}_{\mu=0}^N$ . Dann existiert ein eindeutig bestimmtes  $\hat{\varphi} \in U$ , welches für alle  $\varphi \in U$  die Ungleichung (4.54) erfüllt.*

Nach Satz 4.19 sind die Existenz und Eindeutigkeit eines  $\hat{\varphi} \in U$  gesichert, welches der Ungleichung (4.54) genügt. Damit lässt sich folgendermaßen der Operator der Methode der kleinsten Quadrate im diskreten Fall definieren:

**Definition 4.20** (vgl., z.B., [35, S. 62]). *Seien  $x_\mu \in [a, b]$  paarweise verschiedene Stützstellen für  $\mu = 0, \dots, N$ . Weiter sei  $\omega : [a, b] \rightarrow \mathbb{R}$  eine auf  $\{x_\mu\}_{\mu=0}^N$  positive Gewichtsfunktion. Für ein  $n \leq N$  sei  $U$  ein Unterraum von  $\mathcal{C}[a, b]$  mit  $\dim U =$*

$n+1$  auf  $\{x_\mu\}_{\mu=0}^N$ . Der **Operator der Methode der kleinsten Quadrate**  $LS_n^N : \mathcal{C}[a, b] \rightarrow U$  ist definiert durch die Eigenschaft

$$\sum_{\mu=0}^N \left( LS_n^N[f](x_\mu) - f(x_\mu) \right)^2 \omega(x_\mu) = \min_{\varphi \in U} \sum_{\mu=0}^N (\varphi(x_\mu) - f(x_\mu))^2 \omega(x_\mu). \quad (4.58)$$

**Anmerkung 4.21.** Der Operator  $LS_n^N$  ist nach Definition 2.21 eine Projektion. Denn für jedes  $f \in U$  gilt die Ungleichung

$$0 = \sum_{\mu=0}^N (f(x_\mu) - f(x_\mu))^2 \omega(x_\mu) \leq \sum_{\mu=0}^N (\varphi(x_\mu) - f(x_\mu))^2 \omega(x_\mu) \quad (4.59)$$

für alle  $\varphi \in U$ . Nach Satz 4.19 ist diese Lösung eindeutig, womit  $LS_n^N[f] = f$  für jedes  $f \in U$  gilt.

**Anmerkung 4.22** (vgl., z.B., [35, S. 57]). Im Spezialfall für  $N = n$  und  $U = \mathcal{P}_n$  entspricht die Methode der kleinsten Quadrate der Polynominterpolation aus Abschnitt 4.1. D. h. für jede stetige Funktion  $f \in \mathcal{C}[a, b]$  entspricht das Interpolationspolynom  $p_n[f]$  zu  $\{(x_\mu, f(x_\mu))\}_{\mu=0}^n$  dem Polynom  $LS_n^n[f]$ . Also gilt  $p_n[f] = LS_n^n[f]$ .

Um zu einer Funktion  $f \in \mathcal{C}[a, b]$  die entsprechende Funktion  $LS_n^N[f] \in U$  zu bestimmen, benötigt man eine Basis  $\{\varphi_\nu\}_{\nu=0}^n$  von  $U$ .  $LS_n^N[f]$  besitzt dann eine eindeutige Darstellung

$$LS_n^N[f] = \sum_{\nu=0}^n a_\nu \varphi_\nu, \quad (4.60)$$

mit entsprechenden Koeffizienten  $a_0, \dots, a_n$ .

Sind eine Funktion  $f \in \mathcal{C}[a, b]$  sowie eine Basis  $\{\varphi_\nu\}_{\nu=0}^n$  von  $U$  gegeben, so lässt sich, analog zum kontinuierlichen Fall, die zugehörige Lösung  $LS_n^N[f]$  der Methode der kleinsten Quadrate durch Lösen des linearen Gleichungssystems

$$\sum_{\nu=0}^n a_\nu \langle \varphi_\lambda, \varphi_\nu \rangle_\omega = \langle f, \varphi_\lambda \rangle_\omega, \quad \lambda = 0, \dots, n \quad (4.61)$$

bestimmen (vgl., z.B., [35, S. 61]). Dieses Gleichungssystem besteht aus den sogenannten **Normalgleichungen**. Dessen eindeutige Lösung  $\{a_\nu\}_{\nu=0}^n$  ergibt die zugehörige Lösung

$$LS_n^N[f] = \sum_{\nu=0}^n a_\nu \varphi_\nu \quad (4.62)$$

der Methode der kleinsten Quadrate.

Beim Lösen der Normalgleichungen können folgende Probleme auftreten:

Zum einen sind die Koeffizienten  $a_0, \dots, a_n$  alle abhängig von  $n$  und  $N$ , sodass man bei der Erhöhung von  $n$  beziehungsweise  $N$  das entsprechende Gleichungssystem

vollständig neu lösen muss. Zum anderen können, genau wie im kontinuierlichen Fall, bei der Berechnung der Lösung erhebliche numerische Probleme auftreten, wenn beispielsweise die zugehörige Matrix  $(\langle \varphi_\lambda, \varphi_\nu \rangle_\omega)_{\lambda, \nu=0, \dots, n}$  schlecht konditioniert ist (vgl., z.B., [35, S. 62-63]).

Auch im diskreten Fall lässt sich jedoch Abhilfe schaffen, wenn das System  $\{\varphi_\nu\}_{\nu=0}^n$  orthogonal auf  $\{x_\mu\}_{\mu=0}^N$  bezüglich des Skalarprodukts  $\langle f, g \rangle_\omega$  gemäß (4.55) ist, d. h. wenn

$$\langle \varphi_\lambda, \varphi_\nu \rangle_\omega = 0 \text{ für alle } \lambda \neq \nu \text{ und } \langle \varphi_\lambda, \varphi_\lambda \rangle_\omega > 0 \text{ für alle } \lambda = 0, \dots, n \quad (4.63)$$

gilt. Dann vereinfacht sich die Matrix  $(\langle \varphi_\lambda, \varphi_\nu \rangle_\omega)_{\lambda, \nu=0, \dots, n}$  zu einer Diagonalmatrix und der Operator der Methode der kleinsten Quadrate lässt sich darstellen durch

$$LS_n^N[f] = \sum_{\nu=0}^n \frac{\langle f, \varphi_\nu \rangle_\omega}{\langle \varphi_\nu, \varphi_\nu \rangle_\omega} \varphi_\nu \quad (4.64)$$

(vgl., z.B., [35, S. 63]).

In dieser Arbeit betrachten wir folgenden, in der Praxis wichtigen, Standardfall

- $[a, b] = [-1, 1]$ ,
- $U = \mathcal{P}_n$  ist der Raum der Polynome vom Höchstgrad  $n$ ,
- $\{x_0, \dots, x_N\}$  ist ein äquidistantes Gitter auf dem Intervall  $[-1, 1]$ , d. h.  $x_\mu = -1 + 2\mu/N$  für  $\mu = 0, \dots, N$ ,
- $\omega(x) = \binom{\alpha+x}{x} \binom{\beta+N-x}{N-x}$  mit  $\alpha, \beta > -1$ .

Das Orthogonalsystem besteht dann aus den in Unterabschnitt 3.2.2 untersuchten Hahn-Polynomen  $Q_n(\cdot; \alpha, \beta, N)$ . In diesem Fall hat der Operator der Methode der kleinsten Quadrate die Gestalt

$$LS_n^N[f] = \sum_{k=0}^n \frac{\langle Q_k(\cdot, \alpha, \beta, N), f\left(\frac{2}{N}(\cdot) - 1\right) \rangle_\omega}{\langle Q_k(\cdot, \alpha, \beta, N), Q_k(\cdot, \alpha, \beta, N) \rangle_\omega} Q_k\left(\frac{N}{2}(1 + \cdot), \alpha, \beta, N\right). \quad (4.65)$$

Man beachte dabei die lineare Transformation  $x \mapsto \frac{N}{2}(1+x)$  des Intervalls  $[-1, 1]$  auf das Intervall  $[0, N]$ , da die Hahn-Polynome  $Q_n(\cdot; \alpha, \beta, N)$  auf dem Intervall  $[0, N]$  definiert sind, siehe dazu Unterabschnitt 3.2.2.

Die Approximationseigenschaften dieser Methode untersuchen wir ausführlich in den folgenden Kapiteln 5, 6 und 7. Dabei vergleichen wir diese Methode auch mit allen anderen in diesem Kapitel vorgestellten Approximationsmethoden.

## 4.4 Polynome bester Approximation

Abschließend geben wir einen kurzen Einblick in die Bestapproximation durch Polynome. Bei der Bestapproximation handelt es sich um die bestmögliche Approximation bezüglich der Supremumsnorm. Aus diesem Grund kann sie sehr gut als Maßstab für jede Approximationsmethode heran gezogen werden.

Wir vergleichen unsere Abschätzungen bei der Methode der kleinsten Quadrate mit einer Abschätzung bei der Bestapproximation in Abschnitt 7.7.

**Definition 4.23** (vgl., z.B., [65, S. 113]). *Sei eine Funktion  $f \in \mathcal{C}[a, b]$  gegeben. Ein Polynom  $p_n \in \mathcal{P}_n$  höchstens  $n$ -ten Grades heißt **Polynom bester Approximation von  $f$** , wenn es der Beziehung*

$$\sup_{x \in [a, b]} |f(x) - p_n(x)| = \inf_{p \in \mathcal{P}_n} \sup_{x \in [a, b]} |f(x) - p(x)| \quad (4.66)$$

genügt.

Es ergeben sich sofort folgende Fragen:

- Gibt es für jedes  $n \in \mathbb{N}_0$  ein Polynom bester Approximation  $p_n \in \mathcal{P}_n$  von  $f$ ?
- Ist ein Polynom bester Approximation immer eindeutig?
- Wie lässt sich ein Polynom bester Approximation bestimmen?
- Wie gut ist die Approximation bezüglich der Supremumsnorm?

Die ersten beiden Fragen lassen sich durch das sogenannte Tschebyscheff-Theorem beantworten:

**Satz 4.24** (vgl., z.B., [65, S. 113]). *Sei eine Funktion  $f \in \mathcal{C}[a, b]$  gegeben. Dann existiert für jedes  $n \in \mathbb{N}_0$  das Polynom bester Approximation  $p_n \in \mathcal{P}_n$  von  $f$  und ist eindeutig.*

Wäre die Bestimmung des Polynoms bester Approximation einfach und schnell möglich, so wäre diese Approximationsmethode in den meisten Fällen anderen Approximationsmethoden vorzuziehen. Doch genau hierin besteht das Problem. Das Polynom bester Approximation lässt sich zwar durch den sogenannten Remez-Algorithmus annähern, dieser ist jedoch sehr aufwendig und für die Praxis nur selten verwendbar (vgl., z.B., [7]). Aus diesem Grund gehen wir nicht weiter auf ihn ein.

Seit den grundlegenden Untersuchungen über Bestapproximation von D. Jackson zu Beginn des 20. Jahrhunderts (vgl. [42]), sind zahlreiche Ergebnisse über Polynome bester Approximation publiziert worden (vgl., z.B., [25], [42], [53], [59], [74]).

Hier beschränken wir uns für die Diskussion unserer Ergebnisse in Abschnitt 7.7 auf die folgende unverbesserbare Abschätzung:

**Satz 4.25** (vgl., z.B., [56, S. 78]). *Sei  $f \in \mathcal{C}^{n+1}[-1, 1]$  gegeben und sei  $p_n \in \mathcal{P}_n$  das Polynom bester Approximation von  $f$ . Dann gilt*

$$\sup_{x \in [-1, 1]} |f(x) - p_n(x)| \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{1}{(n+1)!2^n}. \quad (4.67)$$

*Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $\frac{1}{(n+1)!2^n}$  in Ungleichung (4.67) durch keine kleinere Zahl ersetzt werden kann.*



## 5 Divergenz der Methode der kleinsten Quadrate

Bezüglich der in dieser Arbeit zu untersuchenden Operatoren  $LS_n^N$  gemäß (4.65) zeigen wir hier auf, dass nur unter Voraussetzung der Stetigkeit der zu approximierenden Funktion für keine Folge  $(LS_n^N)_n$  Konvergenz in  $\mathcal{C}[-1, 1]$  gewährleistet werden kann.

In den letzten 60 Jahren wurden in der Theorie der Orthogonalreihen Divergenzergebnisse meist dadurch erzielt, dass man die Divergenz der zugehörigen Operatornormen nachwies. Grundlegend hierfür war das Resultat [11] von D. Berman:

**Satz 5.1** (vgl. [11] sowie z.B. [59, S. 189]). *Für jedes  $n \in \mathbb{N}$  gilt für die Norm eines beschränkten linearen Projektionsoperators*

$$L_n : (\mathcal{C}_{2\pi}, \|\cdot\|_\infty) \rightarrow (\mathcal{T}_n, \|\cdot\|_\infty) \quad (5.1)$$

die Ungleichung

$$\|L_n\| \geq \|S_n\| = \frac{4}{\pi^2} \ln n + \mathcal{O}(1). \quad (5.2)$$

Hierbei ist  $\mathcal{C}_{2\pi}$  die Menge der  $2\pi$ -periodischen stetigen Funktionen,  $\mathcal{T}_n$  die Menge der trigonometrischen Ausdrücke  $n$ -ter Ordnung und  $S_n$  der  $n$ -te Fourier-Teilsummen Operator.

Die Operatornorm  $\|\cdot\|$  ist dabei gemäß Satz 2.20 definiert und  $\|\cdot\|_\infty$  bezeichnet die Supremumsnorm. Den Begriff des Projektionsoperators verwenden wir gemäß Definition 2.21.

Im periodischen Fall ist  $S_n$  also die sogenannte „Minimal-Projektion“. Für den Fall, dass der Bildraum von  $L_n$  nicht wie in (5.1), sondern die Menge  $\mathcal{P}_n$  der Polynome vom Höchstgrad  $n$  ist, kennt man bisher noch keine Minimal-Projektion. Die Divergenz der Normen jeder Folge  $(L_n)_n$  beschränkter linearer Projektionsoperatoren

$$L_n : (\mathcal{C}[a, b], \|\cdot\|_\infty) \rightarrow (\mathcal{P}_n[a, b], \|\cdot\|_\infty) \quad (5.3)$$

haben schon 1948 Harsiladze und Lozinskii bewiesen (vgl., z.B., [59, S. 194]). Das asymptotisch korrekte Ergebnis bezüglich der Minimal-Projektion im Polynomfall hat K. Petras 1988 erzielt:

**Satz 5.2** (vgl. [62]). *Für jedes  $n \geq 2$  gilt für die Norm eines beschränkten linearen Projektionsoperators*

$$L_n : (\mathcal{C}[a, b], \|\cdot\|_\infty) \rightarrow (\mathcal{P}_n[a, b], \|\cdot\|_\infty) \quad (5.4)$$

die Ungleichung

$$\|L_n\| \geq \frac{4}{\pi^2} (\ln n - \ln \ln n). \quad (5.5)$$

Unter obigen Voraussetzungen kann in (5.5) die Konstante  $\frac{4}{\pi^2}$  durch keine kleinere Zahl ersetzt werden.

Aus dem obigen Divergenzergbnis erhält man das folgende Korollar:

**Korollar 5.3** (vgl. [70, S. 106, Satz 4.10]). *Es gibt keine Folge beschränkter linearer Projektionsoperatoren*

$$L_n : (\mathcal{C}[a, b], \|\cdot\|_\infty) \rightarrow (\mathcal{P}_n[a, b], \|\cdot\|_\infty) \quad (5.6)$$

mit  $\|L_n f - f\|_\infty \rightarrow 0$  für alle  $f \in \mathcal{C}[a, b]$ .

Für eine Folge von Projektionsoperatoren  $(L_n)_n$  des Typs (5.6) reicht also der Nachweis der Linearität und Beschränktheit aus, um die Existenz einer stetigen Funktion  $f$  nachzuweisen, für die diese Folge  $(L_n f)_n$  nicht gleichmäßig konvergiert. Diese Vorgehensweise ist bei der Methode der kleinsten Quadrate, sowohl im kontinuierlichen Fall als auch im diskreten Fall, leicht durchzuführen. Für den kontinuierlichen Fall sind zahlreiche tiefliegende Divergenzergbnisse bekannt, man vgl. z.B. die Ausführungen in [4], [8], [46], [55]. Für den diskreten Fall scheint ein Korollar 5.3 entsprechendes Ergebnis in der Literatur nicht explizit genannt zu sein. Wir führen daher im Folgenden kurz den einfachen Nachweis durch, dass die Operatoren der Methode der kleinsten Quadrate aus Gleichung (4.65) beschränkte lineare Projektionen sind.

**Lemma 5.4.** *Seien  $\alpha, \beta > -1$  und  $N, n \in \mathbb{N}$  mit  $n \leq N$ . Dann ist*

$$LS_n^N : (\mathcal{C}[-1, 1], \|\cdot\|_\infty) \rightarrow (\mathcal{P}_n[-1, 1], \|\cdot\|_\infty) \quad (5.7)$$

definiert durch

$$LS_n^N[f] = \sum_{k=0}^n \frac{\langle Q_k(\cdot, \alpha, \beta, N), f\left(\frac{2}{N}(\cdot) - 1\right) \rangle_\omega}{\langle Q_k(\cdot, \alpha, \beta, N), Q_k(\cdot, \alpha, \beta, N) \rangle_\omega} Q_k\left(\frac{N}{2}(1 + \cdot), \alpha, \beta, N\right), \quad (5.8)$$

für alle  $f \in \mathcal{C}[-1, 1]$ , ein linearer beschränkter Projektionsoperator.

*Beweis.* Zur Linearität: Seien  $\lambda_1, \lambda_2 \in \mathbb{R}$  und  $f_1, f_2 \in \mathcal{C}[-1, 1]$  gegeben. Dann gilt mit der Linearität des Skalarprodukts

$$\begin{aligned}
& LS_n^N [\lambda_1 f_1 + \lambda_2 f_2] \\
&= \sum_{k=0}^n \frac{\langle Q_k(\cdot, \alpha, \beta, N), (\lambda_1 f_1 + \lambda_2 f_2) \left(\frac{2}{N}(\cdot) - 1\right) \rangle_\omega}{\langle Q_k(\cdot, \alpha, \beta, N), Q_k(\cdot, \alpha, \beta, N) \rangle_\omega} Q_k \left( \frac{N}{2}(1 + \cdot), \alpha, \beta, N \right) \\
&= \lambda_1 \sum_{k=0}^n \frac{\langle Q_k(\cdot, \alpha, \beta, N), f_1 \left(\frac{2}{N}(\cdot) - 1\right) \rangle_\omega}{\langle Q_k(\cdot, \alpha, \beta, N), Q_k(\cdot, \alpha, \beta, N) \rangle_\omega} Q_k \left( \frac{N}{2}(1 + \cdot), \alpha, \beta, N \right) \\
&\quad + \lambda_2 \sum_{k=0}^n \frac{\langle Q_k(\cdot, \alpha, \beta, N), f_2 \left(\frac{2}{N}(\cdot) - 1\right) \rangle_\omega}{\langle Q_k(\cdot, \alpha, \beta, N), Q_k(\cdot, \alpha, \beta, N) \rangle_\omega} Q_k \left( \frac{N}{2}(1 + \cdot), \alpha, \beta, N \right) \\
&= \lambda_1 LS_n^N [f_1] + \lambda_2 LS_n^N [f_2].
\end{aligned} \tag{5.9}$$

Somit ist die Linearität gezeigt.

Zur Beschränktheit: Sei nun ein  $f \in \mathcal{C}[-1, 1]$  mit  $\|f\|_\infty \leq 1$  gegeben. Dann gilt

$$\begin{aligned}
\|LS_n^N [f]\|_\infty &\leq \sum_{k=0}^n \frac{\left| \langle Q_k(\cdot, \alpha, \beta, N), f \left(\frac{2}{N}(\cdot) - 1\right) \rangle_\omega \right|}{\langle Q_k(\cdot, \alpha, \beta, N), Q_k(\cdot, \alpha, \beta, N) \rangle_\omega} \left\| Q_k \left( \frac{N}{2}(1 + \cdot), \alpha, \beta, N \right) \right\|_\infty \\
&\leq \sum_{k=0}^n \frac{\|Q_k(\cdot, \alpha, \beta, N)\|_\omega \cdot \|f \left(\frac{2}{N}(\cdot) - 1\right)\|_\omega}{\langle Q_k(\cdot, \alpha, \beta, N), Q_k(\cdot, \alpha, \beta, N) \rangle_\omega} \left\| Q_k \left( \frac{N}{2}(1 + \cdot), \alpha, \beta, N \right) \right\|_\infty \\
&= \sum_{k=0}^n \frac{\|f \left(\frac{2}{N}(\cdot) - 1\right)\|_\omega}{\|Q_k(\cdot, \alpha, \beta, N)\|_\omega} \left\| Q_k \left( \frac{N}{2}(1 + \cdot), \alpha, \beta, N \right) \right\|_\infty \\
&\leq \sum_{k=0}^n \frac{\|1\|_\omega}{\|Q_k(\cdot, \alpha, \beta, N)\|_\omega} \left\| Q_k \left( \frac{N}{2}(1 + \cdot), \alpha, \beta, N \right) \right\|_\infty \\
&=: C(n, N, \alpha, \beta) < \infty,
\end{aligned} \tag{5.10}$$

woraus die Beschränktheit von  $LS_n^N$  folgt. Denn die Konstante  $C(n, N, \alpha, \beta)$  ist unabhängig von  $f$ .

Die Projektionseigenschaft entnehmen wir der Anmerkung 4.21 und damit folgt die Behauptung.  $\square$

Wir sind nun in der Lage Korollar 5.3 anzuwenden. Damit erhalten wir folgendes Resultat bezüglich der Divergenz der Methode der kleinsten Quadrate im diskreten Fall.

**Satz 5.5** (zur Methode vgl., z.B., [4], [46, S. 293 ff.], [59, S. 178 ff.]). *Seien  $\alpha, \beta > -1$  und eine Folge  $(N_n)_{n \in \mathbb{N}}$  mit  $N_n \geq n$  für alle  $n \in \mathbb{N}$  gegeben. Weiter seien  $LS_n^{N_n}$  die Operatoren der Methode der kleinsten Quadrate wie in Lemma 5.4 definiert. Dann existiert eine stetige Funktion  $f \in \mathcal{C}[-1, 1]$ , sodass  $LS_n^{N_n}[f]$  nicht gleichmäßig gegen  $f$  konvergiert, d. h. es gilt*

$$\limsup_{n \rightarrow \infty} \|LS_n^{N_n}[f] - f\|_\infty \neq 0. \tag{5.11}$$

Wir halten an dieser Stelle also fest, dass es im Fall von äquidistanten Stützstellen kein Verhältnis  $N/n$  zwischen der Anzahl der Stützstellen  $N + 1$  und dem Polynomgrad  $n$  geben kann, sodass die Methode der kleinsten Quadrate  $LS_n^N[f]$  für jedes stetige  $f \in \mathcal{C}[-1, 1]$  gleichmäßig gegen  $f$  konvergiert. Um Konvergenz garantieren zu können, müssen wir also die Funktionenklasse verkleinern. Untersuchungen hierzu stellen wir in den folgenden Kapiteln 6 und 7 an.

Dass es allerdings häufig nicht zufriedenstellend ist, nur die Funktionenklasse zu verkleinern, sei an dieser Stelle kurz ausgeführt. In Unterabschnitt 4.3.2 ist in Anmerkung 4.22 dargestellt, dass die Methode der kleinsten Quadrate im Spezialfall  $N = n$  der Polynominterpolation in  $N + 1$  Stützstellen entspricht. Im Fall von äquidistanten Stützstellen haben wir in Unterabschnitt 4.1.1 für zwei ausgewählte Funktionen die Divergenz der Polynominterpolation dargestellt: Zum einen für die analytisch fortsetzbare Runge-Funktion (4.16), zum anderen für die elementare Betragsfunktion (4.20). Diese beiden Beispiele zeigen uns, dass man selbst für die Klasse der analytischen Funktionen keine gleichmäßige Konvergenz der Methode der kleinsten Quadrate im diskreten Fall für äquidistante Stützstellen garantieren kann. Es liegt daher nahe, sich mit der Stützstellenanzahl  $N + 1$  im Verhältnis zum Polynomgrad  $n$  zu beschäftigen. Dies wird in den folgenden Kapiteln 6 und 7 untersucht.

# 6 Punktweise Konvergenz der Methode der kleinsten Quadrate

In diesem Kapitel diskutieren wir die Frage, unter welchen Bedingungen die Methode der kleinsten Quadrate punktweise konvergiert. Dabei gehen wir insbesondere auf den in Unterabschnitt 4.3.2 erläuterten diskreten Fall ein. Das Hauptergebnis dieses Kapitels ist Theorem 6.25. Dieses Ergebnis stellt eine Abschätzung der Methode auf einem äquidistanten Gitter mit  $N + 1$  Stützstellen auf dem Intervall  $[-1, 1]$  für die Menge der dort stetigen Funktionen von beschränkter Variation bereit. Der Fehler im diskreten Fall wird hierbei durch den Fehler im kontinuierlichen Fall und einem zusätzlichen Term abgeschätzt. Dazu verwenden wir die Operatoren der Methode der kleinsten Quadrate  $LS_n$  beziehungsweise  $LS_n^N$  gemäß Gleichung (4.52) beziehungsweise (4.65). Unsere Resultate erzielen wir für den symmetrischen Fall  $\alpha = \beta$ . Wir untersuchen folgendes Problem: Für welche Funktionen  $f \in K \subset \mathcal{C}[-1, 1]$  und welches Verhältnis  $N/n$  konvergiert die Folge  $(LS_n^N[f])$  punktweise?

## 6.1 Hauptergebnis zur punktweisen Konvergenz

In diesem Abschnitt beweisen wir unser Hauptergebnis zur punktweisen Konvergenz der Methode der kleinsten Quadrate. Als ersten Schritt untersuchen wir die Beschränktheit der Hahn-Polynome in Abhängigkeit vom Verhältnis  $N/n$ . Darüber gibt folgendes Lemma Aufschluss:

**Lemma 6.1.** *Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$*

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (6.1)$$

*Dann gilt für beliebiges  $n \leq n(\alpha, N)$*

$$\max_{x \in [0, N]} |Q_n(x; \alpha, \alpha, N)| = 1. \quad (6.2)$$

Dieses Resultat folgt unmittelbar aus dem folgenden von H. Dette bewiesenen Ergebnis, durch das Setzen von  $\alpha = \beta$ :

**Satz 6.2** (vgl. [26]). *Seien  $\alpha + \beta > -1$  und*

$$n(\alpha, \beta, N) := -\frac{1}{2} \left( \alpha + \beta - 1 - \sqrt{(\alpha + \beta + 1)(\alpha + \beta + 2N + 1)} \right). \quad (6.3)$$

Dann erfüllen die Hahn-Polynome die Ungleichungen

$$|Q_n(x; \alpha, \beta, N)| \leq \max \left\{ 1, \frac{(\beta + 1)_n}{(\alpha + 1)_n} \right\} = \max \{ |Q_n(0; \alpha, \beta, N)|, |Q_n(N; \alpha, \beta, N)| \}, \quad (6.4)$$

für alle  $x \in [0, N]$  und alle  $n \leq n(\alpha, \beta, N)$ .

In Theorem 6.25 wird der Fehler der Methode der kleinsten Quadrate im diskreten Fall durch den Fehler im kontinuierlichen Fall abgeschätzt. Dazu untersuchen wir im Folgenden den Zusammenhang zwischen den Jacobi-Polynomen  $P_n^{\alpha, \beta}$  und den Hahn-Polynomen  $Q_n(\cdot; \alpha, \beta, N)$ . Die folgende Beziehung zwischen den Hahn-Polynomen  $Q_n$  und den Jacobi-Polynomen  $P_n \equiv P_n^{\alpha, \beta}$  ist bereits bekannt, siehe Gleichung (3.56). Es gilt für fest gewähltes  $n$

$$\lim_{N \rightarrow \infty} (-1)^n \binom{n + \alpha}{n} Q_n \left( \frac{N}{2}(1 + x); \alpha, \beta, N \right) = P_n^{\beta, \alpha}(x), \quad (6.5)$$

für jedes  $x \in [-1, 1]$ . Wir untersuchen nun in Bezug auf Gleichung (6.5) folgende, allgemeinere Situation: Der Polynomgrad  $n$  sei im Folgenden nicht länger fest gewählt, sondern wir interessieren uns für den Fall, dass  $n$  in Abhängigkeit von  $N$  wächst. Um diesbezüglich die Verbindung zwischen den Jacobi-Polynomen und den Hahn-Polynomen genauer zu untersuchen, betrachten wir die Rekursionsgleichungen (3.16) beziehungsweise (3.48), welche die Jacobi-Polynome beziehungsweise die Hahn-Polynome erfüllen. Dabei formen wir diese Gleichungen so um, dass wir anschließend die Koeffizienten miteinander vergleichen können. Aufgrund der Eindeutigkeit der Lösungen der Rekursionsgleichungen erhalten wir auf diese Weise ein Verhältnis  $N/n$  für die Konvergenz der Hahn-Polynome gegen die Jacobi-Polynome.

Der Kürze halber schreiben wir

$$\tilde{Q}_n(x) := (-1)^n \binom{n + \alpha}{n} Q_n \left( \frac{N}{2}(1 + x); \alpha, \alpha, N \right). \quad (6.6)$$

Durch Substitution in der Rekursionsgleichung (3.48) erhalten wir für  $\tilde{Q}_n$  mit  $x \in [-1, 1]$  die Rekursion

$$\begin{aligned} & \frac{N}{2}(1 + x)\tilde{Q}_n(x) \\ &= A_n \frac{\binom{n + \alpha}{n}}{\binom{n + 1 + \alpha}{n + 1}} \tilde{Q}_{n + 1}(x) + (A_n + C_n) \tilde{Q}_n(x) + C_n \frac{\binom{n + \alpha}{n}}{\binom{n - 1 + \alpha}{n - 1}} \tilde{Q}_{n - 1}(x) \quad (6.7) \\ &= A_n \frac{n + 1}{n + 1 + \alpha} \tilde{Q}_{n + 1}(x) + (A_n + C_n) \tilde{Q}_n(x) + C_n \frac{n + \alpha}{n} \tilde{Q}_{n - 1}(x). \end{aligned}$$

Einfache Umformungen der Gleichung ergeben mit den Konstanten  $A_n$  und  $C_n$  aus den Gleichungen (3.49) und (3.50)

$$\begin{aligned}
 & x\tilde{Q}_n(x) \\
 &= \frac{2}{N}A_n \frac{n+1}{n+1+\alpha} \tilde{Q}_{n+1}(x) + \frac{2}{N}(A_n + C_n) \tilde{Q}_n(x) - \tilde{Q}_n(x) + \frac{2}{N}C_n \frac{n+\alpha}{n} \tilde{Q}_{n-1}(x) \\
 &= \frac{n+1}{n+1+\alpha} \frac{n+2\alpha+1}{2n+2\alpha+1} \left(1 - \frac{n}{N}\right) \tilde{Q}_{n+1}(x) + \frac{n+\alpha}{n} \frac{n(n+2\alpha+N+1)}{N(2n+2\alpha+1)} \tilde{Q}_{n-1}(x) \\
 &= \alpha_n \left(1 - \frac{n}{N}\right) \tilde{Q}_{n+1}(x) + \gamma_n \left(1 + \frac{n+2\alpha+1}{N}\right) \tilde{Q}_{n-1}(x).
 \end{aligned} \tag{6.8}$$

Wir können nun mit Verwendung von Lemma 6.1 und der Gleichung (6.5) den folgenden Satz beweisen:

**Satz 6.3.** Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha+1)(2\alpha+2N+1)}. \tag{6.9}$$

Dann gelten die beiden folgenden Beziehungen:

$$(-1)^n \binom{n+\alpha}{n} Q_n \left( \frac{N}{2}(1+x); \alpha, \alpha, N \right) = P_n^{\alpha, \alpha}(x) + \mathcal{O} \left( \frac{n^2}{N} \right), \tag{6.10}$$

mit  $n \leq n(\alpha, N)$ , für  $x \in [-1, 1]$  und  $-\frac{1}{2} < \alpha < 1$ .

$$(-1)^n \binom{n+\alpha}{n} Q_n \left( \frac{N}{2}(1+x); \alpha, \alpha, N \right) = P_n^{\alpha, \alpha}(x) + \mathcal{O} \left( \frac{n^{1+\alpha}}{N} \right), \tag{6.11}$$

mit  $n \leq n(\alpha, N)$ , für  $x \in [-1, 1]$  und  $1 \leq \alpha$ .

*Beweis.* Aus den Gleichungen (3.16) - (3.19) und (6.8) erhalten wir die Rekursionsgleichungen

$$xP_n^{\alpha, \alpha}(x) = \alpha_n P_{n+1}^{\alpha, \alpha}(x) + \beta_n P_n^{\alpha, \alpha}(x) + \gamma_n P_{n-1}^{\alpha, \alpha}(x) \tag{6.12}$$

sowie

$$x\tilde{Q}_n(x) = \alpha_n \left(1 - \frac{n}{N}\right) \tilde{Q}_{n+1}(x) + \gamma_n \left(1 + \frac{n+2\alpha+1}{N}\right) \tilde{Q}_{n-1}(x), \tag{6.13}$$

mit den Konstanten

$$\alpha_n = \frac{(n+1)(n+2\alpha+1)}{(2n+2\alpha+1)(n+\alpha+1)}, \tag{6.14}$$

$$\beta_n = 0, \tag{6.15}$$

$$\gamma_n = \frac{n+\alpha}{2n+2\alpha+1}. \tag{6.16}$$

Vergleicht man die beiden Rekursionsgleichungen, so ergibt sich der zusätzliche Term

$$-\alpha_n \frac{n}{N} \tilde{Q}_{n+1}(x) + \gamma_n \frac{n+2\alpha+1}{N} \tilde{Q}_{n-1}(x). \quad (6.17)$$

Wir untersuchen nun für welches Verhältnis  $N/n$  dieser verschwindet. Zunächst gelten

$$|\alpha_n| = \left| \frac{(n+1)(n+2\alpha+1)}{(2n+2\alpha+1)(n+\alpha+1)} \right| \leq \left| \frac{n+1}{n+\alpha+1} \right| \leq \left| \frac{n+1}{n+1/2} \right| \leq 2, \quad (6.18)$$

und

$$|\gamma_n| = \left| \frac{n+\alpha}{2n+2\alpha+1} \right| \leq \left| \frac{n}{2n+2\alpha+1} \right| + \left| \frac{\alpha}{2n+2\alpha+1} \right| \leq \frac{1}{2} + \frac{|\alpha|}{2\alpha+1}. \quad (6.19)$$

Weiterhin gilt mit Lemma 6.1

$$\max_{x \in [-1,1]} |\tilde{Q}_n(x)| \leq \binom{n+\alpha}{n} \max_{x \in [-1,1]} \left| Q_n \left( \frac{N}{2}(1+x); \alpha, \alpha, N \right) \right| \leq \binom{n+\alpha}{n} \quad (6.20)$$

und mit Anmerkung 2.15 folgt

$$\max_{x \in [-1,1]} |\tilde{Q}_n(x)| = \mathcal{O}(n^\alpha). \quad (6.21)$$

Wir erhalten insgesamt für den untersuchten zusätzliche Term (6.17)

$$\max_{x \in [-1,1]} \left| -\alpha_n \frac{n}{N} \tilde{Q}_{n+1}(x) + \gamma_n \frac{n+2\alpha+1}{N} \tilde{Q}_{n-1}(x) \right| = \mathcal{O} \left( \frac{n^{1+\alpha}}{N} \right). \quad (6.22)$$

Aufgrund der Eindeutigkeit der Lösungen der Rekursionsgleichungen ergibt sich die Behauptung.  $\square$

**Anmerkung 6.4.** *Vergleichen wir diese Aussage nochmals mit Gleichung (6.5), in der  $n$  fest war, so haben wir nun ein Verhältnis  $N/n$  für die Konvergenz der Hahn-Polynome zu den Jacobi-Polynomen gefunden.*

- *Im Fall  $-\frac{1}{2} < \alpha < 1$  können wir die Konvergenz garantieren, wenn die Anzahl der Stützstellen  $N+1$  schneller als  $n^2$  wächst.*
- *Im Fall  $1 \leq \alpha$  können wir die Konvergenz garantieren, wenn die Anzahl der Stützstellen  $N+1$  schneller als  $n^{1+\alpha}$  wächst.*

Im Folgenden beweisen wir mit den Lemmata, Sätzen und Korollaren 6.5 - 6.15 Resultate, die eher von technischer Natur sind und die uns wichtige Eigenschaften der Jacobi-Polynome und Hahn-Polynome für den Beweis unseres Hauptergebnisses bereitstellen.

Das nächste Lemma gibt zunächst das Normierungsverhältnis der Polynome an.

**Lemma 6.5.** *Es gilt*

$$\begin{aligned} & \langle Q_n(\cdot; \alpha, \beta, N), Q_n(\cdot; \alpha, \beta, N) \rangle_\omega \\ &= \frac{n!n!\Gamma(n + \alpha + \beta + 2 + N)\Gamma(\alpha + 1)(N - n)!}{\Gamma(\beta + 1)\Gamma(1 + \alpha + n)\Gamma(1 + \alpha + n)N!N!2^{\alpha+\beta+1}} \left( P_n^{\alpha,\beta}, P_n^{\alpha,\beta} \right)_\varrho. \end{aligned} \quad (6.23)$$

*Beweis.* Mit den Gleichungen (3.15) und (3.47) gilt

$$\begin{aligned} & \frac{\langle Q_n(\cdot; \alpha, \beta, N), Q_n(\cdot; \alpha, \beta, N) \rangle_\omega}{\left( P_n^{\alpha,\beta}, P_n^{\alpha,\beta} \right)_\varrho} \\ &= \frac{(-1)^n(n + \alpha + \beta + 1)_{N+1}(\beta + 1)_n n!}{(2n + \alpha + \beta + 1)(\alpha + 1)_n (-N)_n N!} \cdot \frac{(2n + \alpha + \beta + 1)n!\Gamma(n + \alpha + \beta + 1)}{2^{\alpha+\beta+1}\Gamma(n + \alpha + 1)\Gamma(n + \beta + 1)} \\ &= \frac{n!n!\Gamma(n + \alpha + \beta + 2 + N)\Gamma(\alpha + 1)(N - n)!}{\Gamma(\beta + 1)\Gamma(1 + \alpha + n)\Gamma(1 + \alpha + n)N!N!2^{\alpha+\beta+1}}. \end{aligned} \quad (6.24)$$

□

Nun ermitteln wir das Verhältnis zwischen den Gewichtsfunktionen  $\omega$  und  $\varrho$ .

**Lemma 6.6.** *Für  $\alpha = \beta$  gilt*

$$\omega\left(\frac{N}{2}(1+x)\right) = \frac{N^{2\alpha}}{2^{2\alpha}\Gamma(\alpha+1)\Gamma(\alpha+1)} \left(1 + \mathcal{O}\left(\frac{1}{N}\right)\right) \varrho(x), \quad (6.25)$$

für  $x \in [-1, 1]$ .

*Beweis.* Sei  $\alpha = \beta$ . Dann gilt für jedes  $x \in [-1, 1]$

$$\begin{aligned} & \frac{\omega\left(\frac{N}{2}(1+x)\right)}{\varrho(x)} \\ &= \binom{\alpha + \frac{N}{2}(1+x)}{\frac{N}{2}(1+x)} \binom{\alpha + N - \frac{N}{2}(1+x)}{N - \frac{N}{2}(1+x)} (1-x)^{-\alpha} (1+x)^{-\alpha} \\ &= \frac{\Gamma\left(\alpha + \frac{N}{2}(1+x) + 1\right) \Gamma\left(\alpha + N - \frac{N}{2}(1+x) + 1\right) (1-x)^{-\alpha} (1+x)^{-\alpha}}{\Gamma\left(\frac{N}{2}(1+x) + 1\right) \Gamma(\alpha + 1) \Gamma\left(N - \frac{N}{2}(1+x) + 1\right) \Gamma(\alpha + 1)} \\ &= \frac{\Gamma\left(\alpha + 1 + \frac{N}{2}(1+x)\right) \Gamma\left(\alpha + 1 + \frac{N}{2}(1-x)\right)}{\Gamma\left(\frac{N}{2}(1+x) + 1\right) \Gamma(\alpha + 1) \Gamma\left(\frac{N}{2}(1-x) + 1\right) \Gamma(\alpha + 1) (1-x)^\alpha (1+x)^\alpha}. \end{aligned} \quad (6.26)$$

Mit Lemma 2.10 folgt

$$\begin{aligned} & \frac{\omega\left(\frac{N}{2}(1+x)\right)}{\varrho(x)} \\ & \stackrel{(2.22)}{=} \frac{\left(\frac{N}{2}(1+x)\right)^\alpha \left(\frac{N}{2}(1-x)\right)^\alpha}{\Gamma(\alpha+1)\Gamma(\alpha+1)(1-x)^\alpha(1+x)^\alpha} \left(1 + \mathcal{O}\left(\frac{1}{N}\right)\right) \\ &= \frac{N^{2\alpha}}{2^{2\alpha}\Gamma(\alpha+1)\Gamma(\alpha+1)} \left(1 + \mathcal{O}\left(\frac{1}{N}\right)\right). \end{aligned} \quad (6.27)$$

□

Im nächsten Lemma verwenden wir die totale Variation einer reellwertigen Funktion  $f : [a, b] \rightarrow \mathbb{R}$  gemäß Definition 2.3. Das Lemma gibt uns eine Abschätzung der totalen Variation der Jacobi-Polynome  $P_n^{\alpha, \alpha}$ .

**Lemma 6.7.** *Für  $\alpha \geq 0$  gilt*

$$\mathcal{V}[P_n^{\alpha, \alpha}] \leq 2n \binom{n + \alpha}{n}. \quad (6.28)$$

*Beweis.* Sei  $\alpha \geq 0$ . Dann gilt für die Ableitung

$$\frac{d}{dx} P_n^{\alpha, \alpha}(x) = \frac{2\alpha + n + 1}{2} P_{n-1}^{\alpha+1, \alpha+1}(x), \quad (6.29)$$

siehe Gleichung (3.22). Das Polynom  $P_{n-1}^{\alpha+1, \alpha+1}$  hat nach Anmerkung 3.4 genau  $n-1$  einfache Nullstellen  $z_i$ ,  $i = 1, \dots, n-1$  in dem Intervall  $(-1, 1)$ . Damit hat das Polynom  $P_n^{\alpha, \alpha}$  genau  $n-1$  lokale Extrema in dem Intervall  $(-1, 1)$ . Also ist  $P_n^{\alpha, \alpha}$  jeweils auf den  $n$  Teilintervallen  $[-1, z_1]$ ,  $[z_i, z_{i+1}]$ ,  $i = 1, \dots, n-2$  und  $[z_{n-1}, 1]$  monoton. Seien nun  $z_0 := -1$  und  $z_n := 1$ . Mit Satz 3.8 folgt dann

$$\max_{x \in [-1, 1]} |P_n^{\alpha, \alpha}(x)| = \binom{n + \alpha}{n} \quad (6.30)$$

und daraus erhalten wir

$$\begin{aligned} \mathcal{V}[P_n^{\alpha, \alpha}] &\leq \sum_{i=0}^{n-1} \mathcal{V}[P_n^{\alpha, \alpha}|_{[z_i, z_{i+1}]}] \\ &\leq \sum_{i=0}^{n-1} |P_n^{\alpha, \alpha}(z_{i+1}) - P_n^{\alpha, \alpha}(z_i)| \\ &\leq \sum_{i=0}^{n-1} |P_n^{\alpha, \alpha}(z_{i+1})| + |P_n^{\alpha, \alpha}(z_i)| \\ &\leq \sum_{i=0}^{n-1} 2 \max_{x \in [-1, 1]} |P_n^{\alpha, \alpha}(x)| \\ &\leq 2n \max_{x \in [-1, 1]} |P_n^{\alpha, \alpha}(x)| \\ &\leq 2n \binom{n + \alpha}{n}. \end{aligned} \quad (6.31)$$

□

Durch das folgende Lemma ergibt sich eine weitere wichtige Abschätzung der totalen Variation der Produkte der Jacobi-Polynome  $P_n^{\alpha, \alpha}$  mit deren zugehöriger Gewichtsfunktion  $\varrho$ .

**Lemma 6.8.** *Für  $\alpha > 0$  gilt*

$$\mathcal{V}[P_n^{\alpha, \alpha} \varrho] \leq 2(n+1) \max_{x \in [-1, 1]} |P_n^{\alpha, \alpha}(x) \varrho(x)|. \quad (6.32)$$

*Beweis.* Sei  $\alpha > 0$ . Dann gilt nach Gleichung (3.22) für die Ableitung:

$$\begin{aligned}
 & \frac{d}{dx} [P_n^{\alpha,\alpha} \varrho](x) \\
 &= (1-x^2)^\alpha \frac{d}{dx} P_n^{\alpha,\alpha}(x) + P_n^{\alpha,\alpha}(x) \frac{d}{dx} (1-x^2)^\alpha \\
 &= (1-x^2)^\alpha \frac{2\alpha+n+1}{2} P_{n-1}^{\alpha+1,\alpha+1}(x) - P_n^{\alpha,\alpha}(x) 2x\alpha (1-x^2)^{\alpha-1} \\
 &= (1-x^2)^{\alpha-1} \left[ \frac{2\alpha+n+1}{2} (1-x^2) P_{n-1}^{\alpha+1,\alpha+1}(x) - 2x\alpha P_n^{\alpha,\alpha}(x) \right] \\
 &=: (1-x^2)^{\alpha-1} \tilde{P}(x).
 \end{aligned} \tag{6.33}$$

Da  $\tilde{P} \in \mathcal{P}_{n+1}$  ein Polynom von höchstens dem Grad  $n+1$  ist, besitzt es höchstens  $k \leq n+1$  Nullstellen  $z_i$ ,  $i = 1, \dots, k$  in dem Intervall  $(-1, 1)$ . Dadurch hat  $P_n^{\alpha,\alpha} \varrho$  höchstens  $l \leq k \leq n+1$  lokale Extrema im Intervall  $(-1, 1)$ . Also ist  $P_n^{\alpha,\alpha} \varrho$  jeweils auf den  $l+1$  Teilintervallen  $[-1, z_1]$ ,  $[z_i, z_{i+1}]$ ,  $i = 1, \dots, l-1$  und  $[z_l, 1]$  monoton. Seien nun  $z_0 := -1$  und  $z_{l+1} := 1$ . Dann gilt

$$\begin{aligned}
 \mathcal{V} [P_n^{\alpha,\alpha} \varrho] &\leq \sum_{i=0}^l \mathcal{V} [P_n^{\alpha,\alpha} \varrho|_{[z_i, z_{i+1}]}] \\
 &\leq \sum_{i=0}^l |P_n^{\alpha,\alpha}(z_{i+1}) \varrho(z_{i+1}) - P_n^{\alpha,\alpha}(z_i) \varrho(z_i)| \\
 &\leq \sum_{i=0}^l |P_n^{\alpha,\alpha}(z_{i+1}) \varrho(z_{i+1})| + |P_n^{\alpha,\alpha}(z_i) \varrho(z_i)| \\
 &\leq |P_n^{\alpha,\alpha}(z_0) \varrho(z_0)| + |P_n^{\alpha,\alpha}(z_{l+1}) \varrho(z_{l+1})| + \sum_{i=1}^l 2 |P_n^{\alpha,\alpha}(z_i) \varrho(z_i)| \\
 &\leq |P_n^{\alpha,\alpha}(-1) \varrho(-1)| + |P_n^{\alpha,\alpha}(1) \varrho(1)| + \sum_{i=1}^l 2 \max_{x \in [-1, 1]} |P_n^{\alpha,\alpha}(x) \varrho(x)| \\
 &\leq \sum_{i=1}^l 2 \max_{x \in [-1, 1]} |P_n^{\alpha,\alpha}(x) \varrho(x)| \\
 &\leq 2l \max_{x \in [-1, 1]} |P_n^{\alpha,\alpha}(x) \varrho(x)| \\
 &\leq 2(n+1) \max_{x \in [-1, 1]} |P_n^{\alpha,\alpha}(x) \varrho(x)|.
 \end{aligned} \tag{6.34}$$

□

Der folgende Satz stellt scharfe Abschätzungen für die Maxima der Produkte der Jacobi-Polynome  $P_n^{\alpha,\alpha}$  mit deren zugehöriger Gewichtsfunktion  $\varrho$  bereit.

**Satz 6.9** (vgl. auch [27, S. 362-363]). *Für  $\alpha \geq \frac{1}{2}$  gilt*

$$\max_{x \in [-1, 1]} |P_n^{\alpha,\alpha}(x) \varrho(x)| \leq \left| \frac{\Gamma(2\alpha+1) \Gamma(n+\alpha+1) \Gamma\left(\frac{n}{2} + \alpha + \frac{1}{2}\right)}{\Gamma(\alpha+1) \Gamma\left(\alpha + \frac{1}{2}\right) \Gamma(n+2\alpha+1) \Gamma\left(\frac{n}{2} + 1\right)} \right|. \tag{6.35}$$

Für  $\frac{1}{2} \geq \alpha > 0$  gilt

$$\max_{x \in [-1, 1]} |P_n^{\alpha, \alpha}(x) \varrho(x)| \leq \left| \frac{\Gamma(\alpha) \Gamma(2\alpha + 1) \Gamma(n + \alpha + 1)}{\sqrt{\pi} \Gamma\left(\alpha + \frac{1}{2}\right) \Gamma(\alpha + 1) \Gamma(n + 2\alpha + 1)} \right|. \quad (6.36)$$

*Beweis.* Im Folgenden benötigen wir die Gegenbauer-Polynome

$$P_n^\lambda(x) := \frac{\Gamma\left(\lambda + \frac{1}{2}\right) \Gamma(n + 2\lambda)}{\Gamma(2\lambda) \Gamma\left(n + \lambda + \frac{1}{2}\right)} P_n^{\lambda - \frac{1}{2}, \lambda - \frac{1}{2}}(x), \quad (6.37)$$

aus Definition 3.9 und die Abschätzungen

$$\left[ P_n^\lambda(x) \right]^2 \leq (1 - x^2)^{1-2\lambda} \left[ \frac{\Gamma\left(\frac{n}{2} + \lambda\right)}{\Gamma(\lambda) \Gamma\left(\frac{n}{2} + 1\right)} \right]^2, \quad \text{für } \lambda \geq 1, \quad (6.38)$$

und

$$\left[ P_n^\lambda(x) \right]^2 \leq \frac{1}{\pi} \left[ \frac{\Gamma\left(\lambda - \frac{1}{2}\right)}{\Gamma(\lambda)} \right]^2 (1 - x^2)^{1-2\lambda}, \quad \text{für } \frac{1}{2} < \lambda \leq 1, \quad (6.39)$$

(vgl., z.B., [27, S. 362-363]). Weitere Ergebnisse dieser Form findet man in [32]. Seien zunächst  $\alpha \geq \frac{1}{2}$  und  $\lambda := \alpha + \frac{1}{2}$  gegeben. Dann ist  $\lambda \geq 1$ . Weiter sei  $x \in [-1, 1]$ , dann gilt

$$\begin{aligned} |P_n^{\alpha, \alpha}(x) \varrho(x)| &= \left| P_n^{\lambda - \frac{1}{2}, \lambda - \frac{1}{2}}(x) (1 - x^2)^{\lambda - \frac{1}{2}} \right| \\ &\stackrel{(6.37)}{=} \left| \frac{\Gamma(2\lambda) \Gamma\left(n + \lambda + \frac{1}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right) \Gamma(n + 2\lambda)} P_n^\lambda(x) (1 - x^2)^{\lambda - \frac{1}{2}} \right| \\ &\stackrel{(6.38)}{\leq} \left| \frac{\Gamma(2\lambda) \Gamma\left(n + \lambda + \frac{1}{2}\right)}{\Gamma\left(\lambda + \frac{1}{2}\right) \Gamma(n + 2\lambda)} \frac{\Gamma\left(\frac{n}{2} + \lambda\right)}{\Gamma(\lambda) \Gamma\left(\frac{n}{2} + 1\right)} \right| \\ &\leq \left| \frac{\Gamma(2\alpha + 1) \Gamma(n + \alpha + 1)}{\Gamma(\alpha + 1) \Gamma(n + 2\alpha + 1)} \frac{\Gamma\left(\frac{n}{2} + \alpha + \frac{1}{2}\right)}{\Gamma\left(\alpha + \frac{1}{2}\right) \Gamma\left(\frac{n}{2} + 1\right)} \right|. \end{aligned} \quad (6.40)$$

Die rechte Seite dieser Ungleichungskette ist unabhängig von  $x$ , damit erhalten wir (6.35).

Nun seien  $\frac{1}{2} \geq \alpha > 0$  und  $\lambda := \alpha + \frac{1}{2}$  gegeben. Dann ist  $\frac{1}{2} < \lambda \leq 1$ . Weiter sei wieder  $x \in [-1, 1]$ , dann gilt

$$\begin{aligned}
 |P_n^{\alpha,\alpha}(x)\varrho(x)| &= \left| P_n^{\lambda-\frac{1}{2},\lambda-\frac{1}{2}}(x) (1-x^2)^{\lambda-\frac{1}{2}} \right| \\
 &\stackrel{(6.37)}{=} \left| \frac{\Gamma(2\lambda)\Gamma\left(n+\lambda+\frac{1}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)\Gamma(n+2\lambda)} P_n^\lambda(x) (1-x^2)^{\lambda-\frac{1}{2}} \right| \\
 &\stackrel{(6.39)}{\leq} \left| \frac{\Gamma(2\lambda)\Gamma\left(n+\lambda+\frac{1}{2}\right)}{\Gamma\left(\lambda+\frac{1}{2}\right)\Gamma(n+2\lambda)} \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\lambda-\frac{1}{2}\right)}{\Gamma(\lambda)} \right| \\
 &\leq \left| \frac{\Gamma(2\alpha+1)\Gamma(n+\alpha+1)}{\Gamma(\alpha+1)\Gamma(n+2\alpha+1)} \frac{1}{\sqrt{\pi}} \frac{\Gamma(\alpha)}{\Gamma\left(\alpha+\frac{1}{2}\right)} \right|.
 \end{aligned} \tag{6.41}$$

Die rechte Seite dieser Ungleichungskette ist wieder unabhängig von  $x$ , damit erhalten wir (6.36).  $\square$

Aus den vorherigen Ergebnissen lassen sich scharfe Abschätzungen entnehmen, die wir später anwenden. Allerdings sind diese für spätere Betrachtungen teilweise sehr unhandlich und wir benötigen zum Teil nur das asymptotische Verhalten. Die folgenden Korollare geben Aufschluss über das asymptotische Verhalten der Abschätzungen.

Im ersten Fall erhalten wir als Folgerung aus Satz 6.9:

**Korollar 6.10.** Für  $\alpha \geq \frac{1}{2}$  gilt

$$\max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| = \mathcal{O}\left(n^{-\frac{1}{2}}\right). \tag{6.42}$$

Für  $\frac{1}{2} \geq \alpha > 0$  gilt

$$\max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| = \mathcal{O}\left(n^{-\alpha}\right). \tag{6.43}$$

*Beweis.* Sei zunächst  $\alpha \geq \frac{1}{2}$ . Mit Satz 6.9 gilt

$$\max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| \leq \left| \frac{\Gamma(2\alpha+1)\Gamma(n+\alpha+1)\Gamma\left(\frac{n}{2}+\alpha+\frac{1}{2}\right)}{\Gamma(\alpha+1)\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(n+2\alpha+1)\Gamma\left(\frac{n}{2}+1\right)} \right|. \tag{6.44}$$

Wir wenden Lemma 2.10 auf die Quotienten

$$\left| \frac{\Gamma(n+\alpha+1)}{\Gamma(n+2\alpha+1)} \right| \quad \text{und} \quad \left| \frac{\Gamma\left(\frac{n}{2}+\alpha+\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)} \right| \tag{6.45}$$

an und erhalten

$$\begin{aligned}
 \max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| &\leq \left| \frac{\Gamma(2\alpha+1)\Gamma(n+\alpha+1)\Gamma\left(\frac{n}{2}+\alpha+\frac{1}{2}\right)}{\Gamma(\alpha+1)\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(n+2\alpha+1)\Gamma\left(\frac{n}{2}+1\right)} \right| \\
 &\stackrel{(2.22)}{=} \mathcal{O}\left(n^{\alpha+1-(2\alpha+1)}\right) \mathcal{O}\left(\left(\frac{n}{2}\right)^{\alpha+\frac{1}{2}-1}\right) \\
 &= \mathcal{O}\left(n^{-\frac{1}{2}}\right).
 \end{aligned} \tag{6.46}$$

Nun sei  $\frac{1}{2} \geq \alpha > 0$ . Dann gilt mit Satz 6.9

$$\max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| \leq \left| \frac{\Gamma(\alpha)\Gamma(2\alpha+1)\Gamma(n+\alpha+1)}{\sqrt{\pi}\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(\alpha+1)\Gamma(n+2\alpha+1)} \right|. \quad (6.47)$$

Auch hier wenden wir Lemma 2.10 auf den Quotienten

$$\left| \frac{\Gamma(n+\alpha+1)}{\Gamma(n+2\alpha+1)} \right| \quad (6.48)$$

an und erhalten

$$\begin{aligned} \max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| &\leq \left| \frac{\Gamma(\alpha)\Gamma(2\alpha+1)\Gamma(n+\alpha+1)}{\sqrt{\pi}\Gamma\left(\alpha+\frac{1}{2}\right)\Gamma(\alpha+1)\Gamma(n+2\alpha+1)} \right| \\ &\stackrel{(2.22)}{=} \mathcal{O}\left(n^{\alpha+1-(2\alpha+1)}\right) \\ &= \mathcal{O}\left(n^{-\alpha}\right). \end{aligned} \quad (6.49)$$

□

Im zweiten Fall erhalten wir als Folgerung aus Lemma 6.8:

**Korollar 6.11.** Für  $\alpha \geq \frac{1}{2}$  gilt

$$\mathcal{V}[P_n^{\alpha,\alpha}\varrho] = \mathcal{O}\left(n^{\frac{1}{2}}\right). \quad (6.50)$$

Für  $\frac{1}{2} \geq \alpha > 0$  gilt

$$\mathcal{V}[P_n^{\alpha,\alpha}\varrho] = \mathcal{O}\left(n^{1-\alpha}\right). \quad (6.51)$$

*Beweis.* Sei zunächst  $\alpha \geq \frac{1}{2}$ . Mit Lemma 6.8 folgt

$$\mathcal{V}[P_n^{\alpha,\alpha}\varrho] \leq 2(n+1) \max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)|. \quad (6.52)$$

Somit folgt mit Korollar 6.10

$$\max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| = \mathcal{O}\left(n^{-\frac{1}{2}}\right). \quad (6.53)$$

Daraus erhalten wir (6.50).

Sei nun  $\frac{1}{2} \geq \alpha > 0$ . Mit Lemma 6.8 folgt in diesem Fall

$$\mathcal{V}[P_n^{\alpha,\alpha}\varrho] \leq 2(n+1) \max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)|. \quad (6.54)$$

Schließlich folgt mit Korollar 6.10

$$\max_{x \in [-1,1]} |P_n^{\alpha,\alpha}(x)\varrho(x)| = \mathcal{O}\left(n^{-\alpha}\right). \quad (6.55)$$

Wir erhalten (6.51). □



Wir fahren nun mit einem Lemma fort, welches wir für kommende Berechnungen zur Verfügung stellen.

**Lemma 6.12.** Für  $n \leq \frac{N}{2}$  gilt

$$\left| 1 - \frac{N!N!N}{\Gamma(n+2+N)(N-n)!} \right| \leq \frac{1+2n^2}{N+1}. \quad (6.56)$$

*Beweis.* Der Fall  $n = 0$  ist trivial. Für beliebiges  $1 \leq n \leq \frac{N}{2}$  gilt

$$\begin{aligned} \left| 1 - \frac{N!N!N}{\Gamma(n+2+N)(N-n)!} \right| &= \left| 1 - \frac{N}{N+1} \frac{N!(N+1)!}{(n+1+N)!(N-n)!} \right| \\ &= \left| 1 - \frac{N}{N+1} \prod_{k=0}^{n-1} \frac{N-k}{N+k+2} \right| \\ &= 1 - \frac{N}{N+1} \prod_{k=0}^{n-1} \frac{N-k}{N+k+2} \\ &\leq 1 - \frac{N}{N+1} \prod_{k=0}^{n-1} \frac{N-n+1}{N+n+1} \\ &= 1 - \frac{N}{N+1} \left( \frac{N-n+1}{N+n+1} \right)^n \\ &\leq 1 - \frac{N}{N+1} \left( \frac{N-2n}{N} \right)^n \\ &= 1 - \frac{N}{N+1} \left( 1 + \frac{-2n}{N} \right)^n. \end{aligned} \quad (6.57)$$

Wir verwenden nun die Bernoulli-Ungleichung:

$$(1+x)^m \geq 1+mx, \quad (6.58)$$

für alle  $m \in \mathbb{N}$  und  $x \geq -1$  (vgl., z.B., [41, S. 62]). Für  $x := \frac{-2n}{N}$  gilt  $x \geq -1$ , da laut Voraussetzung  $n \leq \frac{N}{2}$  ist. Nach der Bernoulli-Ungleichung ergibt dies

$$\left( 1 + \frac{-2n}{N} \right)^n \geq 1 + n \frac{-2n}{N}. \quad (6.59)$$

Wir erhalten insgesamt

$$\begin{aligned} \left| 1 - \frac{N!N!N}{\Gamma(n+2+N)(N-n)!} \right| &\leq 1 - \frac{N}{N+1} \left( 1 + n \frac{-2n}{N} \right) \\ &= \frac{1}{N+1} + 2 \frac{n^2}{N+1}. \end{aligned} \quad (6.60)$$

□

Unmittelbar aus diesem Lemma erhält man folgende Aussage:

**Lemma 6.13.** Für  $n \leq \frac{N}{2}$  gilt

$$\frac{N!N!N}{\Gamma(n+2+N)(N-n)!} = 1 + \mathcal{O}\left(\frac{n^2}{N}\right). \quad (6.61)$$

*Beweis.* Nach Lemma 6.12 gilt für beliebiges  $n \leq \frac{N}{2}$

$$\left| 1 - \frac{N!N!N}{\Gamma(n+2+N)(N-n)!} \right| \leq \frac{1+2n^2}{N+1} = \mathcal{O}\left(\frac{n^2}{N}\right). \quad (6.62)$$

□

Das folgende elementare Lemma benötigen wir ebenfalls für spätere Betrachtungen. Es ist eine Rechenregel in Bezug auf die totale Variation des Produkts zweier reellwertiger Funktionen von beschränkter Variation.

**Lemma 6.14.** Für alle Funktionen  $f, g \in \mathcal{BV}[-1, 1]$  gilt

$$\mathcal{V}[fg] \leq \max_{x \in [-1, 1]} |f(x)| \mathcal{V}[g] + \max_{x \in [-1, 1]} |g(x)| \mathcal{V}[f]. \quad (6.63)$$

*Beweis.* Seien  $f, g \in \mathcal{BV}[-1, 1]$ . Weiter sei  $\{x_1, \dots, x_K\}$  eine Zerlegung des Intervalls  $[-1, 1]$ , d. h. es ist

$$-1 = x_1 < x_2 < \dots < x_i < x_{i+1} < \dots < x_{K-1} < x_K = 1. \quad (6.64)$$

Dann gilt für jedes  $i \in \{1, \dots, K-1\}$

$$\begin{aligned} & |f(x_{i+1})g(x_{i+1}) - f(x_i)g(x_i)| \\ & \leq |f(x_{i+1})g(x_{i+1}) - f(x_{i+1})g(x_i)| + |f(x_{i+1})g(x_i) - f(x_i)g(x_i)| \\ & \leq \max_{x \in [-1, 1]} |f(x)| |g(x_{i+1}) - g(x_i)| + \max_{x \in [-1, 1]} |g(x)| |f(x_{i+1}) - f(x_i)|. \end{aligned} \quad (6.65)$$

Die Summation über  $i = 1, \dots, K-1$  auf der linken und rechten Seite der Ungleichungskette führt zu

$$\begin{aligned} & \sum_{i=1}^{K-1} |f(x_{i+1})g(x_{i+1}) - f(x_i)g(x_i)| \\ & \leq \max_{x \in [-1, 1]} |f(x)| \sum_{i=1}^{K-1} |g(x_{i+1}) - g(x_i)| + \max_{x \in [-1, 1]} |g(x)| \sum_{i=1}^{K-1} |f(x_{i+1}) - f(x_i)| \\ & \leq \max_{x \in [-1, 1]} |f(x)| \mathcal{V}[g] + \max_{x \in [-1, 1]} |g(x)| \mathcal{V}[f]. \end{aligned} \quad (6.66)$$

Da die Zerlegung beliebig gewählt war, folgt die Behauptung. □

Dem folgenden Lemma entnehmen wir eine Fehlerabschätzung für das Trapezverfahren, welches wir in Definition 2.24 eingeführt haben. Das Lemma folgt unmittelbar aus Satz 2.26 durch das Setzen von  $a = -1$  und  $b = 1$ . Diese Fehlerabschätzung wenden wir in kommenden Berechnungen an.

**Lemma 6.15** (vgl., z.B., [21, S. 218, (7.14)]). Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und alle  $N \in \mathbb{N}$  gilt

$$\left| \int_{-1}^1 f(x) dx - \frac{2}{N} \sum_{i=1}^{N-1} f\left(\frac{2i}{N} - 1\right) - \frac{1}{N} f(-1) - \frac{1}{N} f(1) \right| \leq \frac{1}{N} \mathcal{V}[f]. \quad (6.67)$$

Bis hierhin haben wir einige Resultate bereitgestellt, die eher von technischer Natur sind. Sie stellen unter anderem wichtige Eigenschaften der Jacobi-Polynome und Hahn-Polynome zusammen.

Im Folgenden betrachten wir zunächst den Fall  $\alpha = \beta = 0$  und im Anschluss den Fall  $\alpha = \beta > 0$ . An einigen Stellen unterscheiden sich diese beiden Fälle im Beweis.

Wir betrachten zuerst den Fall  $\alpha = \beta = 0$  und schreiben der Einfachheit halber kurz  $P_n \equiv P_n^{0,0}$  sowie  $Q_n(x) \equiv Q_n(x, 0, 0, N)$ . Die Gewichtsfunktionen sind in diesem Fall  $\varrho \equiv 1$  und  $\omega \equiv 1$ . Wir schreiben also auch für die inneren Produkte  $(\cdot, \cdot)_\varrho$  und  $\langle \cdot, \cdot \rangle_\omega$  nur kurz  $(\cdot, \cdot)$  und  $\langle \cdot, \cdot \rangle$ . Um die einzelnen Berechnungen einfacher nachvollziehen zu können, schreiben wir darüber hinaus  $\langle f, g \rangle$  für innere Produkte der Gestalt  $\langle f\left(\frac{2}{N}(\cdot) - 1\right), g\left(\frac{2}{N}(\cdot) - 1\right) \rangle$  zwischen zwei Funktionen  $f : [-1, 1] \rightarrow \mathbb{R}$  und  $g : [-1, 1] \rightarrow \mathbb{R}$ .

Im Folgenden sei

$$A_{n,N} := \frac{(-1)^n N! N! 2}{\Gamma(n+2+N)(N-n)!}. \quad (6.68)$$

**Lemma 6.16.** Sei für  $N \in \mathbb{N}$

$$n(N) := \frac{1}{2} + \frac{1}{2} \sqrt{(2N+1)}. \quad (6.69)$$

Für jedes  $f \in \mathcal{BV}[-1, 1]$  gilt

$$\left| \frac{2}{N} \langle f, P_n \rangle Q_n\left(\frac{N}{2}(1+x)\right) - \langle f, Q_n \rangle Q_n\left(\frac{N}{2}(1+x)\right) A_{n,N} \right| = \mathcal{O}\left(\frac{n^2}{N}\right), \quad (6.70)$$

mit  $n \leq n(N)$ , für jedes  $x \in [-1, 1]$ .

*Beweis.* Sei  $f \in \mathcal{BV}[-1, 1]$  und sei  $x \in [-1, 1]$ . Dann gilt für beliebiges  $n \leq n(N)$

$$\begin{aligned} & \left| \frac{2}{N} \langle f, P_n \rangle Q_n\left(\frac{N}{2}(1+x)\right) - \langle f, Q_n \rangle Q_n\left(\frac{N}{2}(1+x)\right) A_{n,N} \right| \\ & \leq \left| Q_n\left(\frac{N}{2}(1+x)\right) \right| \left| \frac{2}{N} \langle f, P_n \rangle - \langle f, Q_n \rangle \frac{(-1)^n N! N! 2}{\Gamma(n+2+N)(N-n)!} \right|. \end{aligned} \quad (6.71)$$

Mit Lemma 6.1 folgt

$$\begin{aligned}
 & \left| \frac{2}{N} \langle f, P_n \rangle Q_n \left( \frac{N}{2}(1+x) \right) - \langle f, Q_n \rangle Q_n \left( \frac{N}{2}(1+x) \right) A_{n,N} \right| \\
 \stackrel{(6.2)}{\leq} & \left| \frac{2}{N} \langle f, P_n \rangle - \langle f, Q_n \rangle \frac{(-1)^n N! N! 2}{\Gamma(n+2+N)(N-n)!} \right| \\
 & \leq \frac{2}{N} \left| \langle f, P_n \rangle - \langle f, (-1)^n Q_n \rangle \frac{N! N! N}{\Gamma(n+2+N)(N-n)!} \right| \\
 & \leq \frac{2}{N} \left| \langle f, P_n - (-1)^n Q_n \rangle - \langle f, (-1)^n Q_n \rangle \left( \frac{N! N! N}{\Gamma(n+2+N)(N-n)!} - 1 \right) \right| \\
 & \leq \max_{x \in [-1,1]} |f(x)| \frac{2}{N} \left[ \langle 1, |P_n - (-1)^n Q_n| \rangle + \langle 1, |(-1)^n Q_n| \rangle \left| (-1)^n \frac{N}{2} A_{n,N} - 1 \right| \right].
 \end{aligned} \tag{6.72}$$

Wir wenden nun Lemma 6.13 und anschließend Satz 6.3 an und erhalten

$$\begin{aligned}
 & \left| \frac{2}{N} \langle f, P_n \rangle Q_n \left( \frac{N}{2}(1+x) \right) - \langle f, Q_n \rangle Q_n \left( \frac{N}{2}(1+x) \right) A_{n,N} \right| \\
 \stackrel{(6.61)}{\leq} & \max_{x \in [-1,1]} |f(x)| \frac{2}{N} \left[ \langle 1, |P_n - (-1)^n Q_n| \rangle + \langle 1, 1 \rangle \mathcal{O} \left( \frac{n^2}{N} \right) \right] \\
 \stackrel{(6.10)}{\leq} & \max_{x \in [-1,1]} |f(x)| \frac{2}{N} \left[ \mathcal{O} \left( \frac{n^2}{N} \right) \langle 1, 1 \rangle + \langle 1, 1 \rangle \mathcal{O} \left( \frac{n^2}{N} \right) \right] \\
 & \leq \max_{x \in [-1,1]} |f(x)| \frac{2}{N} \left[ \mathcal{O} \left( \frac{n^2}{N} \right) (N+1) + (N+1) \mathcal{O} \left( \frac{n^2}{N} \right) \right] \\
 & \leq \mathcal{O} \left( \frac{n^2}{N} \right).
 \end{aligned} \tag{6.73}$$

□

**Lemma 6.17.** Sei für  $N \in \mathbb{N}$

$$n(N) := \frac{1}{2} + \frac{1}{2} \sqrt{(2N+1)}. \tag{6.74}$$

Für jedes  $f \in \mathcal{C}[-1,1] \cap \mathcal{BV}[-1,1]$  gilt

$$\left| (f, P_n) P_n(x) - \frac{2}{N} \langle f, P_n \rangle (-1)^n Q_n \left( \frac{N}{2}(1+x) \right) \right| = \mathcal{O} \left( \frac{n^2}{N} \right), \tag{6.75}$$

mit  $n \leq n(N)$ , für jedes  $x \in [-1,1]$ .

*Beweis.* Sei  $f \in \mathcal{C}[-1,1] \cap \mathcal{BV}[-1,1]$  und sei  $x \in [-1,1]$ . Dann gilt für beliebiges

$n \leq n(N)$

$$\begin{aligned}
 & \left| (f, P_n) P_n(x) - \frac{2}{N} \langle f, P_n \rangle (-1)^n Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \leq \left| (f, P_n) P_n(x) - \frac{2}{N} \langle f, P_n \rangle P_n(x) \right| \\
 & \quad + \left| \frac{2}{N} \langle f, P_n \rangle P_n(x) - \frac{2}{N} \langle f, P_n \rangle (-1)^n Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \stackrel{(3.23)}{\leq} \left| (f, P_n) - \frac{2}{N} \langle f, P_n \rangle \right| + \frac{2}{N} |\langle f, P_n \rangle| \left| P_n(x) - (-1)^n Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \stackrel{(6.10)}{\leq} \left| (f, P_n) - \frac{2}{N} \sum_{i=0}^N f \left( \frac{2i}{N} - 1 \right) P_n \left( \frac{2i}{N} - 1 \right) \right| + \frac{2(N+1)}{N} \max_{x \in [-1,1]} |f(x)| \mathcal{O} \left( \frac{n^2}{N} \right) \\
 & \leq \left| (f, P_n) - \frac{2}{N} \sum_{i=1}^{N-1} f \left( \frac{2i}{N} - 1 \right) P_n \left( \frac{2i}{N} - 1 \right) - \frac{1}{N} f(-1) P_n(-1) - \frac{1}{N} f(1) P_n(1) \right| \\
 & \quad + \left| \frac{1}{N} f(-1) P_n(-1) \right| + \left| \frac{1}{N} f(1) P_n(1) \right| + \mathcal{O} \left( \frac{n^2}{N} \right).
 \end{aligned} \tag{6.76}$$

Mit Lemma 6.15 folgt

$$\begin{aligned}
 & \left| (f, P_n) P_n(x) - \frac{2}{N} \langle f, P_n \rangle (-1)^n Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \stackrel{(6.67)}{\leq} \frac{1}{N} \mathcal{V}[f P_n] + \mathcal{O} \left( \frac{1}{N} \right) + \mathcal{O} \left( \frac{1}{N} \right) + \mathcal{O} \left( \frac{n^2}{N} \right) \\
 & \stackrel{(6.63)}{\leq} \frac{1}{N} \left[ \max_{x \in [-1,1]} |f(x)| \mathcal{V}[P_n] + \max_{x \in [-1,1]} |P_n(x)| \mathcal{V}[f] \right] + \mathcal{O} \left( \frac{n^2}{N} \right) \\
 & \stackrel{(6.28)}{\leq} \frac{1}{N} \left[ \max_{x \in [-1,1]} |f(x)| 2n + \mathcal{V}[f] \right] + \mathcal{O} \left( \frac{n^2}{N} \right) \\
 & \leq \mathcal{O} \left( \frac{n}{N} \right) + \mathcal{O} \left( \frac{n^2}{N} \right) \\
 & \leq \mathcal{O} \left( \frac{n^2}{N} \right).
 \end{aligned} \tag{6.77}$$

□

**Lemma 6.18.** Sei für  $N \in \mathbb{N}$

$$n(N) := \frac{1}{2} + \frac{1}{2} \sqrt{(2N+1)}. \tag{6.78}$$

Für jedes  $f \in \mathcal{C}[-1,1] \cap \mathcal{BV}[-1,1]$  gilt

$$\left| \frac{(f, P_n)}{(P_n, P_n)} P_n(x) - \frac{\langle f, Q_n \rangle}{\langle Q_n, Q_n \rangle} Q_n \left( \frac{N}{2}(1+x) \right) \right| = \mathcal{O} \left( \frac{n^3}{N} \right), \tag{6.79}$$

mit  $n \leq n(N)$ , für jedes  $x \in [-1,1]$ .

*Beweis.* Sei  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und sei  $x \in [-1, 1]$ . Dann gilt für beliebiges  $n \leq n(N)$

$$\begin{aligned}
 & \left| \frac{(f, P_n)}{(P_n, P_n)} P_n(x) - \frac{\langle f, Q_n \rangle}{\langle Q_n, Q_n \rangle} Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \stackrel{(6.23)}{=} \frac{1}{|(P_n, P_n)|} \left| (f, P_n) P_n(x) - \langle f, Q_n \rangle Q_n \left( \frac{N}{2}(1+x) \right) (-1)^n A_{n,N} \right| \\
 & \leq \frac{1}{|(P_n, P_n)|} \left| (f, P_n) P_n(x) - \frac{2}{N} \langle f, P_n \rangle (-1)^n Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \quad + \frac{1}{|(P_n, P_n)|} \left| \frac{2}{N} \langle f, P_n \rangle Q_n \left( \frac{N}{2}(1+x) \right) - \langle f, Q_n \rangle Q_n \left( \frac{N}{2}(1+x) \right) A_{n,N} \right|.
 \end{aligned} \tag{6.80}$$

Wir wenden Lemma 6.16 an und erhalten

$$\begin{aligned}
 & \left| \frac{(f, P_n)}{(P_n, P_n)} P_n(x) - \frac{\langle f, Q_n \rangle}{\langle Q_n, Q_n \rangle} Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \stackrel{(6.70)}{\leq} \frac{1}{|(P_n, P_n)|} \left| (f, P_n) P_n(x) - \frac{2}{N} \langle f, P_n \rangle (-1)^n Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \quad + \frac{1}{|(P_n, P_n)|} \mathcal{O} \left( \frac{n^2}{N} \right).
 \end{aligned} \tag{6.81}$$

Schließlich ergibt sich mit Hilfe von Lemma 6.17

$$\begin{aligned}
 & \left| \frac{(f, P_n)}{(P_n, P_n)} P_n(x) - \frac{\langle f, Q_n \rangle}{\langle Q_n, Q_n \rangle} Q_n \left( \frac{N}{2}(1+x) \right) \right| \\
 & \stackrel{(6.75)}{\leq} \frac{1}{|(P_n, P_n)|} \mathcal{O} \left( \frac{n^2}{N} \right) + \frac{1}{|(P_n, P_n)|} \mathcal{O} \left( \frac{n^2}{N} \right) \\
 & \stackrel{(3.15)}{\leq} \frac{2n+1}{2} \mathcal{O} \left( \frac{n^2}{N} \right) \\
 & \leq \mathcal{O} \left( \frac{n^3}{N} \right).
 \end{aligned} \tag{6.82}$$

□

Wir können nun das angekündigte Hauptergebnis für den Fall  $\alpha = 0$  beweisen.

**Satz 6.19.** *Sei für  $N \in \mathbb{N}$*

$$n(N) := \frac{1}{2} + \frac{1}{2} \sqrt{(2N+1)}. \tag{6.83}$$

*Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  gilt*

$$\begin{aligned}
 & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle}{\langle Q_k, Q_k \rangle} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\
 & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)}{(P_k, P_k)} P_k(x) \right| + \mathcal{O} \left( \frac{n^4}{N} \right),
 \end{aligned} \tag{6.84}$$

*mit  $n \leq n(N)$ , für jedes  $x \in [-1, 1]$ .*

*Beweis.* Sei  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und sei  $x \in [-1, 1]$ . Dann gilt für beliebiges  $n \leq n(N)$

$$\begin{aligned}
 & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle}{\langle Q_k, Q_k \rangle} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\
 & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)}{(P_k, P_k)} P_k(x) \right| \\
 & + \left| \sum_{k=0}^n \frac{(f, P_k)}{(P_k, P_k)} P_k(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle}{\langle Q_k, Q_k \rangle} Q_k \left( \frac{N}{2}(1+x) \right) \right| \quad (6.85) \\
 & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)}{(P_k, P_k)} P_k(x) \right| \\
 & + \sum_{k=0}^n \left| \frac{(f, P_k)}{(P_k, P_k)} P_k(x) - \frac{\langle f, Q_k \rangle}{\langle Q_k, Q_k \rangle} Q_k \left( \frac{N}{2}(1+x) \right) \right|.
 \end{aligned}$$

Lemma 6.18 ergibt

$$\begin{aligned}
 & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle}{\langle Q_k, Q_k \rangle} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\
 & \stackrel{(6.79)}{\leq} \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)}{(P_k, P_k)} P_k(x) \right| + \sum_{k=0}^n \mathcal{O} \left( \frac{k^3}{N} \right) \quad (6.86) \\
 & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)}{(P_k, P_k)} P_k(x) \right| + \mathcal{O} \left( \frac{n^4}{N} \right).
 \end{aligned}$$

□

Wir behandeln im Folgenden nun den Fall  $\alpha = \beta > 0$ . Auch hier schreiben wir zur Vereinfachung kurz  $P_n \equiv P_n^{\alpha, \alpha}$  und  $Q_n(x) \equiv Q_n(x, \alpha, \alpha, N)$ . Des Weiteren nutzen wir folgende Abkürzung

$$\tilde{Q}_n(x) := (-1)^n \binom{n+\alpha}{n} Q_n \left( \frac{N}{2}(1+x); \alpha, \alpha, N \right), \quad (6.87)$$

aus Gleichung (6.6).

**Lemma 6.20.** Sei  $\alpha > 0$ . Für jedes  $n \leq \frac{N}{2}$  gilt

$$\frac{N^{2\alpha} N! N! N}{\Gamma(n+2\alpha+2+N)(N-n)!} = 1 + \mathcal{O} \left( \frac{n^2}{N} \right). \quad (6.88)$$

*Beweis.* Sei  $\alpha > 0$ . Für jedes  $n \leq \frac{N}{2}$  gilt

$$\begin{aligned}
 \left| 1 - \frac{N^{2\alpha}\Gamma(n+2+N)}{\Gamma(n+2\alpha+2+N)} \right| &= 1 - \frac{N^{2\alpha}\Gamma(n+2+N)}{\Gamma(n+2\alpha+2+N)} \\
 &= 1 - \frac{N^{2\alpha}}{(n+2+N)^{2\alpha}} \frac{\Gamma(n+2+N)(n+2+N)^{2\alpha}}{\Gamma(n+2\alpha+2+N)} \\
 &\stackrel{(2.22)}{=} 1 - \left( 1 + \mathcal{O}\left(\frac{n^2}{N}\right) \right) \left( 1 + \mathcal{O}\left(\frac{1}{n+N}\right) \right) \\
 &= \mathcal{O}\left(\frac{n^2}{N}\right).
 \end{aligned} \tag{6.89}$$

Mit Lemma 6.12 folgt für jedes  $n \leq \frac{N}{2}$

$$\begin{aligned}
 \frac{N^{2\alpha}N!N!N}{\Gamma(n+2\alpha+2+N)(N-n)!} &= \frac{N^{2\alpha}\Gamma(n+2+N)}{\Gamma(n+2\alpha+2+N)} \frac{N!N!N}{\Gamma(n+2+N)(N-n)!} \\
 &= \left( 1 + \mathcal{O}\left(\frac{n^2}{N}\right) \right) \left( 1 + \mathcal{O}\left(\frac{n^2}{N}\right) \right) \\
 &= 1 + \mathcal{O}\left(\frac{n^2}{N}\right).
 \end{aligned} \tag{6.90}$$

□

Nun sei im Folgenden

$$A_{n,N}^\alpha := \frac{(-1)^n N!N!2^{2\alpha+1}}{\Gamma(n+2\alpha+2+N)(N-n)!}. \tag{6.91}$$

**Lemma 6.21.** Sei  $\alpha > 0$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha+1)(2\alpha+2N+1)}. \tag{6.92}$$

Für jedes  $f \in \mathcal{BV}[-1, 1]$  gilt

$$\begin{aligned}
 &\left| Q_n \left( \frac{N}{2}(1+x) \right) \left[ \frac{2}{N} \langle f, P_n \rangle_\varrho \binom{n+\alpha}{n} - \langle f, Q_n \rangle_\omega \frac{\Gamma^2(\alpha+1+n)}{n!n!} A_{n,N}^\alpha \right] \right| \\
 &= \mathcal{O} \left( \frac{n^{1+\alpha+\max\{1,\alpha\}}}{N} \right),
 \end{aligned} \tag{6.93}$$

mit  $n \leq n(\alpha, N)$ , für jedes  $x \in [-1, 1]$ .

*Beweis.* Sei  $f \in \mathcal{BV}[-1, 1]$  und sei  $x \in [-1, 1]$ . Dann gilt für beliebiges  $n \leq n(\alpha, N)$

$$\begin{aligned}
 & \left| Q_n \left( \frac{N}{2}(1+x) \right) \left[ \frac{2}{N} \langle f, P_n \rangle_{\varrho} \binom{n+\alpha}{n} - \langle f, Q_n \rangle_{\omega} \frac{\Gamma^2(\alpha+1+n)}{n!n!} A_{n,N}^{\alpha} \right] \right| \\
 & \leq \left| Q_n \left( \frac{N}{2}(1+x) \right) \right| \left| \frac{2}{N} \langle f, P_n \rangle_{\varrho} \binom{n+\alpha}{n} - \langle f, Q_n \rangle_{\omega} \frac{\Gamma^2(\alpha+1+n)}{n!n!} A_{n,N}^{\alpha} \right| \\
 & \stackrel{(6.2)}{\leq} \binom{n+\alpha}{n} \left| \frac{2}{N} \langle f, P_n \rangle_{\varrho} - \langle f, Q_n \rangle_{\omega} \frac{\Gamma(\alpha+1+n)\Gamma(\alpha+1)}{n!} A_{n,N}^{\alpha} \right| \tag{6.94} \\
 & \stackrel{(6.25)}{\leq} \binom{n+\alpha}{n} \left| \frac{2}{N} \langle f, P_n \rangle_{\varrho} - \langle f, Q_n \rangle_{\varrho} \frac{N^{2\alpha} \left(1 + \mathcal{O}\left(\frac{1}{N}\right)\right) \Gamma(\alpha+1+n)}{2^{2\alpha} \Gamma(\alpha+1) n!} A_{n,N}^{\alpha} \right| \\
 & \leq \binom{n+\alpha}{n} \left| \frac{2}{N} \langle f, P_n \rangle_{\varrho} - \frac{2}{N} \langle f, \tilde{Q}_n \rangle_{\varrho} \frac{N^{2\alpha} N! N! N \left(1 + \mathcal{O}\left(\frac{1}{N}\right)\right)}{\Gamma(n+2\alpha+2+N)(N-n)!} \right|.
 \end{aligned}$$

Mit Lemma 6.20 folgt

$$\begin{aligned}
 & \left| Q_n \left( \frac{N}{2}(1+x) \right) \left[ \frac{2}{N} \langle f, P_n \rangle_{\varrho} \binom{n+\alpha}{n} - \langle f, Q_n \rangle_{\omega} \frac{\Gamma^2(\alpha+1+n)}{n!n!} A_{n,N}^{\alpha} \right] \right| \\
 & \stackrel{(6.88)}{\leq} \binom{n+\alpha}{n} \left| \frac{2}{N} \langle f, P_n \rangle_{\varrho} - \frac{2}{N} \langle f, \tilde{Q}_n \rangle_{\varrho} \left(1 + \mathcal{O}\left(\frac{n^2}{N}\right)\right) \left(1 + \mathcal{O}\left(\frac{1}{N}\right)\right) \right| \tag{6.95} \\
 & \leq \binom{n+\alpha}{n} \left| \frac{2}{N} \langle f, P_n \rangle_{\varrho} + \frac{2}{N} \langle f, P_n - \tilde{Q}_n - P_n \rangle_{\varrho} \left(1 + \mathcal{O}\left(\frac{n^2}{N}\right)\right) \right| \\
 & \leq \frac{2}{N} \binom{n+\alpha}{n} \left[ \left| \langle f, P_n \rangle_{\varrho} \right| \mathcal{O}\left(\frac{n^2}{N}\right) + \left| \langle f, P_n - \tilde{Q}_n \rangle_{\varrho} \right| \left(1 + \mathcal{O}\left(\frac{n^2}{N}\right)\right) \right].
 \end{aligned}$$

Wir erhalten mit Satz 6.3

$$\begin{aligned}
 & \left| Q_n \left( \frac{N}{2}(1+x) \right) \left[ \frac{2}{N} \langle f, P_n \rangle_{\varrho} \binom{n+\alpha}{n} - \langle f, Q_n \rangle_{\omega} \frac{\Gamma^2(\alpha+1+n)}{n!n!} A_{n,N}^{\alpha} \right] \right| \\
 & \stackrel{(6.10)}{\leq} \frac{2}{N} \binom{n+\alpha}{n} \left| \langle f, P_n \rangle_{\varrho} \right| \mathcal{O}\left(\frac{n^2}{N}\right) \\
 & \quad + \frac{2}{N} \binom{n+\alpha}{n} \max_{x \in [-1,1]} |f(x)\varrho(x)| \mathcal{O}\left(\frac{n^{1+\max\{1,\alpha\}}}{N}\right) \langle 1, 1 \rangle_1 \left(1 + \mathcal{O}\left(\frac{n^2}{N}\right)\right) \\
 & \leq \frac{2}{N} \binom{n+\alpha}{n} \left| \langle f, P_n \rangle_{\varrho} \right| \mathcal{O}\left(\frac{n^2}{N}\right) + \mathcal{O}\left(\frac{n^{1+\alpha+\max\{1,\alpha\}}}{N}\right) \tag{6.96} \\
 & \leq \frac{2}{N} \binom{n+\alpha}{n} \max_{x \in [-1,1]} |f(x)| \max_{x \in [-1,1]} |P_n(x)\varrho(x)| \langle 1, 1 \rangle_1 \mathcal{O}\left(\frac{n^2}{N}\right) \\
 & \quad + \mathcal{O}\left(\frac{n^{1+\alpha+\max\{1,\alpha\}}}{N}\right) \\
 & \leq \max_{x \in [-1,1]} |P_n(x)\varrho(x)| \mathcal{O}\left(\frac{n^{2+\alpha}}{N}\right) + \mathcal{O}\left(\frac{n^{1+\alpha+\max\{1,\alpha\}}}{N}\right).
 \end{aligned}$$

Mit Korollar 6.10 folgt die Behauptung.  $\square$

**Lemma 6.22.** Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (6.97)$$

Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und  $\frac{1}{2} \geq \alpha > 0$  gilt

$$\left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| = \mathcal{O} \left( \frac{n^{2-\alpha}}{N} \right), \quad (6.98)$$

mit  $n \leq n(\alpha, N)$ , für jedes  $x \in [-1, 1]$ .

Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und  $\alpha \geq \frac{1}{2}$  gilt

$$\left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| = \mathcal{O} \left( \frac{n^{\frac{1}{2} + \max\{1, \alpha\}}}{N} \right), \quad (6.99)$$

mit  $n \leq n(\alpha, N)$ , für jedes  $x \in [-1, 1]$ .

*Beweis.* Sei  $\alpha > 0$ . Weiter sei  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und sei  $x \in [-1, 1]$ . Dann gilt für beliebiges  $n \leq n(\alpha, N)$

$$\begin{aligned} & \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\ & \leq \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho P_n(x) \right| \\ & \quad + \left| \frac{2}{N} \langle f, P_n \rangle_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\ & \stackrel{(3.23)}{\leq} \binom{n+\alpha}{n} \left| (f, P_n)_\varrho - \frac{2}{N} \langle f, P_n \rangle_\varrho \right| + \frac{2}{N} |\langle f, P_n \rangle_\varrho| |P_n(x) - \tilde{Q}_n(x)| \\ & \stackrel{(6.67)}{\leq} \binom{n+\alpha}{n} \frac{1}{N} \mathcal{V}[fP_n\varrho] + \frac{2}{N} |\langle f, P_n \rangle_\varrho| |P_n(x) - \tilde{Q}_n(x)|. \end{aligned} \quad (6.100)$$

Mit Satz 6.3 folgt

$$\begin{aligned} & \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\ & \stackrel{(6.10)}{\leq} \stackrel{(6.11)}{\leq} \binom{n+\alpha}{n} \frac{1}{N} \mathcal{V}[fP_n\varrho] \\ & \quad + \frac{2(N+1)}{N} \max_{x \in [-1, 1]} |f(x)| \max_{x \in [-1, 1]} |P_n(x)\varrho(x)| \mathcal{O} \left( \frac{n^{1+\max\{1, \alpha\}}}{N} \right). \end{aligned} \quad (6.101)$$

Die Anwendung von Lemma 6.14 ergibt

$$\begin{aligned}
 & \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\
 & \stackrel{(6.63)}{\leq} \binom{n+\alpha}{n} \frac{1}{N} \left[ \max_{x \in [-1,1]} |f(x)| \mathcal{V}[P_n \varrho] + \max_{x \in [-1,1]} |P_n(x) \varrho(x)| \mathcal{V}[f] \right] \\
 & + \max_{x \in [-1,1]} |P_n(x) \varrho(x)| \mathcal{O} \left( \frac{n^{1+\max\{1,\alpha\}}}{N} \right) \\
 & \leq \left[ \mathcal{V}[P_n \varrho] + \max_{x \in [-1,1]} |P_n(x) \varrho(x)| \right] \mathcal{O} \left( \frac{n^\alpha}{N} \right) \\
 & + \max_{x \in [-1,1]} |P_n(x) \varrho(x)| \mathcal{O} \left( \frac{n^{1+\max\{1,\alpha\}}}{N} \right) \\
 & \leq \mathcal{V}[P_n \varrho] \mathcal{O} \left( \frac{n^\alpha}{N} \right) + \max_{x \in [-1,1]} |P_n(x) \varrho(x)| \mathcal{O} \left( \frac{n^{1+\max\{1,\alpha\}}}{N} \right).
 \end{aligned} \tag{6.102}$$

Damit folgt insbesondere für  $\frac{1}{2} \geq \alpha > 0$

$$\begin{aligned}
 & \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\
 & \leq \mathcal{V}[P_n \varrho] \mathcal{O} \left( \frac{n^\alpha}{N} \right) + \max_{x \in [-1,1]} |P_n(x) \varrho(x)| \mathcal{O} \left( \frac{n^{1+\max\{1,\alpha\}}}{N} \right) \\
 & \stackrel{(6.43)}{\leq} \stackrel{(6.51)}{\mathcal{O}} \left( n^{1-\alpha} \right) \mathcal{O} \left( \frac{n^\alpha}{N} \right) + \mathcal{O} \left( n^{-\alpha} \right) \mathcal{O} \left( \frac{n^{1+\max\{1,\alpha\}}}{N} \right) \\
 & \leq \mathcal{O} \left( \frac{n^{2-\alpha}}{N} \right).
 \end{aligned} \tag{6.103}$$

Außerdem folgt insbesondere für  $\alpha \geq \frac{1}{2}$

$$\begin{aligned}
 & \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\
 & \leq \mathcal{V}[P_n \varrho] \mathcal{O} \left( \frac{n^\alpha}{N} \right) + \max_{x \in [-1,1]} |P_n(x) \varrho(x)| \mathcal{O} \left( \frac{n^{1+\max\{1,\alpha\}}}{N} \right) \\
 & \stackrel{(6.42)}{\leq} \stackrel{(6.50)}{\mathcal{O}} \left( n^{\frac{1}{2}} \right) \mathcal{O} \left( \frac{n^\alpha}{N} \right) + \mathcal{O} \left( n^{-\frac{1}{2}} \right) \mathcal{O} \left( \frac{n^{1+\max\{1,\alpha\}}}{N} \right) \\
 & \leq \mathcal{O} \left( \frac{n^{\frac{1}{2}+\max\{1,\alpha\}}}{N} \right).
 \end{aligned} \tag{6.104}$$

□

**Lemma 6.23.** Sei  $\alpha > 0$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \tag{6.105}$$

Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  gilt

$$\left| \frac{(f, P_n)_\varrho}{(P_n, P_n)_\varrho} P_n(x) - \frac{\langle f, Q_n \rangle_\omega}{\langle Q_n, Q_n \rangle_\omega} Q_n\left(\frac{N}{2}(1+x)\right) \right| = \mathcal{O}\left(\frac{n^{2+\alpha+\max\{1, \alpha\}}}{N}\right), \quad (6.106)$$

mit  $n \leq n(\alpha, N)$ , für jedes  $x \in [-1, 1]$ .

*Beweis.* Sei  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und sei  $x \in [-1, 1]$ . Dann gilt für beliebiges  $n \leq n(\alpha, N)$

$$\begin{aligned} & \left| \frac{(f, P_n)_\varrho}{(P_n, P_n)_\varrho} P_n(x) - \frac{\langle f, Q_n \rangle_\omega}{\langle Q_n, Q_n \rangle_\omega} Q_n\left(\frac{N}{2}(1+x)\right) \right| \\ & \stackrel{(6.23)}{=} \frac{1}{|(P_n, P_n)_\varrho|} \left| (f, P_n)_\varrho P_n(x) - \langle f, Q_n \rangle_\omega \tilde{Q}_n(x) \frac{\Gamma(\alpha+1+n)\Gamma(\alpha+1)}{n!} A_{n,N}^\alpha \right| \\ & \leq \frac{1}{|(P_n, P_n)_\varrho|} \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\ & \quad + \frac{1}{|(P_n, P_n)_\varrho|} \left| \tilde{Q}_n(x) \left[ \frac{2}{N} \langle f, P_n \rangle_\varrho - \langle f, Q_n \rangle_\omega \frac{\Gamma(\alpha+1+n)\Gamma(\alpha+1)}{n!} A_{n,N}^\alpha \right] \right|. \end{aligned} \quad (6.107)$$

Wir wenden Lemma 6.21 an und erhalten

$$\begin{aligned} & \left| \frac{(f, P_n)_\varrho}{(P_n, P_n)_\varrho} P_n(x) - \frac{\langle f, Q_n \rangle_\omega}{\langle Q_n, Q_n \rangle_\omega} Q_n\left(\frac{N}{2}(1+x)\right) \right| \\ & \stackrel{(6.93)}{\leq} \frac{1}{|(P_n, P_n)_\varrho|} \left| (f, P_n)_\varrho P_n(x) - \frac{2}{N} \langle f, P_n \rangle_\varrho \tilde{Q}_n(x) \right| \\ & \quad + \frac{1}{|(P_n, P_n)_\varrho|} \mathcal{O}\left(\frac{n^{1+\alpha+\max\{1, \alpha\}}}{N}\right). \end{aligned} \quad (6.108)$$

Schließlich ergibt sich mit Hilfe von Lemma 6.22

$$\begin{aligned} & \left| \frac{(f, P_n)_\varrho}{(P_n, P_n)_\varrho} P_n(x) - \frac{\langle f, Q_n \rangle_\omega}{\langle Q_n, Q_n \rangle_\omega} Q_n\left(\frac{N}{2}(1+x)\right) \right| \\ & \stackrel{(6.98)}{\leq} \frac{1}{|(P_n, P_n)_\varrho|} \mathcal{O}\left(\frac{n^{1+\max\{1, \alpha\}}}{N}\right) + \frac{1}{|(P_n, P_n)_\varrho|} \mathcal{O}\left(\frac{n^{1+\alpha+\max\{1, \alpha\}}}{N}\right) \\ & \stackrel{(3.15)}{\leq} \frac{(2n+2\alpha+1)n!\Gamma(n+2\alpha+1)}{2^{2\alpha+1}\Gamma(n+\alpha+1)\Gamma(n+\alpha+1)} \mathcal{O}\left(\frac{n^{1+\alpha+\max\{1, \alpha\}}}{N}\right) \\ & \stackrel{(2.22)}{\leq} \mathcal{O}\left(\frac{n^{2+\alpha+\max\{1, \alpha\}}}{N}\right). \end{aligned} \quad (6.109)$$

□

Wir können nun das angekündigte Hauptergebnis für den Fall  $\alpha > 0$  beweisen.

**Satz 6.24.** *Sei  $\alpha > 0$  und sei für  $N \in \mathbb{N}$*

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (6.110)$$

Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  gilt

$$\begin{aligned} & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \right| + \mathcal{O} \left( \frac{n^{3+\alpha+\max\{1,\alpha\}}}{N} \right), \end{aligned} \quad (6.111)$$

mit  $n \leq n(\alpha, N)$ , für jedes  $x \in [-1, 1]$ .

*Beweis.* Sei  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und sei  $x \in [-1, 1]$ . Dann gilt für beliebiges  $n \leq n(\alpha, N)$

$$\begin{aligned} & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \right| \\ & + \left| \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \right| \\ & + \sum_{k=0}^n \left| \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) - \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right|. \end{aligned} \quad (6.112)$$

Lemma 6.23 ergibt

$$\begin{aligned} & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \stackrel{(6.106)}{\leq} \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \right| + \sum_{k=0}^n \mathcal{O} \left( \frac{k^{2+\alpha+\max\{1,\alpha\}}}{N} \right) \\ & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \right| + \mathcal{O} \left( \frac{n^{3+\alpha+\max\{1,\alpha\}}}{N} \right). \end{aligned} \quad (6.113)$$

□

Wir fassen nun die Sätze 6.19 und 6.24 als unser Hauptergebnis zusammen:

**Theorem 6.25.** Sei  $\alpha \geq 0$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (6.114)$$

Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  gilt

$$\begin{aligned} & \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \right| + \mathcal{O} \left( \frac{n^{3+\alpha+\max\{1, \alpha\}}}{N} \right), \end{aligned} \quad (6.115)$$

mit  $n \leq n(\alpha, N)$ , für jedes  $x \in [-1, 1]$ .

## 6.2 Folgerungen aus dem Hauptergebnis

Im Folgenden stellen wir einige Resultate dar, die sich aus der Anwendung von Theorem 6.25 ergeben. Insbesondere diskutieren wir also verschiedene Fälle, in denen wir die punktweise Konvergenz der Methode der kleinsten Quadrate erhalten.

**Korollar 6.26.** Sei  $\alpha \geq 0$ . Für jedes  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und jede Folge  $(N_n)_{n \in \mathbb{N}}$  mit  $N_n^{-1} n^{3+\alpha+\max\{1, \alpha\}} \rightarrow 0$ , konvergiert die Methode der kleinsten Quadrate

$$LS_{N_n}^{N_n}[f](x) = \sum_{k=0}^{N_n} \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N_n}{2}(1+x) \right) \rightarrow f(x), \quad n \rightarrow \infty \quad (6.116)$$

für jedes  $x \in [-1, 1]$ , falls die Reihenentwicklung durch Jacobi-Polynome punktweise gegen  $f(x)$  konvergiert, d. h.

$$\sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(x) \rightarrow f(x), \quad n \rightarrow \infty \quad (6.117)$$

für jedes  $x \in [-1, 1]$ .

Die Reihenentwicklung  $\sum_{k=0}^n \hat{f}_k Q_k$  von einer Funktion  $f$  durch Hahn-Polynome  $Q_k$  konvergiert also punktweise, wenn die Reihenentwicklung  $\sum_{k=0}^n \hat{f}_k P_k$  von der Funktion  $f$  durch Jacobi-Polynome  $P_k$  punktweise konvergiert und wenn  $n^{3+\alpha+\max\{1, \alpha\}}/N \rightarrow 0$  für  $n, N \rightarrow \infty$  gilt. Für etliche Klassen von Funktionen wurde in den letzten Jahrzehnten die Konvergenz der zugehörigen Reihenentwicklungen in Jacobi-Polynomen gezeigt. Wir geben diesbezüglich einige Anwendungsbeispiele für Theorem 6.25 an. Im Folgenden verwenden wir das Stetigkeitsmodul  $\omega$  gemäß Definition 2.6.

S. Agakhanov und G. Natanson zeigten 1966 folgendes Resultat:

**Satz 6.27** (vgl. [3]). Seien  $\alpha, \beta > -\frac{1}{2}$ . Für jede Funktion  $f \in \mathcal{C}[-1, 1]$ , mit

$$\lim_{n \rightarrow \infty} \omega\left(f, [-1, 1], \frac{1}{n}\right) n^{\max\{\alpha, \beta\} + 1/2} = 0, \quad (6.118)$$

konvergiert die Reihenentwicklung in Jacobi-Polynomen auf dem Intervall  $[-1, 1]$  gleichmäßig gegen  $f$ .

Unmittelbar aus Satz 6.27 und unserem Korollar 6.26 folgt:

**Korollar 6.28.** Sei  $\alpha \geq 0$ , sei

$$f \in K := \left\{ g \in \mathcal{C}[-1, 1] : \lim_{n \rightarrow \infty} \omega\left(g, [-1, 1], \frac{1}{n}\right) n^{\alpha + 1/2} = 0 \right\} \quad (6.119)$$

und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge mit

$$\lim_{n \rightarrow \infty} \frac{n^{3 + \alpha + \max\{1, \alpha\}}}{N_n} = 0. \quad (6.120)$$

Dann konvergiert die Methode der kleinsten Quadrate  $LS_{N_n}^{N_n}[f]$  für jedes  $x \in [-1, 1]$ .

**Anmerkung 6.29.** Offenbar gilt für eine Lipschitz-stetige Funktion  $f : [-1, 1] \rightarrow \mathbb{R}$ , mit der Lipschitz-Konstanten

$$L := \sup_{x, t \in [-1, 1]} \frac{|f(x) - f(t)|}{|x - t|} \quad (6.121)$$

(vgl., z.B., [71, S. 186]), die Abschätzung  $\omega(f, [-1, 1], \delta) \leq L\delta$ . Also ist für jede Lipschitz-stetige Funktion  $f$  die Folge  $\left(\omega\left(f, [-1, 1], \frac{1}{n}\right) n\right)_n$  beschränkt. Damit erfüllt jede Lipschitz-stetige Funktion  $f : [-1, 1] \rightarrow \mathbb{R}$  im Fall  $0 \leq \alpha < \frac{1}{2}$  die Voraussetzung aus Korollar 6.28.

V. Badkov bewies 1968 folgendes Ergebnis:

**Satz 6.30** (vgl. [8]). Seien  $-1 < \alpha \leq -\frac{1}{2}$  und  $-1 < \beta \leq -\frac{1}{2}$ . Für jede Funktion  $f \in \mathcal{C}[-1, 1]$ , mit

$$\lim_{n \rightarrow \infty} \omega\left(f, [-1, 1], \frac{1}{n}\right) \ln(n) = 0, \quad (6.122)$$

konvergiert die Reihenentwicklung in Jacobi-Polynomen auf dem Intervall  $[-1, 1]$  gleichmäßig gegen  $f$ .

A. Belen'kii hat 1989 folgendes Ergebnis nachgewiesen:

**Satz 6.31** (vgl. [10]). Sei  $\alpha, \beta > -1$ . Für jede Funktion  $f \in \mathcal{C}[-1, 1]$ , mit

$$\lim_{n \rightarrow \infty} \omega\left(f, [-1, 1], \frac{1}{n}\right) \ln(n) = 0 \quad (6.123)$$

und einer am Rand von  $[-1, 1]$  konvergenten Folge

$$\left( \sum_{k=0}^n \frac{(f, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(\pm 1) \right)_n, \quad (6.124)$$

konvergiert die Reihenentwicklung in Jacobi-Polynomen auf dem Intervall  $[-1, 1]$  gleichmäßig gegen  $f$ .

Aus Satz 6.31 und unserem Korollar 6.26 erhalten wir:

**Korollar 6.32.** Sei  $\alpha \geq 0$  und seien

$$K_1 := \left\{ g \in \mathcal{C}[-1, 1] : \lim_{n \rightarrow \infty} \omega \left( g, [-1, 1], \frac{1}{n} \right) \ln(n) = 0 \right\}, \quad (6.125)$$

$$K_2 := \left\{ g \in \mathcal{C}[-1, 1] : \left( \sum_{k=0}^n \frac{(g, P_k)_\varrho}{(P_k, P_k)_\varrho} P_k(\pm 1) \right)_n \text{ konvergiert} \right\}. \quad (6.126)$$

Sei  $f \in K := K_1 \cap K_2$  und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge mit

$$\lim_{n \rightarrow \infty} \frac{n^{3+\alpha+\max\{1, \alpha\}}}{N_n} = 0. \quad (6.127)$$

Dann konvergiert die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  für jedes  $x \in [-1, 1]$ .

Im Spezialfall  $\alpha = \beta = 0$  reduzieren sich die Jacobi-Polynome zu den Legendre-Polynomen, siehe dazu Definition 3.9. In diesem Fall können wir Jackson's Theorem verwenden:

**Satz 6.33** (vgl. [42]). Für jede Funktion  $f \in \mathcal{C}^1[-1, 1]$ , mit  $f' \in \mathcal{BV}[-1, 1]$ , konvergiert die Reihenentwicklung in Legendre-Polynomen auf dem Intervall  $[-1, 1]$  gleichmäßig gegen  $f$ .

Mit Anwendung von Satz 6.33 und unserem Korollar 6.26 ergibt sich:

**Korollar 6.34.** Sei  $f \in K := \{g \in \mathcal{C}^1[-1, 1] : g' \in \mathcal{BV}[-1, 1]\}$  und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge mit

$$\lim_{n \rightarrow \infty} \frac{n^4}{N_n} = 0. \quad (6.128)$$

Dann konvergiert für  $\alpha = 0$  die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  für jedes  $x \in [-1, 1]$ .

Die Klasse von Funktionen wird größer, wenn man nur das offene Intervall  $(-1, 1)$  betrachtet. M. Powierska bewies 2007 das Folgende:

**Satz 6.35** (vgl. [64]). Für jede Funktion  $f \in \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$ , konvergiert die Reihenentwicklung in Legendre-Polynomen auf dem Intervall  $(-1, 1)$  punktweise gegen  $f$ .



Schließlich ergibt sich aus Satz 6.35 und unserem Korollar 6.26:

**Korollar 6.36.** *Sei  $f \in K := \mathcal{C}[-1, 1] \cap \mathcal{BV}[-1, 1]$  und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge mit*

$$\lim_{n \rightarrow \infty} \frac{n^4}{N_n} = 0. \quad (6.129)$$

*Dann konvergiert für  $\alpha = 0$  die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  für jedes  $x \in (-1, 1)$ .*

Weitere Ergebnisse, auf die sich Theorem 6.25 anwenden lässt, findet man z.B. in [17], [50], [61], [64].



# 7 Gleichmäßige Konvergenz der Methode der kleinsten Quadrate

In diesem Kapitel beschäftigen wir uns mit der Frage nach der gleichmäßigen Konvergenz der Methode der kleinsten Quadrate. Dabei gehen wir insbesondere auf den in Unterabschnitt 4.3.2 erläuterten diskreten Fall ein. Den kontinuierlichen Fall aus Unterabschnitt 4.3.1 behandeln wir in Abschnitt 7.3 und vergleichen dort auch beide Fälle. Das Hauptergebnis dieses Kapitels ist Theorem 7.7. Dieses Ergebnis stellt eine unverbesserbare Abschätzung der Methode auf einem äquidistanten Gitter mit  $N + 1$  Stützstellen auf dem Intervall  $[-1, 1]$  für die Menge der dort  $(n + 1)$ -mal stetig differenzierbaren Funktionen bereit. Dazu gehen wir wieder von der gleichen Ausgangssituation wie im vorherigen Kapitel 6 aus und beschreiben den Operator der Methode der kleinsten Quadrate  $LS_n^N$  wieder gemäß Gleichung (4.65), für den symmetrischen Fall  $\alpha = \beta$ .

Wir untersuchen folgendes Problem: Für welche Funktionen  $f \in K \subset C[-1, 1]$  und welches Verhältnis  $N/n$  konvergiert die Folge  $(LS_n^N[f])$  gleichmäßig? Anschließend vergleichen wir unsere Approximationsresultate mit entsprechenden Resultaten bekannter Approximationsmethoden.

## 7.1 Hauptergebnis zur gleichmäßigen Konvergenz

Grundlegend für unsere Untersuchungen ist das folgende von H. Brass bewiesene Resultat:

**Satz 7.1** (vgl. [20]). *Sei  $d\sigma$  eine Distribution auf  $[-1, 1]$  und sei*

$$\{q_k : k = 0, \dots, n + 1\} \tag{7.1}$$

*eine Familie von Orthogonalpolynomen, die orthogonal bezüglich des inneren Produkts*

$$(f, g)_\sigma := \int_{-1}^1 f(x)g(x)d\sigma(x) \tag{7.2}$$

*sind. Weiterhin seien diese durch  $(q_k, q_k)_\sigma = 1$  normiert. Sei die Distribution  $d\sigma$  so definiert, dass die folgenden Eigenschaften gelten:*

- $\int_{-1}^1 f(x)d\sigma(x) = \int_{-1}^1 f(-x)d\sigma(x)$  für jedes  $f \in C[-1, 1]$ ,

- $\sup_{x \in [-1,1]} |q_k(x)| = q_k(1)$  für jedes  $k = 0, \dots, n+1$ .

Sei

$$C_n := \frac{\sup_{x \in [-1,1]} |q_{n+1}(x)|}{\sup_{x \in [-1,1]} |q_{n+1}^{(n+1)}(x)|}. \quad (7.3)$$

Dann gilt für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$

$$\sup_{x \in [-1,1]} \left| f(x) - \sum_{k=0}^n (f, q_k)_\sigma q_k(x) \right| \leq C_n \sup_{x \in [-1,1]} |f^{(n+1)}(x)|. \quad (7.4)$$

Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $C_n$  in Ungleichung (7.4) durch keine kleinere Zahl ersetzt werden kann.

Um dieses Ergebnis anwenden zu können, müssen wir im Folgenden eine entsprechende Distribution und die passende Familie von Orthogonalpolynomen definieren. Als erstes verwenden wir dafür die Darstellung der Hahn-Polynome durch hypergeometrische Reihen gemäß Definition 3.16. Das erste Lemma gibt uns ein Verhältnis  $N/n$  für die Beschränktheit der Hahn-Polynome. Des Weiteren sehen wir, dass die Maxima am Rand angenommen werden.

**Lemma 7.2.** Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (7.5)$$

Dann gilt für beliebiges  $n \leq n(\alpha, N)$

$$\max_{x \in [0, N]} |Q_n(x; \alpha, \alpha, N)| = Q_n(0; \alpha, \alpha, N) = (-1)^n Q_n(N; \alpha, \alpha, N) = 1. \quad (7.6)$$

*Beweis.* Unmittelbar aus dem von H. Dette bewiesenen Resultat [26], siehe auch Satz 6.2, folgt

$$\max_{x \in [0, N]} |Q_n(x; \alpha, \alpha, N)| = \max \{ |Q_n(0; \alpha, \alpha, N)|, |Q_n(N; \alpha, \alpha, N)| \} = 1. \quad (7.7)$$

Weiterhin gilt folgende Symmetrieeigenschaft (vgl., z.B., [26]):

$$Q_n(0; \alpha, \alpha, N) = (-1)^n Q_n(N; \alpha, \alpha, N). \quad (7.8)$$

Mit der Definition 3.16 der Hahn-Polynome ergibt sich die Positivität von  $Q_n(0; \alpha, \alpha, N)$ :

$$Q_n(0; \alpha, \alpha, N) = \sum_{k=0}^n \frac{(-n)_k (n + \alpha + \alpha + 1)_k (0)_k}{(\alpha + 1)_k (-N)_k} \frac{1}{k!} = 1. \quad (7.9)$$

□

**Anmerkung 7.3.** *Im Folgenden sei*

$$\hat{Q}_k(x) := \frac{(-1)^k Q_k\left(\frac{N}{2}(1+x); \alpha, \alpha, N\right)}{\sqrt{\langle Q_k(\cdot; \alpha, \alpha, N), Q_k(\cdot; \alpha, \alpha, N) \rangle_\omega}}. \quad (7.10)$$

Sei  $N \in \mathbb{N}$  und sei  $\alpha = \beta > -\frac{1}{2}$ . Des Weiteren betrachten wir in diesem Abschnitt die durch

$$\int_{-1}^1 f(x) d\sigma(x) = \sum_{i=0}^N f\left(-1 + \frac{2i}{N}\right) \omega(i) \quad (7.11)$$

gegebene Distribution  $d\sigma$  auf  $[-1, 1]$ .

Im nächsten Lemma weisen wir alle erforderlichen Eigenschaften der Polynome  $\hat{Q}_k$  nach, die im Satz 7.1 vorausgesetzt werden.

**Lemma 7.4.** *Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$*

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (7.12)$$

Sei  $N \in \mathbb{N}$  und sei  $d\sigma$  die Distribution aus Anmerkung 7.3. Darüber hinaus sei  $\{\hat{Q}_k : k = 0, \dots, N\}$  die Familie von Polynomen, wie in Anmerkung 7.3 definiert. Dann gelten mit  $n + 1 \leq n(\alpha, N)$  die folgenden Eigenschaften:

1.  $\{\hat{Q}_k : k = 0, \dots, n + 1\}$  ist eine Familie von Orthogonalpolynomen, die orthogonal bezüglich des inneren Produkts  $(f, g)_\sigma := \int_{-1}^1 f(x)g(x)d\sigma(x)$  sind.
2.  $(\hat{Q}_k, \hat{Q}_k)_\sigma = 1$ .
3.  $\int_{-1}^1 f(x)d\sigma(x) = \int_{-1}^1 f(-x)d\sigma(x)$  für jedes  $f \in \mathcal{C}[-1, 1]$ .
4.  $\sup_{x \in [-1, 1]} |\hat{Q}_k(x)| = \hat{Q}_k(1)$  für jedes  $k = 0, \dots, n + 1$ .

*Beweis.* Für alle  $k, l \in \{1, \dots, n + 1\}$  gilt

$$(\hat{Q}_k, \hat{Q}_l)_\sigma = \int_{-1}^1 \hat{Q}_k(x) \hat{Q}_l(x) d\sigma(x) = \sum_{i=0}^N \hat{Q}_k\left(-1 + \frac{2i}{N}\right) \hat{Q}_l\left(-1 + \frac{2i}{N}\right) \omega(i). \quad (7.13)$$

Mit der Definition von  $\{\hat{Q}_k : k = 0, \dots, N\}$  folgt

$$(\hat{Q}_k, \hat{Q}_l)_\sigma = \frac{(-1)^{k+l} \langle Q_k, Q_l \rangle_\omega}{\sqrt{\langle Q_k, Q_k \rangle_\omega \langle Q_l, Q_l \rangle_\omega}}. \quad (7.14)$$

Damit erhalten wir Eigenschaft 1.

Mit  $k = l$  gilt  $(\hat{Q}_k, \hat{Q}_k)_\sigma = 1$  und wir erhalten Eigenschaft 2.

Weiterhin gilt für jedes  $f \in \mathcal{C}[-1, 1]$

$$\int_{-1}^1 f(x) d\sigma(x) = \sum_{i=0}^N f\left(-1 + \frac{2i}{N}\right) \binom{\alpha + i}{i} \binom{\alpha + N - i}{N - i}. \quad (7.15)$$

Mit der Indextransformation  $i \mapsto N - i$  und der Gleichung

$$-1 + \frac{2(N - i)}{N} = 1 - \frac{2i}{N}, \quad (7.16)$$

folgt

$$\int_{-1}^1 f(x) d\sigma(x) = \sum_{i=0}^N f\left(1 - \frac{2i}{N}\right) \binom{\alpha + N - i}{N - i} \binom{\alpha + i}{i} = \int_{-1}^1 f(-x) d\sigma(x). \quad (7.17)$$

Damit ist Eigenschaft 3 bewiesen.

Mit Lemma 7.2 gilt für jedes  $k = 0, \dots, n + 1$

$$\sup_{x \in [-1, 1]} |\hat{Q}_k(x)| = \sup_{x \in [0, N]} |Q_k(x)| \frac{1}{\sqrt{\langle Q_k, Q_k \rangle_\omega}} = \frac{(-1)^k Q_k(N)}{\sqrt{\langle Q_k, Q_k \rangle_\omega}} = \hat{Q}_k(1), \quad (7.18)$$

womit wir Eigenschaft 4 erhalten. □

Nun können wir Satz 7.1 anwenden.

**Lemma 7.5.** Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (7.19)$$

Sei  $N \in \mathbb{N}$  und sei  $d\sigma$  die Distribution aus Anmerkung 7.3. Darüber hinaus sei  $\{\hat{Q}_k : k = 0, \dots, N\}$  die Familie von Polynomen, wie in Anmerkung 7.3 definiert. Weiter sei

$$D_{n, N} := \frac{\sup_{x \in [-1, 1]} |\hat{Q}_{n+1}(x)|}{\sup_{x \in [-1, 1]} |\hat{Q}_{n+1}^{(n+1)}(x)|}. \quad (7.20)$$

Dann gilt für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$  mit  $n + 1 \leq n(\alpha, N)$

$$\sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n (f, \hat{Q}_k)_\sigma \hat{Q}_k(x) \right| \leq D_{n, N} \sup_{x \in [-1, 1]} |f^{(n+1)}(x)|. \quad (7.21)$$

Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $D_{n, N}$  in Ungleichung (7.21) durch keine kleinere Zahl ersetzt werden kann.

*Beweis.* Wir wenden Satz 7.1 und Lemma 7.4 an. Wegen Lemma 7.4, erfüllt die Familie von Orthogonalpolynomen  $\{\hat{Q}_k : k = 0, \dots, N\}$  alle Voraussetzungen von Satz 7.1. Also können wir Satz 7.1 anwenden und erhalten die Behauptung.  $\square$

Im folgenden Lemma ermitteln wir den Faktor  $D_{n,N}$  aus dem vorherigen Lemma 7.5, der in Gleichung (7.20) definiert ist.

**Lemma 7.6.** *Unter den Voraussetzungen von Lemma 7.5 gilt für beliebiges  $n + 1 \leq n(\alpha, N)$*

$$D_{n,N} = \frac{\sup_{x \in [-1,1]} |\hat{Q}_{n+1}(x)|}{\sup_{x \in [-1,1]} |\hat{Q}_{n+1}^{(n+1)}(x)|} = \frac{2^{n+1} \Gamma(n + 2\alpha + 2) \Gamma(n + \alpha + 2)}{(n + 1)! \Gamma(2n + 2\alpha + 3) \Gamma(\alpha + 1)} \frac{N!}{N^{n+1} (N - n - 1)!}. \quad (7.22)$$

*Beweis.* Zunächst gilt für jedes  $x \in [-1, 1]$  und beliebiges  $n + 1 \leq n(\alpha, N)$

$$\begin{aligned} \hat{Q}_{n+1}^{(n+1)}(x) &= \frac{(-1)^{n+1}}{\sqrt{\langle Q_{n+1}, Q_{n+1} \rangle_\omega}} \frac{d^{n+1}}{dx^{n+1}} Q_{n+1} \left( \frac{N}{2} (1 + x) \right) \\ &= \frac{(-1)^{n+1}}{\sqrt{\langle Q_{n+1}, Q_{n+1} \rangle_\omega}} \left( \frac{N}{2} \right)^{n+1} Q_{n+1}^{(n+1)} \left( \frac{N}{2} (1 + x) \right). \end{aligned} \quad (7.23)$$

Mit Lemma 7.2 folgt

$$\sup_{x \in [-1,1]} |\hat{Q}_{n+1}(x)| = \frac{\sup_{x \in [0,N]} |Q_{n+1}(x)|}{\sqrt{\langle Q_{n+1}, Q_{n+1} \rangle_\omega}} = \frac{1}{\sqrt{\langle Q_{n+1}, Q_{n+1} \rangle_\omega}}. \quad (7.24)$$

Dann gilt

$$\frac{\sup_{x \in [-1,1]} |\hat{Q}_{n+1}(x)|}{\sup_{x \in [-1,1]} |\hat{Q}_{n+1}^{(n+1)}(x)|} = \left( \frac{2}{N} \right)^{n+1} \frac{1}{\sup_{x \in [0,N]} |Q_{n+1}^{(n+1)}(x)|}. \quad (7.25)$$

Mit der Darstellung der Hahn-Polynome gemäß Definition 3.16 folgt

$$Q_{n+1}^{(n+1)}(x) = \frac{(-n - 1)_{n+1} (n + 2\alpha + 2)_{n+1}}{(\alpha + 1)_{n+1} (-N)_{n+1}} \frac{1}{(n + 1)!} \frac{d^{n+1}}{dx^{n+1}} (-x)_{n+1}. \quad (7.26)$$

Der Term  $(-x)_{n+1}$  ist ein Polynom vom Grad  $n + 1$ , welches wir durch

$$(-x)_{n+1} = (-x)(-x + 1) \cdot \dots \cdot (-x + n) = (-1)^{n+1} x^{n+1} + \tilde{p}(x) \quad (7.27)$$

zerlegen können. Dabei ist  $\tilde{p} \in \mathcal{P}_n$  ein Polynom vom Grad  $n$ . Daher folgt

$$Q_{n+1}^{(n+1)}(x) = \frac{(-n - 1)_{n+1} (n + 2\alpha + 2)_{n+1}}{(\alpha + 1)_{n+1} (-N)_{n+1}} \frac{1}{(n + 1)!} (-1)^{n+1} (n + 1)!. \quad (7.28)$$

Mit den Umformungen

$$(-n-1)_{n+1} = (-n-1)(-n) \cdot \dots \cdot (-1) = (-1)^{n+1}(n+1)! \quad (7.29)$$

und

$$(-N)_{n+1} = (-N)(-N+1) \cdot \dots \cdot (-N+n) = (-1)^{n+1} \frac{N!}{(N-n-1)!} \quad (7.30)$$

erhalten wir

$$\begin{aligned} Q_{n+1}^{(n+1)}(x) &= \frac{(-1)^{n+1}(n+1)! (n+2\alpha+2)_{n+1} (N-n-1)!}{(\alpha+1)_{n+1} (-1)^{n+1} N!} (-1)^{n+1} \\ &= \frac{(-1)^{n+1}(n+1)! \Gamma(2n+2\alpha+3) \Gamma(\alpha+1) (N-n-1)!}{\Gamma(n+2\alpha+2) \Gamma(n+\alpha+2) N!}. \end{aligned} \quad (7.31)$$

Durch Einsetzen in Gleichung (7.25) erhält man schließlich

$$\frac{\sup_{x \in [-1,1]} |\hat{Q}_{n+1}(x)|}{\sup_{x \in [-1,1]} |\hat{Q}_{n+1}^{(n+1)}(x)|} = \left(\frac{2}{N}\right)^{n+1} \frac{\Gamma(n+2\alpha+2) \Gamma(n+\alpha+2) N!}{(n+1)! \Gamma(2n+2\alpha+3) \Gamma(\alpha+1) (N-n-1)!}, \quad (7.32)$$

woraus Gleichung (7.22) folgt.  $\square$

Nun können wir mit Hilfe der vorherigen Lemmata das Hauptergebnis dieses Kapitels beweisen.

**Theorem 7.7.** Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha+1)(2\alpha+2N+1)}. \quad (7.33)$$

Weiter sei

$$D_{n,N} := \frac{2^{n+1} \Gamma(n+2\alpha+2) \Gamma(n+\alpha+2) N!}{(n+1)! \Gamma(2n+2\alpha+3) \Gamma(\alpha+1) N^{n+1} (N-n-1)!}. \quad (7.34)$$

Dann gilt für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$  mit  $n+1 \leq n(\alpha, N)$

$$\sup_{x \in [-1,1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \leq D_{n,N} \sup_{x \in [-1,1]} |f^{(n+1)}(x)|. \quad (7.35)$$

Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $D_{n,N}$  in Ungleichung (7.35) durch keine kleinere Zahl ersetzt werden kann.

*Beweis.* Sei  $N \in \mathbb{N}$  und sei  $f \in \mathcal{C}^{n+1}[-1, 1]$  mit  $n+1 \leq n(\alpha, N)$ . Dann gilt für jedes  $x \in [-1, 1]$

$$\sum_{k=0}^n (f, \hat{Q}_k)_\sigma \hat{Q}_k(x) = \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right). \quad (7.36)$$

Nun wenden wir Lemma 7.5 und Lemma 7.6 an und erhalten daraus

$$\begin{aligned} & \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{2^{n+1} \Gamma(n+2\alpha+2) \Gamma(n+\alpha+2) N!}{(n+1)! \Gamma(2n+2\alpha+3) \Gamma(\alpha+1) N^{n+1} (N-n-1)!}. \end{aligned} \quad (7.37)$$

Die Unverbesserbarkeit der Abschätzung erhält man aus Lemma 7.5.  $\square$

## 7.2 Folgerungen aus dem Hauptergebnis

Im Folgenden präsentieren wir einige Resultate, die sich aus der Anwendung von Theorem 7.7 ergeben. Insbesondere diskutieren wir verschiedene Fälle, in denen wir die gleichmäßige Konvergenz der Methode der kleinsten Quadrate erhalten. Zunächst untersuchen wir den Faktor  $D_{n,N}$  aus Theorem 7.7. Hierzu stellen wir die folgenden Lemmata bereit.

**Lemma 7.8.** *Sei  $\alpha > -\frac{1}{2}$ . Dann gilt*

$$\frac{\Gamma(n+2\alpha+2) \Gamma(n+\alpha+2)}{\Gamma(2n+2\alpha+3)} = \frac{(n+1)!(n+1)! n^\alpha}{(2n+2)! 2^{2\alpha}} \left( 1 + \mathcal{O}(n^{-1}) \right). \quad (7.38)$$

*Beweis.* Wir wenden Lemma 2.10 an und erhalten

$$\begin{aligned} \frac{\Gamma(n+2\alpha+2) \Gamma(n+\alpha+2)}{\Gamma(2n+2\alpha+3)} &= \frac{\Gamma(n+1)n^{(2\alpha+2)-1} \Gamma(n+1)n^{(\alpha+2)-1}}{\Gamma(2n+1)(2n)^{(2\alpha+3)-1}} \left( 1 + \mathcal{O}(n^{-1}) \right) \\ &= \frac{n!n!}{(2n)!} \frac{n^\alpha}{2^{2\alpha+2}} \left( 1 + \mathcal{O}(n^{-1}) \right) \\ &= \frac{(n+1)!(n+1)! n^\alpha}{(2n+2)! 2^{2\alpha}} \left( 1 + \mathcal{O}(n^{-1}) \right). \end{aligned} \quad (7.39)$$

$\square$

**Lemma 7.9.** *Es gilt*

$$\frac{\sqrt{\pi n}}{2^n n!} e^{\frac{2}{12n+1} - \frac{1}{24n}} \leq \frac{2^n n!}{(2n)!} \leq \frac{\sqrt{\pi n}}{2^n n!} e^{\frac{1}{6n} - \frac{1}{24n+1}}. \quad (7.40)$$

*Beweis.* Wir beweisen die beiden Ungleichungen einzeln und verwenden dabei die verallgemeinerte Version der Stirling-Formel (vgl., z.B., [31, S. 50-53], [67])

$$e^{\frac{1}{12n+1}} \leq \frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} \leq e^{\frac{1}{12n}}. \quad (7.41)$$

Zuerst zeigen wir die erste Ungleichung. Mit Hilfe der Stirling-Formel erhalten wir zunächst

$$\begin{aligned} \frac{2^n n!}{(2n)!} &\geq 2^n \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}}}{\sqrt{2\pi(2n)} \left(\frac{2n}{e}\right)^{2n} e^{\frac{1}{24n}}} \\ &= \frac{\sqrt{\pi n}}{2^n} \frac{1}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} \frac{e^{\frac{1}{12n+1}}}{e^{\frac{1}{24n}}}. \end{aligned} \quad (7.42)$$

Eine wiederholte Anwendung der Stirling-Formel ergibt

$$\frac{2^n n!}{(2n)!} \geq \frac{\sqrt{\pi n}}{2^n} \frac{1}{n!} \frac{e^{\frac{1}{12n+1}} e^{\frac{1}{12n+1}}}{e^{\frac{1}{24n}}} = \frac{\sqrt{\pi n}}{2^n n!} e^{\frac{2}{12n+1} - \frac{1}{24n}}. \quad (7.43)$$

Damit haben wir die erste Ungleichung gezeigt. In Bezug auf die zweite Ungleichung erhalten wir wieder mit Hilfe der Stirling-Formel

$$\begin{aligned} \frac{2^n n!}{(2n)!} &\leq 2^n \frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}}{\sqrt{2\pi(2n)} \left(\frac{2n}{e}\right)^{2n} e^{\frac{1}{24n+1}}} \\ &= \frac{\sqrt{\pi n}}{2^n} \frac{1}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} \frac{e^{\frac{1}{12n}}}{e^{\frac{1}{24n+1}}}. \end{aligned} \quad (7.44)$$

Eine wiederholte Anwendung ergibt

$$\frac{2^n n!}{(2n)!} \leq \frac{\sqrt{\pi n}}{2^n} \frac{1}{n!} \frac{e^{\frac{1}{12n}} e^{\frac{1}{12n}}}{e^{\frac{1}{24n+1}}} = \frac{\sqrt{\pi n}}{2^n n!} e^{\frac{1}{6n} - \frac{1}{24n+1}}. \quad (7.45)$$

Somit ist auch die zweite Ungleichung bewiesen. □

**Anmerkung 7.10.** Mit Lemma 7.9 erhält man

$$\frac{2^n n!}{(2n)!} = \frac{\sqrt{\pi n}}{2^n n!} \left(1 + \mathcal{O}(n^{-1})\right). \quad (7.46)$$

Wir können nun die „unhandliche“ Abschätzung aus Theorem 7.7 vereinfachen.

**Korollar 7.11.** Sei  $\alpha > -\frac{1}{2}$  und sei für  $N \in \mathbb{N}$

$$n(\alpha, N) := \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}. \quad (7.47)$$

Dann gilt für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$

$$\begin{aligned} & \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{\sqrt{\pi n}}{2^{n+1}(n+1)!} \cdot \frac{n^\alpha}{\Gamma(\alpha+1)2^{2\alpha}} \left( 1 + \mathcal{O}(n^{-1}) \right), \end{aligned} \quad (7.48)$$

mit  $n+1 \leq n(\alpha, N)$ .

*Beweis.* Sei  $N \in \mathbb{N}$  und sei  $f \in \mathcal{C}^{n+1}[-1, 1]$  mit  $n+1 \leq n(\alpha, N)$ . Zunächst gilt

$$\frac{N!}{N^{n+1}(N-n-1)!} = \prod_{i=0}^n \left( 1 - \frac{i}{N} \right) \leq 1. \quad (7.49)$$

Mit Lemma 7.8 erhalten wir

$$\begin{aligned} & \frac{2^{n+1}\Gamma(n+2\alpha+2)\Gamma(n+\alpha+2)}{(n+1)!\Gamma(2n+2\alpha+3)\Gamma(\alpha+1)} \frac{N!}{N^{n+1}(N-n-1)!} \\ & \leq \frac{2^{n+1}\Gamma(n+2\alpha+2)\Gamma(n+\alpha+2)}{(n+1)!\Gamma(2n+2\alpha+3)\Gamma(\alpha+1)} \\ & = \frac{2^{n+1}}{(n+1)!\Gamma(\alpha+1)} \frac{(n+1)!(n+1)!}{(2n+2)!} \frac{n^\alpha}{2^{2\alpha}} \left( 1 + \mathcal{O}(n^{-1}) \right) \\ & = \frac{2^{n+1}(n+1)!}{(2n+2)!} \frac{n^\alpha}{\Gamma(\alpha+1)2^{2\alpha}} \left( 1 + \mathcal{O}(n^{-1}) \right). \end{aligned} \quad (7.50)$$

Wir wenden Anmerkung 7.10 an und erhalten

$$\begin{aligned} & \frac{2^{n+1}\Gamma(n+2\alpha+2)\Gamma(n+\alpha+2)}{(n+1)!\Gamma(2n+2\alpha+3)\Gamma(\alpha+1)} \frac{N!}{N^{n+1}(N-n-1)!} \\ & \leq \frac{\sqrt{\pi(n+1)}}{2^{n+1}(n+1)!} \frac{n^\alpha}{\Gamma(\alpha+1)2^{2\alpha}} \left( 1 + \mathcal{O}(n^{-1}) \right). \end{aligned} \quad (7.51)$$

Somit folgt mit Theorem 7.7 die Behauptung.  $\square$

Für den wichtigen Fall  $\alpha = 0$  geben wir ergänzend die folgende Abschätzung an.

**Korollar 7.12.** Sei für  $N \in \mathbb{N}$

$$n(N) := \frac{1}{2} + \frac{1}{2}\sqrt{2N+1}. \quad (7.52)$$

Weiter sei

$$D_n := \frac{\sqrt{\pi(n+1)}}{2^{n+1}(n+1)!} e^{\frac{1}{6(n+1)} - \frac{1}{24(n+1)+1}}. \quad (7.53)$$

Dann gilt für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$  und für alle  $N \in \mathbb{N}$  mit  $n + 1 \leq n(N)$

$$\sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \leq D_n \sup_{x \in [-1, 1]} |f^{(n+1)}(x)|. \quad (7.54)$$

In Ungleichung (7.54) ist unter obigen Voraussetzungen die Konstante  $D_n$  höchstens um den Faktor

$$d_n := e^{\frac{2}{12(n+1)+1} + \frac{1}{24(n+1)+1} - \frac{1}{6(n+1)} - \frac{1}{24(n+1)}} \approx 1 \quad (7.55)$$

verbesserbar.

*Beweis.* Sei  $N \in \mathbb{N}$  und sei  $f \in \mathcal{C}^{n+1}[-1, 1]$  mit  $n + 1 \leq n(N)$ . Für  $\alpha = 0$  reduziert sich das Theorem 7.7 zu

$$\begin{aligned} & \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{2^{n+1} \Gamma(n+2) \Gamma(n+2)}{(n+1)! \Gamma(2n+3) \Gamma(1)} \frac{N!}{N^{n+1} (N-n-1)!} \\ & = \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{2^{n+1} (n+1)!}{(2n+2)!} \frac{N!}{N^{n+1} (N-n-1)!}. \end{aligned} \quad (7.56)$$

Mit der Abschätzung

$$\frac{N!}{N^{n+1} (N-n-1)!} = \prod_{i=0}^n \left( 1 - \frac{i}{N} \right) \leq 1, \quad (7.57)$$

erhalten wir

$$\begin{aligned} & \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{2^{n+1} (n+1)!}{(2n+2)!}. \end{aligned} \quad (7.58)$$

Diese Abschätzung ist wegen

$$\frac{N!}{N^{n+1} (N-n-1)!} = \prod_{i=0}^n \left( 1 - \frac{i}{N} \right) \rightarrow 1, \quad \text{für } N \rightarrow \infty, \quad (7.59)$$

für beliebiges  $N$  unverbesserbar. Mit Lemma 7.9 folgt

$$\begin{aligned} & \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| \\ & \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{\sqrt{\pi(n+1)}}{2^{n+1} (n+1)!} e^{\frac{1}{6(n+1)} - \frac{1}{24(n+1)+1}} \end{aligned} \quad (7.60)$$

sowie die Unverbesserbarkeit dieser Ungleichung bis auf den Faktor

$$d_n = e^{\frac{2}{12(n+1)+1} + \frac{1}{24(n+1)+1} - \frac{1}{6(n+1)} - \frac{1}{24(n+1)}}. \quad (7.61)$$

□

Aus diesem Korollar 7.12 lässt sich nun eine interessante und spezielle Antwort auf unsere ursprüngliche Fragestellung finden:

Für welche Klassen von Funktionen  $K \subset \mathcal{C}[-1, 1]$ , und für welches Verhältnis  $N/n$  erhält man die gleichmäßige Konvergenz der Methode der kleinsten Quadrate  $LS_n^N$ ?

**Korollar 7.13.** Sei  $\alpha > -\frac{1}{2}$ , sei

$$f \in K := \left\{ g \in \mathcal{C}^\infty[-1, 1] : \lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} |g^{(n)}(x)| \frac{n^{\alpha + \frac{1}{2}}}{2^n n!} = 0 \right\} \quad (7.62)$$

und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge mit

$$N_n \geq \frac{2n^2 + (4\alpha + 2)n}{2\alpha + 1}. \quad (7.63)$$

Dann konvergiert die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  gleichmäßig auf dem Intervall  $[-1, 1]$ .

*Beweis.* Zunächst gilt

$$N_n \geq \frac{2n^2 + (4\alpha + 2)n}{2\alpha + 1} = \frac{2\left(n + \frac{1}{2} + \alpha\right)^2}{2\alpha + 1} - \frac{2\alpha + 1}{2}. \quad (7.64)$$

Einfache Umformungen ergeben

$$(2\alpha + 1)(2\alpha + 2N_n + 1) \geq 4\left(n + \frac{1}{2} + \alpha\right)^2. \quad (7.65)$$

Nach weiterem Umformen erhält man

$$n(\alpha, N_n) := \frac{1}{2} - \alpha + \frac{1}{2}\sqrt{(2\alpha + 1)(2\alpha + 2N_n + 1)} \geq n + 1. \quad (7.66)$$

Somit können wir Korollar 7.11 anwenden:

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N_n}{2}(1+x) \right) \right| \\ & \leq \lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{\sqrt{\pi n}}{2^{n+1}(n+1)!} \cdot \frac{n^\alpha}{\Gamma(\alpha+1)2^{2\alpha}} \left(1 + \mathcal{O}(n^{-1})\right) \\ & \leq \frac{\sqrt{\pi}}{\Gamma(\alpha+1)2^{2\alpha}} \lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} |f^{(n)}(x)| \frac{n^{\alpha + \frac{1}{2}}}{2^n n!}. \end{aligned} \quad (7.67)$$

Laut Voraussetzung ist  $f \in K$ , also folgt

$$\lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N_n}{2}(1+x) \right) \right| = 0. \quad (7.68)$$

□

Bezüglich der zuvor angesprochenen Fragestellung lässt sich nun leicht eine von  $\alpha$  unabhängige Folge  $(N_n)_{n \in \mathbb{N}}$  angeben:

**Korollar 7.14.** *Sei  $\alpha \geq 0$ , sei*

$$f \in K := \left\{ g \in \mathcal{C}^\infty[-1, 1] : \lim_{n \rightarrow \infty} \sup_{x \in [-1, 1]} |g^{(n)}(x)| \frac{n^{\alpha + \frac{1}{2}}}{2^n n!} = 0 \right\} \quad (7.69)$$

und sei  $(N_n)_{n \in \mathbb{N}}$  eine Folge mit  $N_n \geq 2n(n + 1)$ . Dann konvergiert die Methode der kleinsten Quadrate  $LS_n^{N_n}[f]$  gleichmäßig auf dem Intervall  $[-1, 1]$ .

*Beweis.* Die Folge  $(N_n)_{n \in \mathbb{N}}$  erfüllt unabhängig von  $\alpha$  die Voraussetzung aus Korollar 7.13. Denn es gilt

$$N_n \geq 2n(n + 1) \geq 2n^2 + 2n \geq \frac{2n^2}{2\alpha + 1} + \frac{(2\alpha + 1)2n}{2\alpha + 1} \geq \frac{2n^2 + (4\alpha + 2)n}{2\alpha + 1}. \quad (7.70)$$

□

### 7.3 Vergleich zum kontinuierlichen Fall

In diesem Abschnitt vergleichen wir unsere Approximationsergebnisse der Methode der kleinsten Quadrate im diskreten Fall aus den Abschnitten 7.1 und 7.2 mit den Ergebnissen im kontinuierlichen Fall. Wie auch in Kapitel 6 verstehen wir hier unter dem kontinuierlichen Fall die Reihenentwicklung in Jacobi-Polynomen  $P_n \equiv P_n^{\alpha, \alpha}$ , welchen wir in Unterabschnitt 4.3.1 eingeführt haben. Dieser wurde von H. Brass bereits in [19] untersucht. Er zeigte folgendes Resultat:

**Satz 7.15** (vgl. [19]). *Sei  $\alpha \geq -\frac{1}{2}$ . Weiter sei*

$$C_n := \frac{\sup_{x \in [-1, 1]} |P_{n+1}(x)|}{\sup_{x \in [-1, 1]} |P_{n+1}^{(n+1)}(x)|}. \quad (7.71)$$

Dann gilt für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$

$$\sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{(f, P_k)_\rho}{(P_k, P_k)_\rho} P_k(x) \right| \leq C_n \sup_{x \in [-1, 1]} |f^{(n+1)}(x)|. \quad (7.72)$$

*Diese Abschätzung ist unverbesserbar in dem Sinn, dass unter obigen Voraussetzungen die Konstante  $C_n$  in Ungleichung (7.72) durch keine kleinere Zahl ersetzt werden kann.*

Dieses Ergebnis geht ebenfalls aus der Anwendung von Satz 7.1 hervor (vgl. [20]). Im folgenden Lemma ermitteln wir den Faktor  $C_n$  aus dem vorherigen Satz 7.15, der in Gleichung (7.71) definiert ist.

**Lemma 7.16.** Sei  $\alpha \geq -\frac{1}{2}$ . Dann gilt für die Konstante  $C_n$  aus Satz 7.15

$$C_n = \frac{\sup_{x \in [-1,1]} |P_{n+1}(x)|}{\sup_{x \in [-1,1]} |P_{n+1}^{(n+1)}(x)|} = \frac{2^{n+1} \Gamma(n + \alpha + 2) \Gamma(n + 2\alpha + 2)}{(n + 1)! \Gamma(2n + 2\alpha + 3) \Gamma(\alpha + 1)}. \quad (7.73)$$

*Beweis.* Zunächst sind die Jacobi-Polynome nach Definition 3.7 gegeben durch

$$P_{n+1}(x) = \frac{(\alpha + 1)_{n+1}}{(n + 1)!} \sum_{k=0}^{n+1} \frac{(-n - 1)_k (n + 2\alpha + 2)_k (1 - x)^k}{(\alpha + 1)_k 2^k k!}. \quad (7.74)$$

Daraus folgt durch  $(n + 1)$ -faches Ableiten

$$\begin{aligned} P_{n+1}^{(n+1)}(x) &= \frac{(\alpha + 1)_{n+1} (-n - 1)_{n+1} (n + 2\alpha + 2)_{n+1}}{(n + 1)! (\alpha + 1)_{n+1} 2^{n+1} (n + 1)!} \frac{d^{n+1}}{dx^{n+1}} (1 - x)^{n+1} \\ &= \frac{(-n - 1)_{n+1} (n + 2\alpha + 2)_{n+1}}{(n + 1)! 2^{n+1} (n + 1)!} (-1)^{n+1} (n + 1)!. \end{aligned} \quad (7.75)$$

Mit der Umformung

$$(-n - 1)_{n+1} = (-n - 1)(-n) \cdots (-1) = (-1)^{n+1} (n + 1)! \quad (7.76)$$

erhalten wir

$$P_{n+1}^{(n+1)}(x) = \frac{(-1)^{n+1} (n + 1)! (n + 2\alpha + 2)_{n+1}}{2^{n+1} (n + 1)!} (-1)^{n+1} = \frac{(n + 2\alpha + 2)_{n+1}}{2^{n+1}}. \quad (7.77)$$

Damit folgt

$$\sup_{x \in [-1,1]} |P_{n+1}^{(n+1)}(x)| = \frac{(n + 2\alpha + 2)_{n+1}}{2^{n+1}} = \frac{\Gamma(2n + 2\alpha + 3)}{2^{n+1} \Gamma(n + 2\alpha + 2)}. \quad (7.78)$$

Weiterhin entnehmen wir Satz 3.8

$$\sup_{x \in [-1,1]} |P_{n+1}(x)| = \binom{n + 1 + \alpha}{n + 1} = \frac{\Gamma(n + \alpha + 2)}{\Gamma(n + 2) \Gamma(\alpha + 1)}. \quad (7.79)$$

Insgesamt erhält man somit Gleichung (7.73).  $\square$

Wir vergleichen nun die beiden jeweils unverbesserbaren „Fehlerkonstanten“  $D_{n,N}$  aus Theorem 7.7 (diskreter Fall) und  $C_n$  aus Satz 7.15 (kontinuierlicher Fall). Für den Quotienten  $D_{n,N}/C_n$  gilt für  $\alpha > -\frac{1}{2}$

$$\frac{D_{n,N}}{C_n} = \frac{N!}{N^{n+1} (N - n - 1)!} = \prod_{i=0}^n \left(1 - \frac{i}{N}\right) \leq 1, \quad (7.80)$$

wobei im diskreten Fall  $n + 1 \leq \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}$  vorausgesetzt wurde. Bezeichnet man mit  $\mathcal{K}_n$  die folgende Funktionenklasse

$$\mathcal{K}_n := \left\{ f \in \mathcal{C}^n[-1, 1] : \sup_{x \in [-1,1]} |f^{(n)}(x)| \leq 1 \right\}, \quad (7.81)$$

so ergibt sich folgendes Korollar:

**Korollar 7.17.** Sei  $\alpha = \beta > -\frac{1}{2}$ . Weiter sei  $LS_n$  der Operator der Methode der kleinsten Quadrate im kontinuierlichen Fall gemäß Gleichung (4.52) und sei  $LS_n^N$  der Operator der Methode der kleinsten Quadrate im diskreten Fall gemäß Gleichung (4.65) mit  $n + 1 \leq \frac{1}{2} - \alpha + \frac{1}{2}\sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}$ . Dann gilt

$$\sup_{f \in \mathcal{K}_{n+1}} \sup_{x \in [-1,1]} |f(x) - LS_n^N[f](x)| = \prod_{i=0}^n \left(1 - \frac{i}{N}\right) \sup_{f \in \mathcal{K}_{n+1}} \sup_{x \in [-1,1]} |f(x) - LS_n[f](x)|. \quad (7.82)$$

**Anmerkung 7.18.** Für die praktische Anwendung erhalten wir für  $n \in \mathbb{N}_0$  mit Korollar 7.17 die Garantie, dass der „worst case“ bezüglich der Klasse  $\mathcal{K}_{n+1}$  im kontinuierlichen Fall schlechter ist als im entsprechenden diskreten Fall, falls für den Polynomgrad  $n$  und die Stützstellenanzahl  $N + 1$  die Ungleichung  $n + 1 \leq \frac{1}{2} - \alpha + \frac{1}{2}\sqrt{(2\alpha + 1)(2\alpha + 2N + 1)}$  erfüllt ist.

**Anmerkung 7.19.** Betrachtet man wie in der Anmerkung zuvor den „worst case“

$$\sup_{f \in \mathcal{K}_{n+1}} \sup_{x \in [-1,1]} |f(x) - LS_n^N[f](x)|, \quad (7.83)$$

so erhält man folgende Aussage:

Ein Verhältnis  $n^k/N \rightarrow 0$  mit einem  $k > 2$  führt zu keiner besseren Approximation im Sinn von (7.83) als das Verhältnis  $n^2/N \rightarrow 0$ .

## 7.4 Vergleich zur Polynominterpolation

In diesem Abschnitt vergleichen wir unsere Approximationsergebnisse der Methode der kleinsten Quadrate im diskreten Fall aus dem Abschnitt 7.2 mit der Polynominterpolation. Wir haben bereits in Abschnitt 4.1 die Polynominterpolation behandelt und einige bekannte Approximationsresultate dargestellt.

Insbesondere haben wir in Unterabschnitt 4.1.1 aufgezeigt, dass für die Praxis die Verwendung äquidistanter Stützstellen im Fall einer großen Stützstellenanzahl häufig ungeeignet ist.

Für die in diesem Kapitel untersuchte Funktionenklasse  $\mathcal{K}_{n+1}$ , vgl. (7.81), ziehen wir für einen Vergleich die Interpolation in Tschebyscheff-Stützstellen, siehe Anmerkung 4.8, heran.

Im Folgenden werde mit  $intpol(x_1, \dots, x_n)$  derjenige Operator bezeichnet, der jedem  $f \in \mathcal{C}[-1, 1]$  das Interpolationspolynom bezüglich der Stützstellen  $x_1, \dots, x_n$  zuordnet. Des Weiteren bezeichnen wir mit  $Z_n$  die wie folgt gegebene Menge reeller Zahlentupel:

$$Z_n := \{(x_1, \dots, x_n) : -1 \leq x_1 < \dots < x_n \leq 1\}. \quad (7.84)$$

Die Interpolation in Tschebyscheff-Stützstellen ist in dem Sinn in der Klasse  $\mathcal{K}_{n+1}$  optimal, dass der Wert

$$I_n := \inf_{(x_1, \dots, x_{n+1}) \in Z_{n+1}} \sup_{f \in \mathcal{K}_{n+1}} \sup_{x \in [-1, 1]} |f(x) - \text{intpol}(x_1, \dots, x_{n+1})[f](x)| = \frac{1}{2^n (n+1)!} \quad (7.85)$$

nur durch den Interpolationsoperator angenommen wird, der als Stützstellen die Nullstellen des Tschebyscheff-Polynoms  $T_{n+1}$  verwendet, siehe Satz 4.10 und Korollar 4.11.

Wir vergleichen nun den in Gleichung (7.85) definierten Wert  $I_n$  für die Polynominterpolation mit dem in Theorem 7.7 definierten Wert  $D_{n,N}$  für die Methode der kleinsten Quadrate im diskreten Fall. Eine grobe Abschätzung mit Korollar 7.11 ergibt sofort für  $\alpha > -\frac{1}{2}$

$$\frac{D_{n,N}}{I_n} \leq \frac{\sqrt{\pi}}{\Gamma(\alpha+1)} \cdot \frac{1}{2^{2\alpha+1}} n^{\alpha+\frac{1}{2}} (1 + \mathcal{O}(n^{-1})) = \mathcal{O}(n^{\alpha+\frac{1}{2}}), \quad (7.86)$$

unter der Voraussetzung  $n+1 \leq \frac{1}{2} - \alpha + \frac{1}{2} \sqrt{(2\alpha+1)(2\alpha+2N+1)}$ .

Der Unterschied ist also nur von polynomieller Größenordnung und ist daher in Bezug auf die Größenordnung von  $I_n$  nur von geringem praktischen Interesse.

## 7.5 Vergleich zur Approximation mit Bernsteinpolynomen

In diesem Abschnitt vergleichen wir unsere Approximationsergebnisse der Methode der kleinsten Quadrate im diskreten Fall aus dem Abschnitt 7.2 mit der Approximation mit Bernsteinpolynomen aus Abschnitt 4.2.

Ein Vergleich bezüglich des Maximalfehlers in der Klasse  $\mathcal{K}_{n+1}$ , wie in den vorhergehenden Abschnitten 7.3 und 7.4 untersucht, führt hier nicht weiter. Denn es gilt für jedes  $n > 1$

$$\sup_{f \in \mathcal{K}_{n+1}} \sup_{x \in [0, 1]} |f(x) - B_n[f](x)| = \infty. \quad (7.87)$$

Dies sieht man wie folgt leicht ein. Für die Funktion  $p_2$ , definiert durch  $p_2(x) := x^2$  erhält man nach längerer, elementarer Rechnung unter Verwendung der Definition 4.12 des Bernsteinpolynoms:

$$\sup_{x \in [0, 1]} |p_2(x) - B_n[p_2](x)| = \sup_{x \in [0, 1]} \frac{1}{n} |x^2 - x| = \frac{1}{4n}. \quad (7.88)$$

Wir wählen nun für festes  $n > 1$  und beliebiges  $A > 0$  die Funktion  $f := A \cdot p_2$ . Dann gilt  $f^{(n+1)} \equiv 0$ , also ist  $f \in \mathcal{K}_{n+1}$ . Der Maximalfehler ist  $\frac{A}{4n}$ . Es ist aber  $n$  fest gewählt und  $A > 0$  beliebig wählbar!

Die Entscheidung, ob die Approximation mit Bernsteinpolynomen der Methode der kleinsten Quadrate im diskreten Fall vorzuziehen ist, hängt erheblich von der praktischen Situation ab.

- Ist beispielsweise  $f$  eine zweimal stetig differenzierbare Funktion mit beschränkter zweiter Ableitung und  $f$  kein Polynom vom Höchstgrad 1, so ist der Maximalfehler

$$\sup_{x \in [0,1]} |f(x) - B_n[f](x)| \quad (7.89)$$

exakt von der Ordnung  $n^{-1}$ .

- Ist beispielsweise  $f \in C^\infty[-1, 1]$  und sind alle Ableitungen der Funktion betragsmäßig durch die gleiche Konstante beschränkt, so gilt für den Maximalfehler gemäß Korollar 7.11

$$\sup_{x \in [-1,1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, Q_k \rangle_\omega}{\langle Q_k, Q_k \rangle_\omega} Q_k \left( \frac{N}{2}(1+x) \right) \right| = \mathcal{O} \left( \frac{n^{\alpha+\frac{1}{2}}}{2^n(n+1)!} \right). \quad (7.90)$$

Andererseits erfüllen die Bernsteinpolynome verschiedene formerhaltende Eigenschaften, weshalb sie für einige Anwendungen besonders interessant sind. Diese Eigenschaften haben wir in Anmerkung 4.15 zusammengefasst.

Zudem konvergiert die Folge der Bernsteinpolynome gleichmäßig für jede stetige Funktion, dies kann für die Methode der kleinsten Quadrate gemäß Kapitel 5 nicht gewährleistet werden.

## 7.6 Vergleich zu anderer Stützstellenwahl

Wir haben in dieser Arbeit bisher die Methode der kleinsten Quadrate im diskreten Fall mit äquidistanten Stützstellen untersucht. In diesem Abschnitt gehen wir auf Ergebnisse aus der Literatur ein, bei denen andere Stützstellen verwendet werden.

Ein ähnliches Resultat zu Korollar 7.12 haben R. Barnard, G. Dahlquist, K. Pearce, L. Reichel und K. Richards in [9] für die Gram-Polynome  $\phi_k$  gezeigt. Die Gram-Polynome  $\phi_k$  sind diskrete orthogonale Polynome auf dem Intervall  $[-1, 1]$  vom Grad  $k$ . Sie sind orthogonal auf  $[-1, 1]$  bezüglich des Skalarprodukts

$$(f, g)_d := \frac{1}{N} \sum_{i=1}^N f \left( -1 + \frac{2i-1}{N} \right) g \left( -1 + \frac{2i-1}{N} \right). \quad (7.91)$$

Sie sind durch  $(\phi_k, \phi_k)_d = 1$  normiert (vgl., z.B., [9, S. 129]). Die Gram-Polynome sind also mit den Hahn-Polynomen mit den Parametern  $\alpha = \beta = 0$  vergleichbar. Der Unterschied liegt in der Stützstellenverteilung des Skalarprodukts am Rand.

Die äußersten Stützstellen liegen hier einen halben Stützstellenabstand nach innen versetzt. Für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$  gilt

$$\begin{aligned} & \sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n (f, \phi_k)_d \phi_k(x) \right| \\ & \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{n^{1/2}}{2^n(n+1)!} \cdot \frac{\pi^{1/2}}{2} \left(1 + \mathcal{O}(n^{-1})\right) + \hat{c}_n \mathcal{O}(N^{-2}), \end{aligned} \quad (7.92)$$

wobei  $N > n$ . Dabei ist die Konstante  $\hat{c}_n$  zwar unabhängig von  $f$  und  $N$ , aber abhängig von  $n$  (vgl. [9, S. 131]). Jedoch ist das Problem dieser Ungleichung eben genau diese Konstante  $\hat{c}_n$ , denn man hat keine Information über das Wachstum von  $\hat{c}_n$  wenn  $n$  ansteigt. Die uns interessierenden Aussagen über die Konvergenz bei wachsendem  $n$ , und damit natürlich bei wachsendem  $N$ , sind also mit diesem Ergebnis nicht zu erzielen.

Man kann die Frage stellen, ob eine geschickte Stützstellenwahl schon dann eine gute Approximation ermöglicht, wenn nur  $N > n$  vorausgesetzt wird. Diese Überlegung führt zu diskreten orthogonalen Polynomen auf nicht äquidistanten Stützstellen. Hierzu wollen wir im Folgenden auf die Arbeit [28] von A. Eisinberg und G. Fedele aus dem Jahr 2007 eingehen. In dieser haben die Autoren den Ansatz verfolgt, diskrete orthogonale Polynome mit Hilfe der Quadraturtheorie zu bestimmen. Diese Methode basiert auf der Arbeit [49] von T. Koornwinder. In dieser hat Koornwinder orthogonale Polynome untersucht, die orthogonal bezüglich der Gewichtsfunktion

$$(1-x)^\alpha(1+x)^\beta + M_1\delta(x+1) + M_2\delta(x-1) \quad (7.93)$$

sind. Diese setzt sich aus der Gewichtsfunktion  $\varrho(x) = (1-x)^\alpha(1+x)^\beta$  der Jacobi-Polynome sowie der Delta-Distribution bei  $x = -1$  und  $x = 1$  zusammen. Er hat unter anderem gezeigt, dass sich diese Polynome in Termen von Jacobi-Polynomen  $P_n^{\alpha, \beta}$  in der Gestalt

$$\left( (a_n x + b_n) \frac{d}{dx} + c_n \right) P_n^{\alpha, \beta}(x) \quad (7.94)$$

mit gewissen Konstanten  $a_n, b_n$  und  $c_n$  ausdrücken lassen.

In der Arbeit [28] haben Eisinberg und Fedele orthogonale Polynome bezüglich eines diskreten Skalarprodukts auf den sogenannten Gauss-Lobatto Chebyshev Punkten

$$X_n = \left\{ x_k = -\cos\left(\frac{k-1}{n-1}\pi\right) : k = 1, 2, \dots, n \right\} \quad (7.95)$$

betrachtet. Das Skalarprodukt ist auf der Knotenmenge  $X_n$  durch

$$\langle f, g \rangle = \sum_{k=1}^n f(x_k) g(x_k) \quad (7.96)$$

definiert. Weiterhin definierten die Autoren entsprechend der Gewichtsfunktion (7.93) ein Maß  $\mu$  auf dem Intervall  $[-1, 1]$  durch

$$\int_{-1}^1 f(x) d\mu = \frac{\Gamma(\alpha + \beta + 2)}{2^{\alpha+\beta+1} \Gamma(\alpha + 1) \Gamma(\beta + 1)} \int_{-1}^1 f(x) (1-x)^\alpha (1+x)^\beta dx + M_1 f(-1) + M_2 f(1). \quad (7.97)$$

Für den Fall  $\alpha = \beta = -\frac{1}{2}$  und  $M_1 = M_2 = \frac{1}{2(n-1)}$  lässt sich nun mit Hilfe der zugehörigen Gauss-Lobatto Quadraturformel folgende Beziehung zwischen (7.96) und (7.97) herstellen:

$$\begin{aligned} & \frac{1}{\pi} \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx + \frac{1}{2(n-1)} [f(x_1) + f(x_n)] \\ &= \frac{1}{n-1} \sum_{k=1}^n f(x_k) - \frac{1}{2^{2n-3} (2n-2)!} f^{(2n-2)}(\xi), \end{aligned} \quad (7.98)$$

mit  $x_k = -\cos\left(\frac{k-1}{n-1}\pi\right)$ ,  $k = 1, 2, \dots, n$  und  $\xi \in [-1, 1]$ . Diese Formel integriert also Polynome bis zum Grad  $2n-3$  exakt. Definiert man nun folgendes Skalarprodukt

$$(f, g) = \frac{1}{\pi} \int_{-1}^1 \frac{f(x)g(x)}{\sqrt{1-x^2}} dx + \frac{1}{2(n-1)} [f(x_1)g(x_1) + f(x_n)g(x_n)], \quad (7.99)$$

so gilt

$$\langle f, g \rangle = (n-1) (f, g) \quad (7.100)$$

mit

$$\text{grad } f + \text{grad } g \leq 2n - 3. \quad (7.101)$$

Die explizite Darstellung (7.94) der orthogonalen Polynome bezüglich des Skalarprodukts (7.99) ist gegeben durch

$$p_k(x) = \frac{(n+k-2)^2}{(n-1)^2} P_{k-1}^{-\frac{1}{2}, -\frac{1}{2}}(x) - \frac{n+k-2}{(k-1)(n-1)^2} x \frac{d}{dx} \left[ P_{k-1}^{-\frac{1}{2}, -\frac{1}{2}}(x) \right] \quad (7.102)$$

(vgl. [49]). In numerischen Tests bezogen auf die zugehörige Methode der kleinsten Quadrate mit dem Skalarprodukt (7.96) konnten Eisenberg und Fedele interessante Approximationsergebnisse für die ausgewählten Beispielfunktionen erzielen. Eine Fehlerabschätzung, die wir mit unseren Abschätzungen aus Abschnitt 7.2 vergleichen können, wurde nicht bewiesen.

Wir betrachten nun den Fall, dass wir das Skalarprodukt (7.96) an den Rändern modifizieren. Wir definieren ein Skalarprodukt auf der oben definierten Knotenmenge

$$X_N = \left\{ x_k = -\cos\left(\frac{k-1}{N-1}\pi\right) : k = 1, 2, \dots, N \right\} \quad (7.103)$$

durch

$$\langle f, g \rangle_1 = \frac{1}{N-1} f(x_1) g(x_1) + \frac{2}{N-1} \sum_{k=2}^{N-1} f(x_k) g(x_k) + \frac{1}{N-1} f(x_N) g(x_N). \quad (7.104)$$

Dann bilden die in Definition 3.9 eingeführten Tschebyscheff-Polynome

$$T_n(x) = \frac{P_n^{-\frac{1}{2}, -\frac{1}{2}}(x)}{P_n^{-\frac{1}{2}, -\frac{1}{2}}(1)} \quad (7.105)$$

für  $n = 0, \dots, N-1$  ein System von diskreten orthogonalen Polynomen bezüglich des Skalarprodukts (7.104) (vgl., z.B., [66, S. 50]). In [20] zeigte Brass durch Anwenden von Satz 7.1 die folgende Abschätzung:

Für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$  gilt

$$\sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, T_k \rangle_1}{\langle T_k, T_k \rangle_1} T_k(x) \right| \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{1}{(n+1)! 2^n}, \quad (7.106)$$

mit  $n+2 \leq N$ .

Die gleiche Abschätzung erhält man auch, wenn man das Skalarprodukt auf den Tschebyscheff-Stützstellen definiert. Die Tschebyscheff-Stützstellen sind nach Anmerkung 4.8 gegeben durch

$$X_N^T = \left\{ x_k^T = \cos\left(\frac{2k-1}{2N}\pi\right) : k = 1, 2, \dots, N \right\}. \quad (7.107)$$

Wir definieren also ein weiteres Skalarprodukt durch

$$\langle f, g \rangle_2 = \frac{2}{N} \sum_{k=1}^N f(x_k^T) g(x_k^T). \quad (7.108)$$

Dann bildet das System  $\{T_k\}_{k=0}^{N-1}$  von Tschebyscheff-Polynomen bezüglich dieses Skalarprodukts ein System von diskreten orthogonalen Polynomen (vgl., z.B., [66, S. 49]). Auch in diesem Fall zeigte Brass in [20] durch Anwenden von Satz 7.1 die folgende Abschätzung:

Für jedes  $f \in \mathcal{C}^{n+1}[-1, 1]$  gilt

$$\sup_{x \in [-1, 1]} \left| f(x) - \sum_{k=0}^n \frac{\langle f, T_k \rangle_2}{\langle T_k, T_k \rangle_2} T_k(x) \right| \leq \sup_{x \in [-1, 1]} |f^{(n+1)}(x)| \frac{1}{(n+1)! 2^n}, \quad (7.109)$$

mit  $n + 2 \leq N$ .

Die rechten Seiten der beiden obigen Abschätzungen (7.106) und (7.109) entsprechen genau der rechten Seite der Abschätzung (4.31) im Fall der Polynominterpolation in Tschebyscheff-Stützstellen. Wir verweisen also auf den dortigen Vergleich mit der Methode der kleinsten Quadrate bezüglich der Funktionenklasse  $\mathcal{K}_{n+1}$  in Abschnitt 7.4.

## 7.7 Vergleich zum Polynom bester Approximation

In diesem Abschnitt vergleichen wir unsere Approximationsergebnisse der Methode der kleinsten Quadrate im diskreten Fall aus dem Abschnitt 7.2 mit der Bestapproximation durch Polynome. Wir haben bereits in Abschnitt 4.4 diese Approximationsmethode behandelt und einige bekannte Eigenschaften dieser Methode dargestellt.

Für Polynome bester Approximation ist der Maximalfehler in der Klasse  $\mathcal{K}_{n+1}$  nach Satz 4.25 der Wert  $\frac{1}{2^n(n+1)!}$ . Wir können also festhalten, dass die in der Klasse  $\mathcal{K}_{n+1}$  unverbesserbare Fehlerabschätzung für Polynome bester Approximation aus Satz 4.25 mit den Abschätzungen

- (7.106) bei der Methode der kleinsten Quadrate in mindestens  $n + 2$  Gauss-Lobatto Chebyshev Stützstellen,
- (7.109) bei der Methode der kleinsten Quadrate in mindestens  $n + 2$  Tschebyscheff-Stützstellen,
- (4.31) bei der Polynominterpolation in  $n + 1$  Tschebyscheff-Stützstellen

übereinstimmt. Unseren Ergebnissen in den Abschnitten 7.1 und 7.2 entnimmt man, dass die Methode der kleinsten Quadrate im diskreten Fall  $LS_n^N$  in der Klasse  $\mathcal{K}_{n+1}$  konkurrenzfähig ist, wenn Polynomgrad  $n$  und Stützstellenanzahl  $N + 1$  geeignet gewählt werden. Man vergleiche hierzu die Ausführungen in Abschnitt 7.4.

## 8 Numerische Resultate

In diesem Kapitel präsentieren wir numerische Resultate zur Methode der kleinsten Quadrate. Dabei gehen wir insbesondere auf den in Unterabschnitt 4.3.2 erläuterten diskreten Fall ein und verwenden den Operator der Methode der kleinsten Quadrate  $LS_n^N$  wieder gemäß Gleichung (4.65).

In Abschnitt 8.1 vergleichen wir numerisch den diskreten mit dem kontinuierlichen Fall der Methode der kleinsten Quadrate. Außerdem vergleichen wir die Methode im diskreten Fall mit der Polynominterpolation in Tschebyscheff-Stützstellen sowie mit der Approximation mit Bernsteinpolynomen. Diese Methoden haben wir in Kapitel 4 bereitgestellt.

In Abschnitt 8.2 betrachten wir zum einen Fälle mit einem asymptotischen Verhältnis  $n^k/N \rightarrow 0$  mit einem  $k > 2$ . Wir haben in Abschnitt 7.3 gezeigt, dass ein Verhältnis  $n^k/N \rightarrow 0$  mit einem  $k > 2$  zu keiner Verbesserung des Maximalfehlers in der Klasse  $\mathcal{K}_{n+1}$  als das Verhältnis  $n^2/N \rightarrow 0$  führt. Wir untersuchen numerisch, ob für ausgewählte Testfunktionen entsprechende Aussagen gelten können.

Zum anderen untersuchen wir numerisch bei welchem Verhältnis  $N/n$  der Maximalfehler bei wachsenden  $n$  und  $N$  ansteigt. Für den Grenzfall  $N = n$  handelt es sich, wie in Anmerkung 4.22 dargestellt, bei dem Operator  $LS_n^N$  um den Interpolationsoperator in äquidistanten Stützstellen. Man vergleiche diesbezüglich unsere Ausführungen zum Runge-Phänomen in Unterabschnitt 4.1.1. Andererseits wissen wir aus unseren Ergebnissen in Kapitel 7, dass für Funktionen  $f \in \mathcal{C}^\infty[-1, 1]$  mit Zusatzforderungen an die Beschränkung der Ableitungen das asymptotische Verhältnis  $n^2/N \rightarrow 0$  für die gleichmäßige Konvergenz der Methode der kleinsten Quadrate ausreicht. Wir untersuchen daher numerisch das Approximationsverhalten bei einem Verhältnis innerhalb dieser beiden Grenzfälle  $N = n$  und  $n^2/N \rightarrow 0$ .

In beiden Abschnitten wenden wir die Methoden exemplarisch auf die Runge-Funktion

$$f(x) = \frac{1}{1 + 25x^2}, \quad x \in [-1, 1] \quad (8.1)$$

sowie auf die Betragsfunktion

$$g(x) = |x|, \quad x \in [-1, 1] \quad (8.2)$$

an, zur Motivation vgl. Unterabschnitt 4.1.1.

Wir berechnen jeweils die durchschnittlichen sowie maximalen Fehler numerisch auf einem äquidistanten Gitter  $\tilde{X}$  mit  $\tilde{N} = 10001$  Knoten im Intervall  $[-1, 1]$ :

$$\tilde{X} = \left\{ \tilde{x}_\mu = -1 + \mu \frac{2}{\tilde{N} - 1} : \mu = 0, 1, \dots, \tilde{N} - 1 \right\}. \quad (8.3)$$

## 8.1 Vergleich zu anderen Approximationsmethoden

Bei der Methode der kleinsten Quadrate im diskreten Fall verwenden wir die Stützstellenanzahl  $N = 2n(n + 1)$  aus Korollar 7.13, welche wir dort als hinreichende untere Grenze für die gleichmäßige Konvergenz ermittelt haben.

Wir verwenden mit  $h = f$  aus Gleichung (8.1) beziehungsweise  $h = g$  aus Gleichung (8.2) für die Fehler die folgenden Bezeichnungen:

$$\begin{aligned} \varepsilon_M^D &= \max_{\tilde{x} \in \tilde{X}} |h(\tilde{x}) - LS_n^N[h](\tilde{x})|, \\ \varepsilon_D^D &= \frac{1}{\tilde{N}} \sum_{\tilde{x} \in \tilde{X}} |h(\tilde{x}) - LS_n^N[h](\tilde{x})|, \\ \varepsilon_M^C &= \max_{\tilde{x} \in \tilde{X}} |h(\tilde{x}) - LS_n[h](\tilde{x})|, \\ \varepsilon_D^C &= \frac{1}{\tilde{N}} \sum_{\tilde{x} \in \tilde{X}} |h(\tilde{x}) - LS_n[h](\tilde{x})|, \\ \varepsilon_M^B &= \max_{\tilde{x} \in \tilde{X}} \left| h\left(\frac{\tilde{x} + 1}{2}\right) - B_n[h(2(\cdot) - 1)]\left(\frac{\tilde{x} + 1}{2}\right) \right|, \\ \varepsilon_D^B &= \frac{1}{\tilde{N}} \sum_{\tilde{x} \in \tilde{X}} \left| h\left(\frac{\tilde{x} + 1}{2}\right) - B_n[h(2(\cdot) - 1)]\left(\frac{\tilde{x} + 1}{2}\right) \right|, \\ \varepsilon_M^T &= \max_{\tilde{x} \in \tilde{X}} |h(\tilde{x}) - p_n[h](\tilde{x})|, \\ \varepsilon_D^T &= \frac{1}{\tilde{N}} \sum_{\tilde{x} \in \tilde{X}} |h(\tilde{x}) - p_n[h](\tilde{x})|. \end{aligned}$$

- Dabei ist  $LS_n^N$  der Operator der Methode der kleinsten Quadrate im diskreten Fall aus Unterabschnitt 4.3.2.
- Weiter ist  $LS_n$  der Operator der Methode der kleinsten Quadrate im kontinuierlichen Fall aus Unterabschnitt 4.3.1.
- $B_n[h(2(\cdot) - 1)]$  ist das Bernsteinpolynom aus Abschnitt 4.2 von der transformierten Funktion  $h(2(\cdot) - 1)$ . Man beachte dabei die lineare Transformation  $x \mapsto \frac{x+1}{2}$  des Intervalls  $[-1, 1]$  auf das Intervall  $[0, 1]$ , da die Bernsteinpolynome auf dem Intervall  $[0, 1]$  definiert sind, siehe dazu Abschnitt 4.2.
- $p_n[h]$  ist das Interpolationspolynom von  $h$  in den Tschebyscheff-Stützstellen aus Unterabschnitt 4.1.2.

Im Folgenden sind die jeweiligen durchschnittlichen beziehungsweise maximalen relativen Fehler stets in Bezug auf  $\varepsilon_D^D$  beziehungsweise  $\varepsilon_M^D$  angegeben.

### 8.1.1 Vergleich zum kontinuierlichen Fall

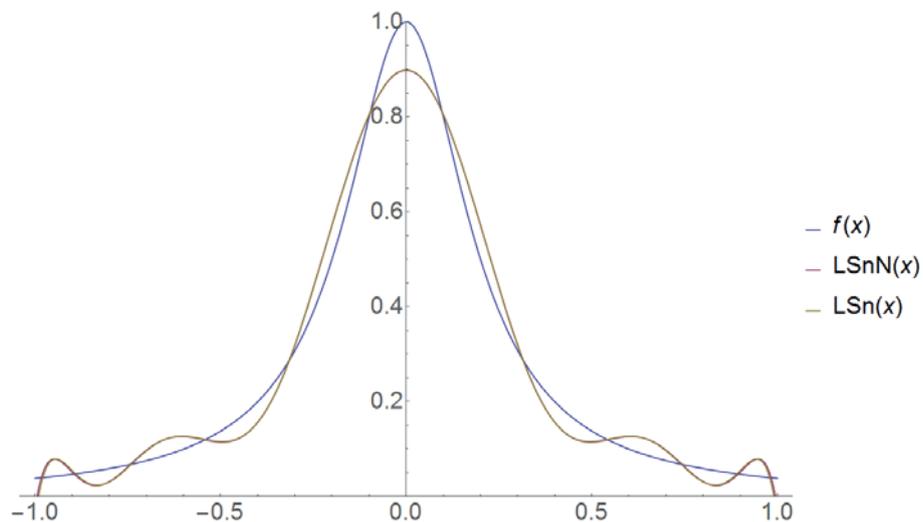


Abbildung 8.1

Funktionen:  $f, LS_n^N[f], LS_n[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

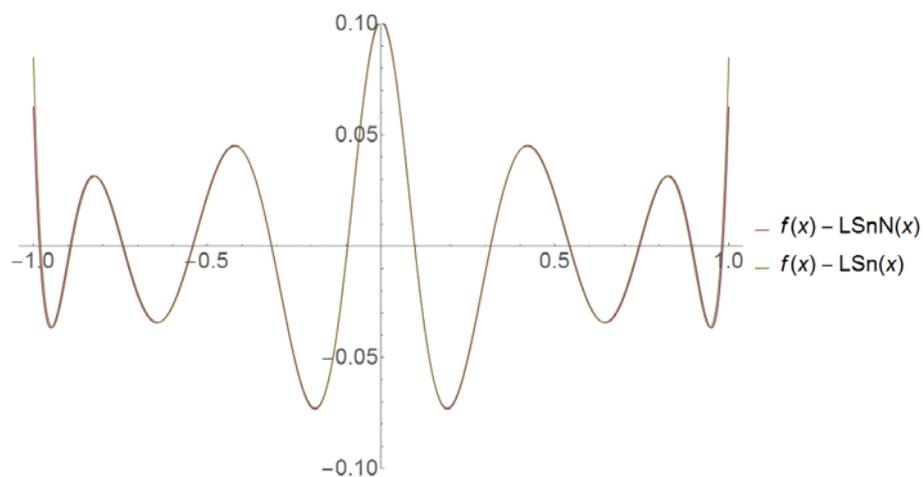


Abbildung 8.2

Funktionen:  $f - LS_n^N[f], f - LS_n[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

Fehler:  $\varepsilon_M^D = 0.101513, \varepsilon_D^D = 0.0327692,$   
 $\varepsilon_M^C = 0.100575, \varepsilon_D^C = 0.0327074.$

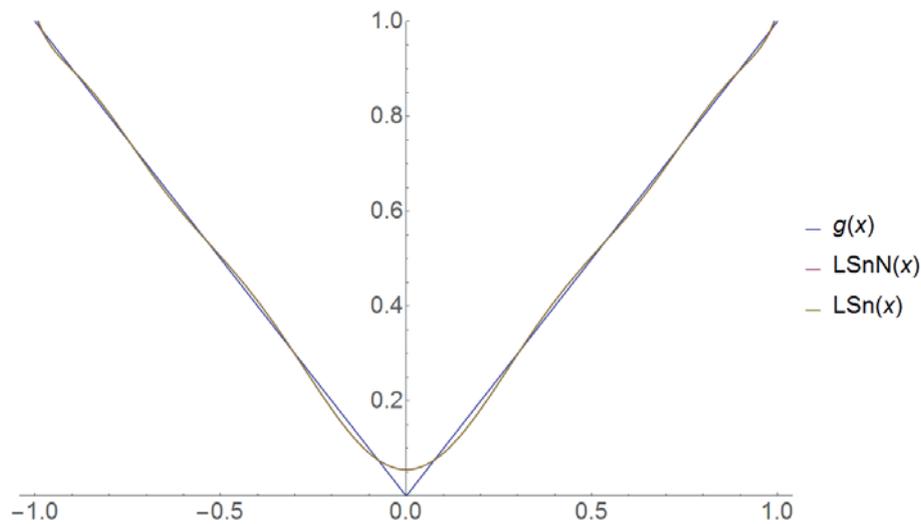


Abbildung 8.3

Funktionen:  $g, LS_n^N[g], LS_n[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

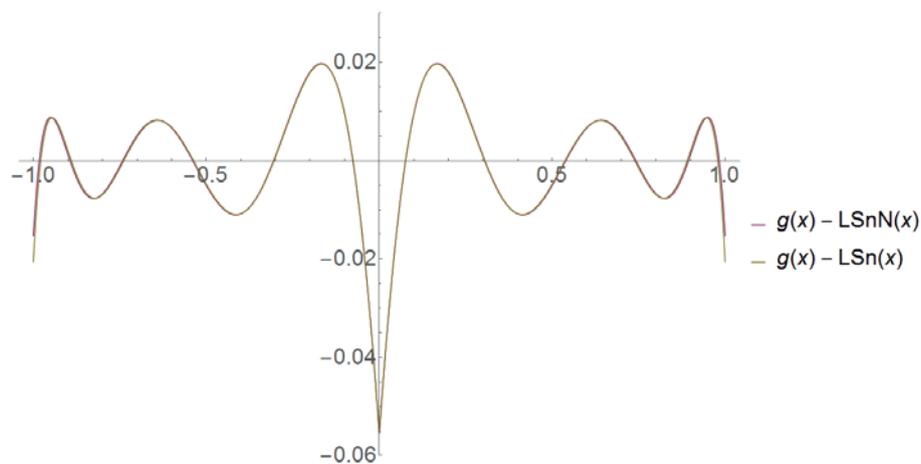


Abbildung 8.4

Funktionen:  $g - LS_n^N[g], g - LS_n[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

Fehler:  $\varepsilon_M^D = 0.0556922, \varepsilon_D^D = 0.00883205,$   
 $\varepsilon_M^C = 0.0555153, \varepsilon_D^C = 0.00882663.$

Bei beiden Testfunktionen liegt der durchschnittliche Fehler  $\varepsilon_D^D$  und der maximale Fehler  $\varepsilon_M^D$  im diskreten Fall geringfügig über den Fehlern  $\varepsilon_D^C$  und  $\varepsilon_M^C$  im kontinuierlichen Fall. Der Maximalfehler unterscheidet sich jeweils um höchstens 1.0%, der durchschnittliche Fehler jeweils sogar nur um höchstens 0.2%. Dies kann man auch in den Abbildungen optisch erkennen. Die Ergebnisse für die Runge-Funktion  $f$  verdeutlichen unsere Resultate aus Abschnitt 7.3: Wählt man  $N \geq 2n(n + 1)$ , dann ist der Qualitätsunterschied bei  $(n + 1)$ -mal stetig differenzierbaren Funktionen nur

gering. Erstaunlich gut ist das Ergebnis für die Betragsfunktion  $g$ . Die Resultate aus Kapitel 7 sind nicht anwendbar, da die Betragsfunktion nicht differenzierbar ist. Die Ergebnisse in Kapitel 6 weisen solch ein Verhalten erst bei einem asymptotischen Verhältnis  $n^4/N \rightarrow 0$  nach. Diese Situation untersuchen wir weiter in Abschnitt 8.2.

### 8.1.2 Vergleich zur Approximation mit Bernsteinpolynomen

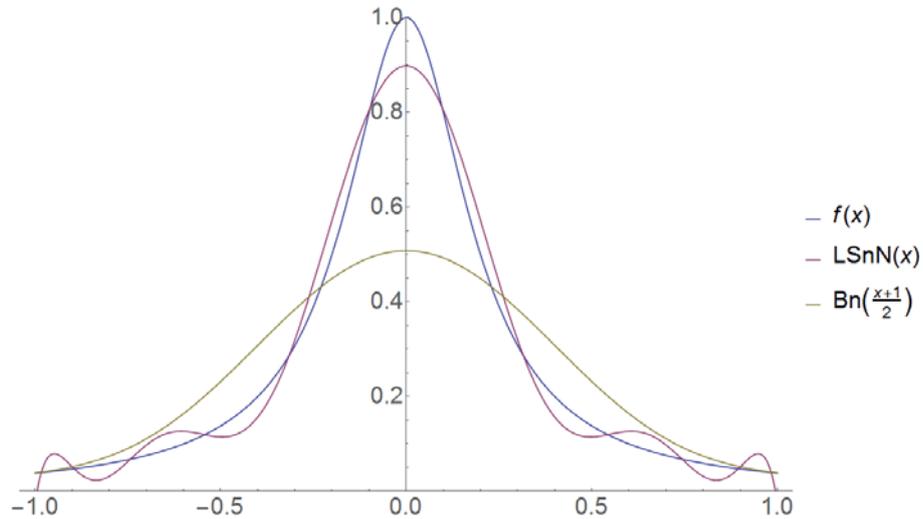


Abbildung 8.5

Funktionen:  $f, LS_n^N[f], B_n[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

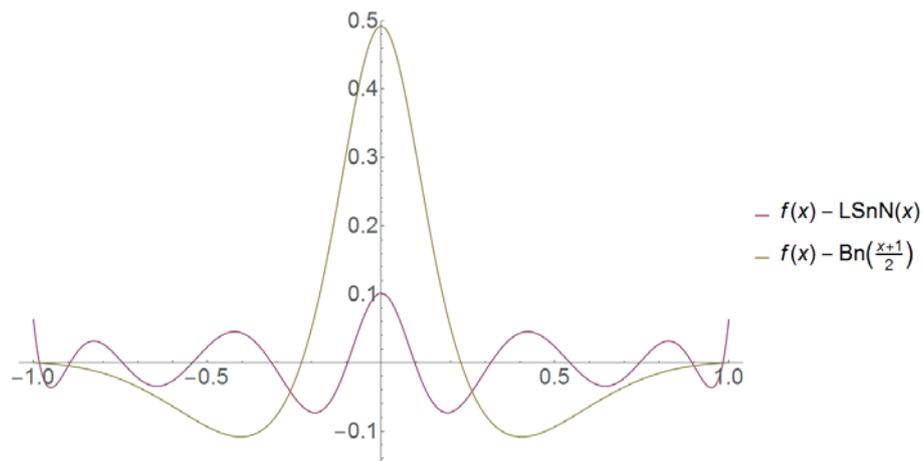


Abbildung 8.6

Funktionen:  $f - LS_n^N[f], f - B_n[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

Fehler:  $\varepsilon_M^D = 0.101513, \varepsilon_D^D = 0.0327692,$   
 $\varepsilon_M^B = 0.491940, \varepsilon_D^B = 0.0997603.$

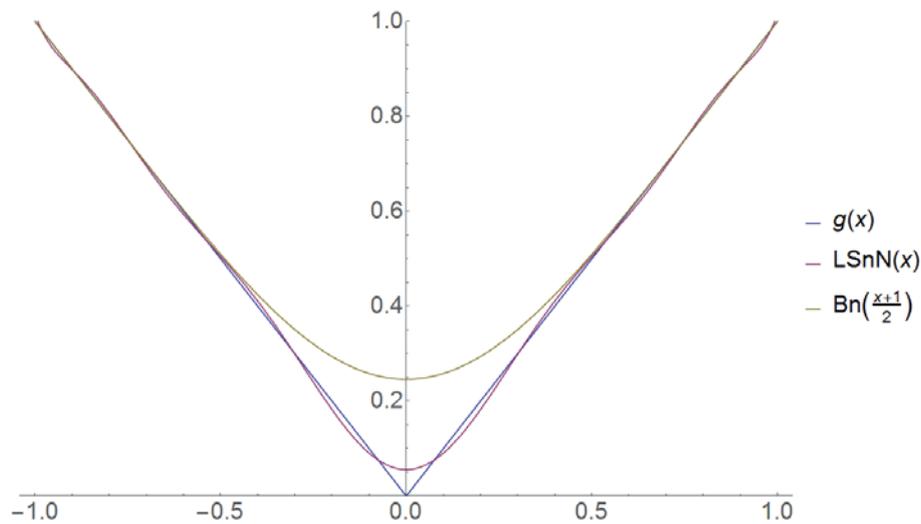


Abbildung 8.7

Funktionen:  $g, LS_n^N[g], B_n[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

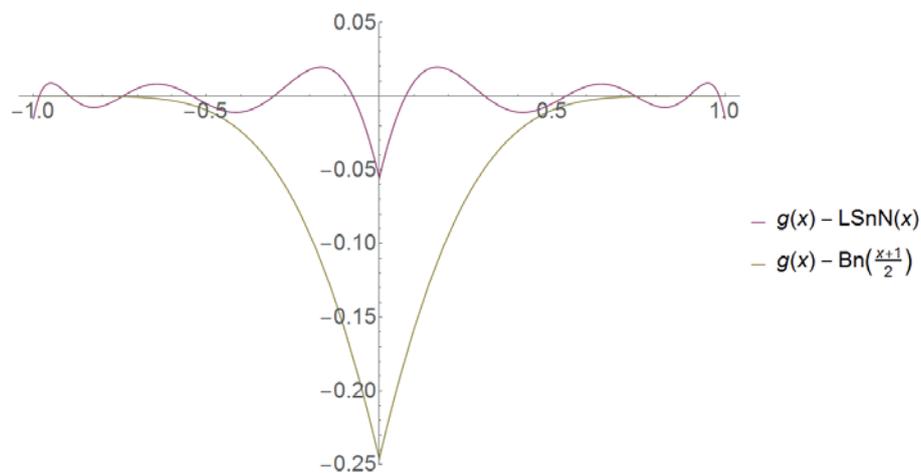


Abbildung 8.8

Funktionen:  $g - LS_n^N[g], g - B_n[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

Fehler:  $\varepsilon_M^D = 0.0556922, \varepsilon_D^D = 0.00883205,$   
 $\varepsilon_M^B = 0.246094, \varepsilon_D^B = 0.0454500$ .

Bei beiden Testfunktionen liegt sowohl der durchschnittliche Fehler  $\varepsilon_D^D$  als auch der maximale Fehler  $\varepsilon_M^D$  bei der Methode der kleinsten Quadrate im diskreten Fall deutlich unter den Fehlern  $\varepsilon_D^B$  und  $\varepsilon_M^B$  bei der Approximation mit Bernsteinpolynomen. Der Maximalfehler unterscheidet sich jeweils mindestens um 340%, der durchschnittliche Fehler jeweils mindestens um 200%. Die große Differenz kann man auch in den Abbildungen besonders um den Punkt  $x = 0$  optisch erkennen. Der relativ große

Unterschied entspricht auch unseren Erwartungen, da der Maximalfehler bei der Approximation mit Bernsteinpolynomen nur von der Ordnung  $n^{-1}$  ist, siehe Abschnitt 4.2.

### 8.1.3 Vergleich zur Polynominterpolation

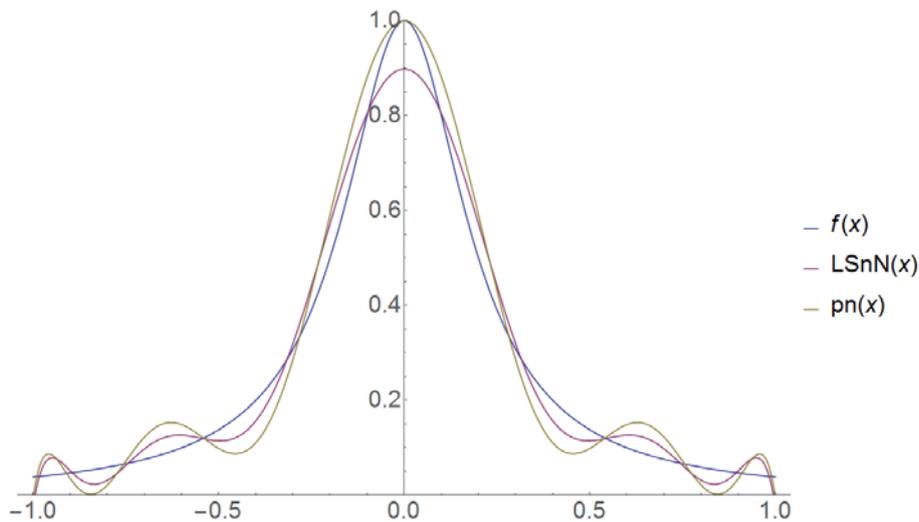


Abbildung 8.9

Funktionen:  $f, LS_n^N[f], p_n[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

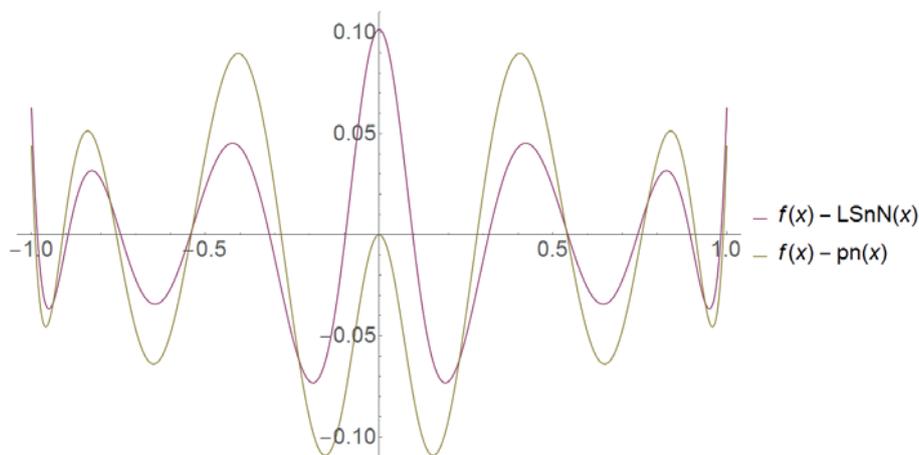


Abbildung 8.10

Funktionen:  $f - LS_n^N[f], f - p_n[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

Fehler:  $\varepsilon_M^D = 0.101513, \varepsilon_D^D = 0.0327692,$   
 $\varepsilon_M^T = 0.109153, \varepsilon_D^T = 0.0485122.$

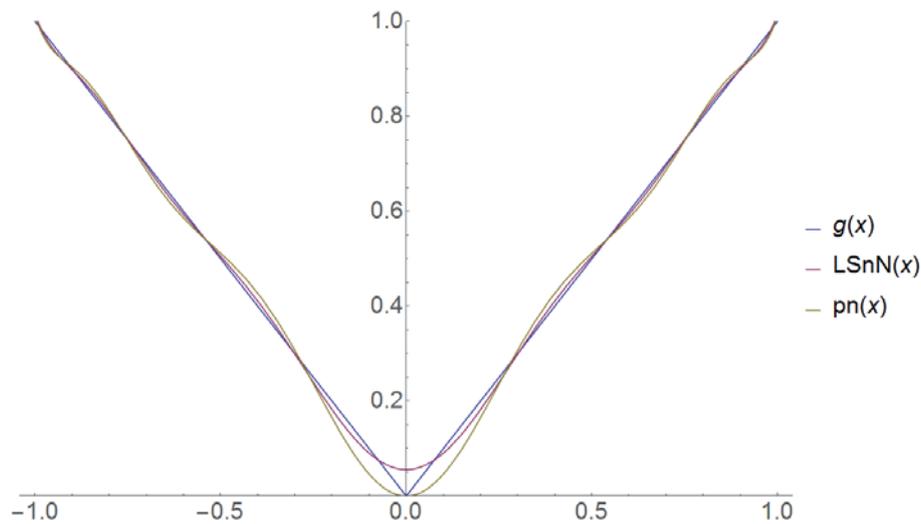


Abbildung 8.11

Funktionen:  $g, LS_n^N[g], p_n[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

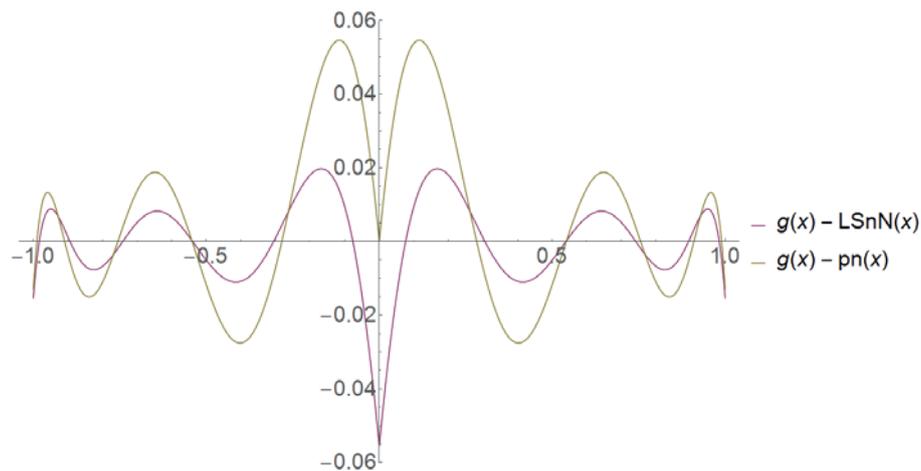


Abbildung 8.12

Funktionen:  $g - LS_n^N[g], g - p_n[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10, N = 2n(n + 1) = 220$ .

Fehler:  $\varepsilon_M^D = 0.0556922, \varepsilon_D^D = 0.00883205,$   
 $\varepsilon_M^T = 0.0546222, \varepsilon_D^T = 0.0191434$ .

Bei beiden Testfunktionen liegt der maximale Fehler  $\varepsilon_M^D$  bei der Methode der kleinsten Quadrate im diskreten Fall nah an dem Fehler  $\varepsilon_M^T$  bei der Polynominterpolation in Tschebyscheff-Stützstellen. Der Maximalfehler unterscheidet sich jeweils um höchstens 8% zuungunsten, der durchschnittliche Fehler jeweils sogar um mindestens 48% zugunsten der Methode der kleinsten Quadrate. Die Ergebnisse für die Runge-Funktion  $f$  verdeutlichen unsere Resultate aus Abschnitt 7.4. Allerdings fällt

auf, dass in beiden Fällen der durchschnittliche Fehler  $\varepsilon_D^D$  bei der Methode der kleinsten Quadrate signifikant geringer, als der durchschnittliche Fehler  $\varepsilon_D^T$  bei der Polynominterpolation, ist. Eine mögliche Erklärung könnte darin bestehen, dass die Methode der kleinsten Quadrate die Fehlerquadrate auf einem äquidistanten Gitter im Durchschnitt minimiert. In den Abbildungen fällt zudem eine starke Ähnlichkeit der „Funktionsverläufe“ beider Methoden auf. Insbesondere liegen die Nullstellen der Fehlerfunktionen sehr nah beieinander.

## 8.2 Variation der Stützstellenanzahl

In diesem Abschnitt untersuchen wir numerisch für unsere beiden Testfunktionen  $f$  und  $g$  die Methode der kleinsten Quadrate im diskreten Fall in Abhängigkeit des Verhältnisses  $N/n$ .

Zum einen verwenden wir die Methode mit einer Größenordnung der Stützstellenanzahl oberhalb von  $N_1 = 2n(n+1)$ , konkret mit  $N_2 = n^3$  und  $N_3 = n^4$ . Hier erwarten wir, wie schon am Anfang des Kapitels diskutiert, keine großen Veränderungen der jeweiligen Lösungen.

Zum anderen verwenden wir die Methode mit einer Stützstellenanzahl im Bereich zwischen  $N = n$  und  $N = 2n(n+1)$ , konkret mit  $N_1 = 2n(n+1)$ ,  $N_2 = n^{\frac{3}{2}}$  und  $N_3 = 2n$ . In diesem Fall sind wir vor allem daran interessiert, wie sich bei unseren beiden Testfunktionen ein kleineres Verhältnis  $N/n$  auf die Qualität der Approximation auswirkt.

Wir verwenden mit  $h = f$  aus Gleichung (8.1) beziehungsweise  $h = g$  aus Gleichung (8.2) und mit  $i = 1, 2, 3$  für die Fehler die folgenden Bezeichnungen:

$$\begin{aligned}\varepsilon_M^i &= \max_{\tilde{x} \in \tilde{X}} \left| h(\tilde{x}) - LS_n^{N_i}[h](\tilde{x}) \right|, \\ \varepsilon_D^i &= \frac{1}{\tilde{N}} \sum_{\tilde{x} \in \tilde{X}} \left| h(\tilde{x}) - LS_n^{N_i}[h](\tilde{x}) \right|.\end{aligned}$$

Dabei sind  $LS_n^{N_i}$  für  $i = 1, 2, 3$  die Operatoren der Methode der kleinsten Quadrate im diskreten Fall aus Unterabschnitt 4.3.2.

Im Folgenden sind die jeweiligen durchschnittlichen beziehungsweise maximalen relativen Fehler stets in Bezug auf  $\varepsilon_D^1$  beziehungsweise  $\varepsilon_M^1$  angegeben.

### 8.2.1 Vergleich mit größerer Stützstellenanzahl

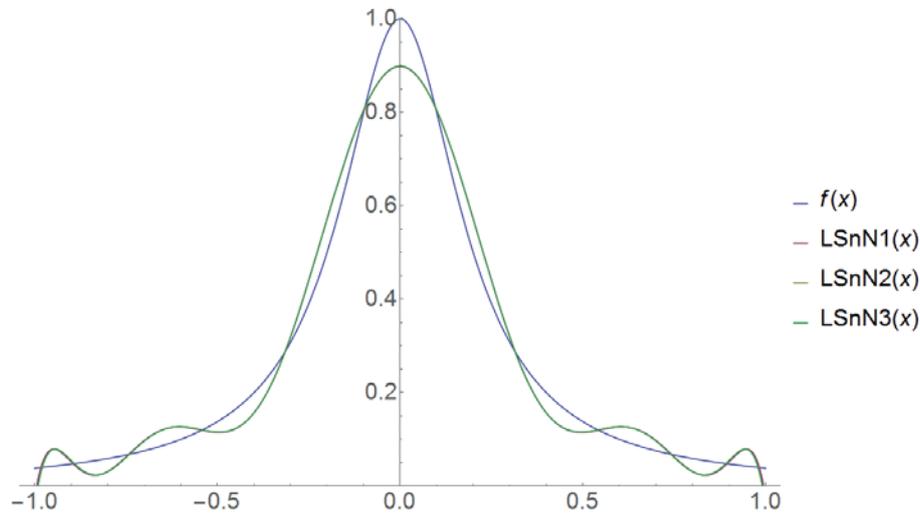


Abbildung 8.13

Funktionen:  $f, LS_n^{N_1}[f], LS_n^{N_2}[f], LS_n^{N_3}[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$N_1 = 2n(n+1) = 220, N_2 = n^3 = 1000, N_3 = n^4 = 10000.$

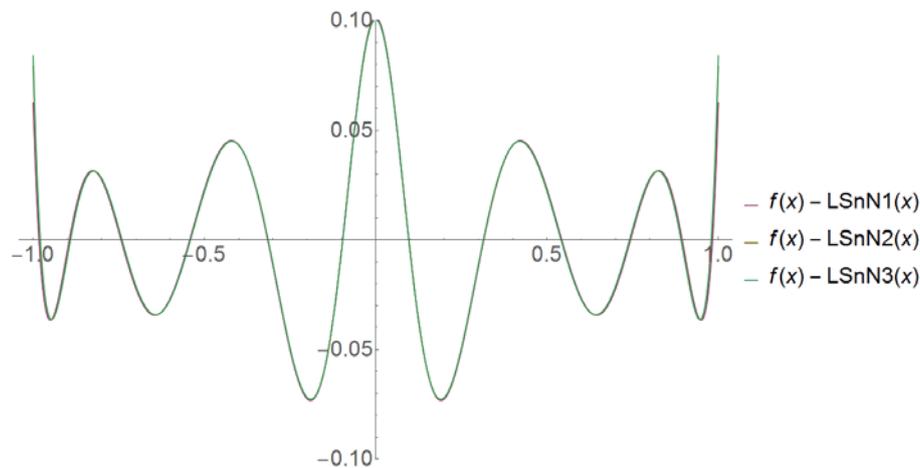


Abbildung 8.14

Funktionen:  $f - LS_n^{N_1}[f], f - LS_n^{N_2}[f], f - LS_n^{N_3}[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$N_1 = 2n(n+1) = 220, N_2 = n^3 = 1000, N_3 = n^4 = 10000.$

Fehler:  $\varepsilon_M^1 = 0.101513, \varepsilon_D^1 = 0.0327692,$

$\varepsilon_M^2 = 0.100799, \varepsilon_D^2 = 0.0327125,$

$\varepsilon_M^3 = 0.100598, \varepsilon_D^3 = 0.0327077.$

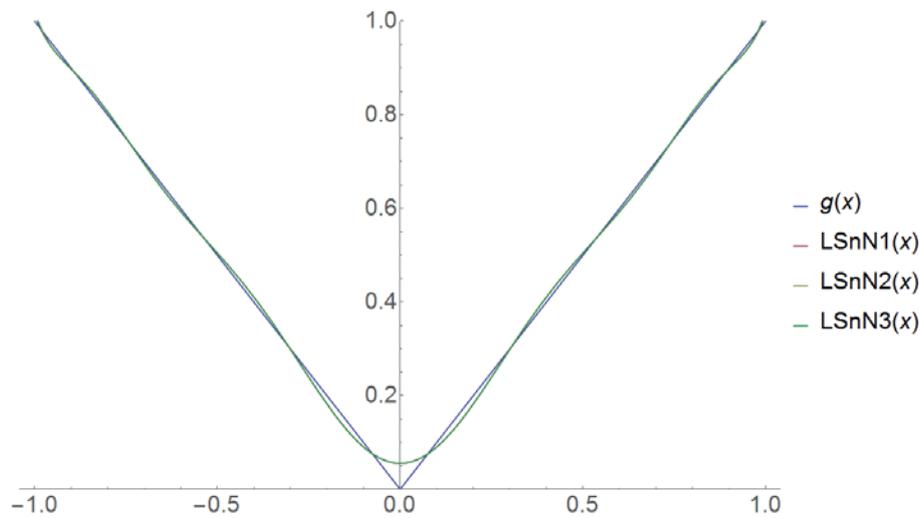


Abbildung 8.15

Funktionen:  $g, LS_n^{N_1}[g], LS_n^{N_2}[g], LS_n^{N_3}[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$N_1 = 2n(n+1) = 220, N_2 = n^3 = 1000, N_3 = n^4 = 10000.$

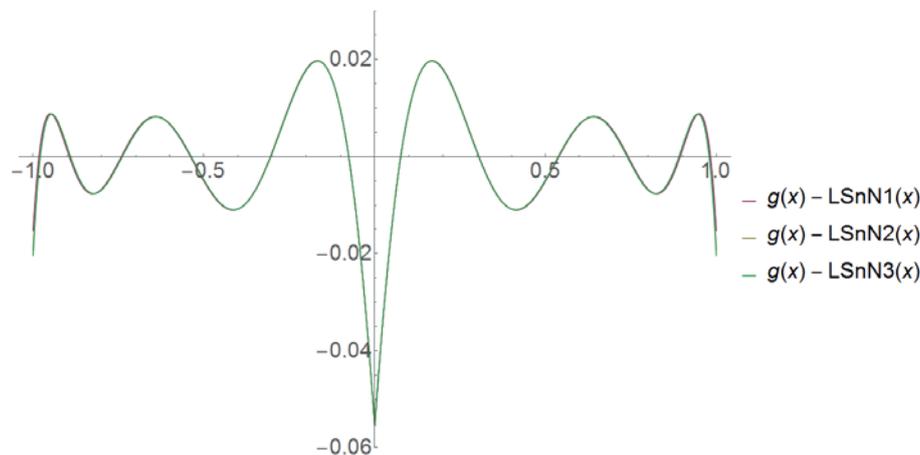


Abbildung 8.16

Funktionen:  $g - LS_n^{N_1}[g], g - LS_n^{N_2}[g], g - LS_n^{N_3}[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$N_1 = 2n(n+1) = 220, N_2 = n^3 = 1000, N_3 = n^4 = 10000.$

Fehler:  $\varepsilon_M^1 = 0.0556922, \varepsilon_D^1 = 0.00883205,$

$\varepsilon_M^2 = 0.0555671, \varepsilon_D^2 = 0.00882465,$

$\varepsilon_M^3 = 0.0555208, \varepsilon_D^3 = 0.00882635.$

Wie zu Beginn des Abschnitts erwartet, ist bei  $N_2$  und  $N_3$  zumindest aus praktischer Sicht der Unterschied vernachlässigbar klein. Der Maximalfehler unterscheidet sich jeweils um höchstens 1.0%, der durchschnittliche Fehler jeweils sogar nur um höchstens 0.2%.

## 8.2.2 Vergleich mit geringerer Stützstellenanzahl

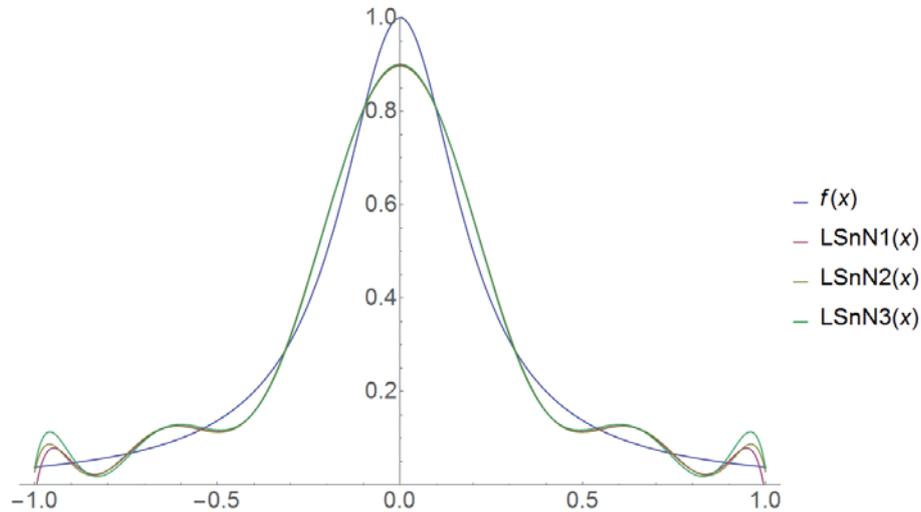


Abbildung 8.17

Funktionen:  $f, LS_n^{N_1}[f], LS_n^{N_2}[f], LS_n^{N_3}[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$$N_1 = 2n(n+1) = 220, N_2 = n^{\frac{3}{2}} = 32, N_3 = 2n = 20.$$

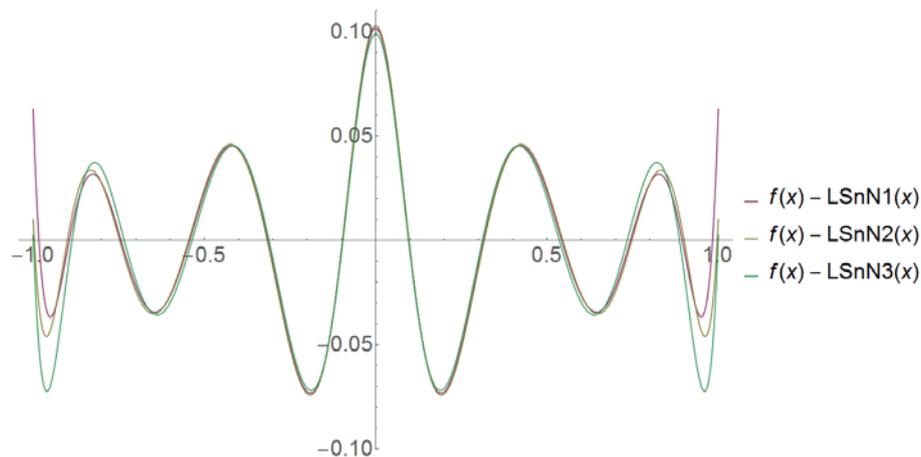


Abbildung 8.18

Funktionen:  $f - LS_n^{N_1}[f], f - LS_n^{N_2}[f], f - LS_n^{N_3}[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$$N_1 = 2n(n+1) = 220, N_2 = n^{\frac{3}{2}} = 32, N_3 = 2n = 20.$$

Fehler:  $\varepsilon_M^1 = 0.101513, \varepsilon_D^1 = 0.0327692,$   
 $\varepsilon_M^2 = 0.102780, \varepsilon_D^2 = 0.0338114,$   
 $\varepsilon_M^3 = 0.0987744, \varepsilon_D^3 = 0.0353820.$

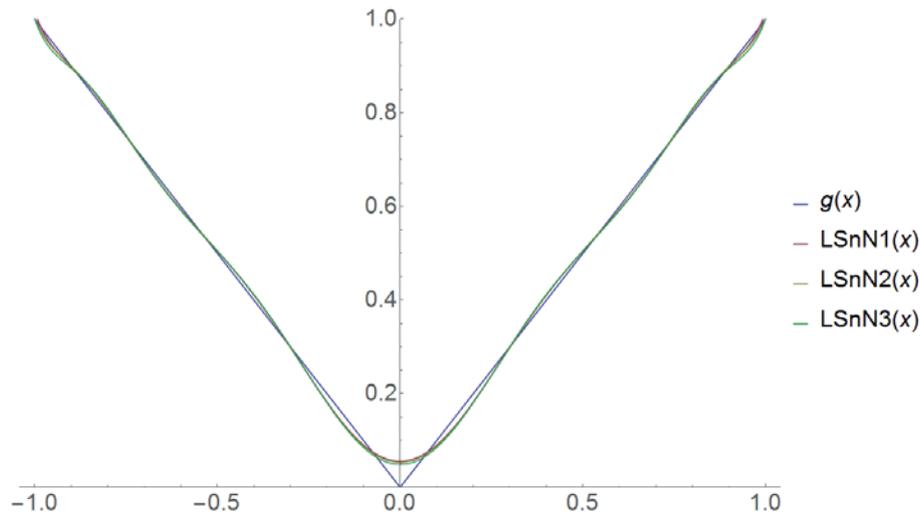


Abbildung 8.19

Funktionen:  $g, LS_n^{N_1}[g], LS_n^{N_2}[g], LS_n^{N_3}[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$$N_1 = 2n(n+1) = 220, N_2 = n^{\frac{3}{2}} = 32, N_3 = 2n = 20.$$

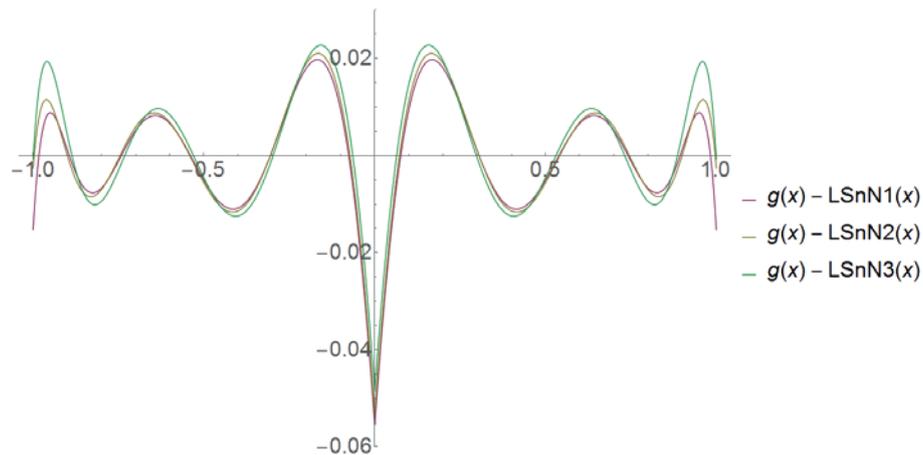


Abbildung 8.20

Funktionen:  $g - LS_n^{N_1}[g], g - LS_n^{N_2}[g], g - LS_n^{N_3}[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 10,$

$$N_1 = 2n(n+1) = 220, N_2 = n^{\frac{3}{2}} = 32, N_3 = 2n = 20.$$

Fehler:  $\varepsilon_M^1 = 0.0556922, \varepsilon_D^1 = 0.00883205,$

$$\varepsilon_M^2 = 0.0536703, \varepsilon_D^2 = 0.00930164,$$

$$\varepsilon_M^3 = 0.0488102, \varepsilon_D^3 = 0.0102552.$$

Sowohl die durchschnittlichen Fehler  $\varepsilon_D^1, \varepsilon_D^2$  und  $\varepsilon_D^3$  als auch die maximalen Fehler  $\varepsilon_M^1, \varepsilon_M^2$  und  $\varepsilon_M^3$  weichen nicht erheblich voneinander ab. Der Maximalfehler unterscheidet sich jeweils um höchstens 13%, der durchschnittliche Fehler jeweils um höchstens 17%.

Da die numerischen Ergebnisse für  $n = 10$  somit wenig aussagekräftig sind, verdoppeln wir im Folgenden den Polynomgrad von  $n = 10$  auf  $n = 20$  und belassen das jeweilige Verhältnis zur Anzahl der Stützstellen bei  $N_1 = 2n(n + 1)$ ,  $N_2 = n^{\frac{3}{2}}$  und  $N_3 = 2n$ :

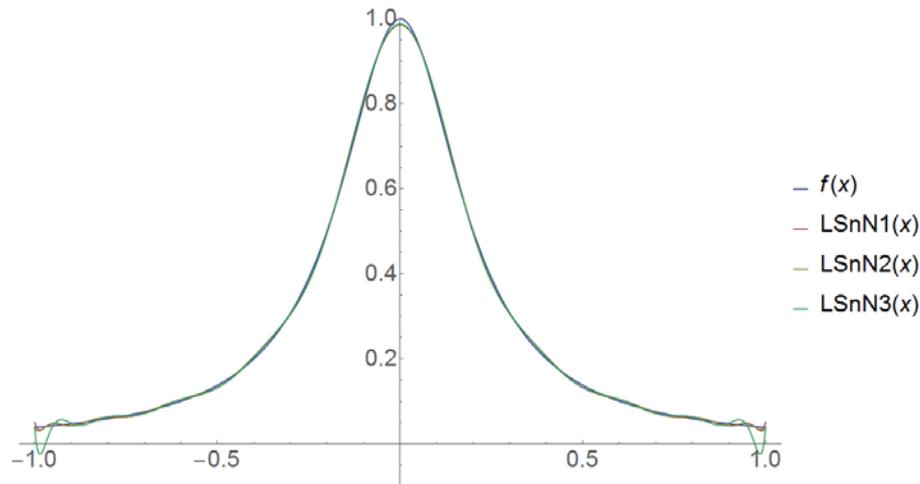


Abbildung 8.21

Funktionen:  $f, LS_n^{N_1}[f], LS_n^{N_2}[f], LS_n^{N_3}[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 20$ ,

$N_1 = 2n(n + 1) = 840, N_2 = n^{\frac{3}{2}} = 90, N_3 = 2n = 40$ .

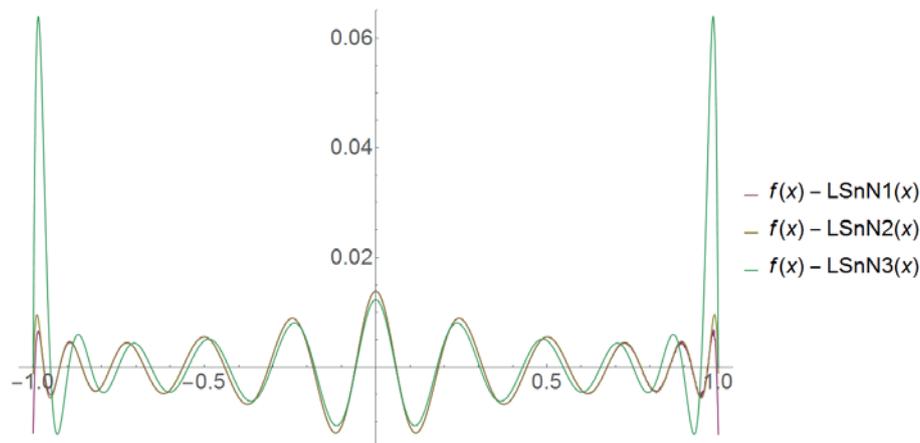


Abbildung 8.22

Funktionen:  $f - LS_n^{N_1}[f], f - LS_n^{N_2}[f], f - LS_n^{N_3}[f]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 20$ ,

$N_1 = 2n(n + 1) = 840, N_2 = n^{\frac{3}{2}} = 90, N_3 = 2n = 40$ .

Fehler:  $\varepsilon_M^1 = 0.0138676, \varepsilon_D^1 = 0.00449756$ ,

$\varepsilon_M^2 = 0.0138927, \varepsilon_D^2 = 0.00458770$ ,

$\varepsilon_M^3 = 0.0638648, \varepsilon_D^3 = 0.00612865$ .

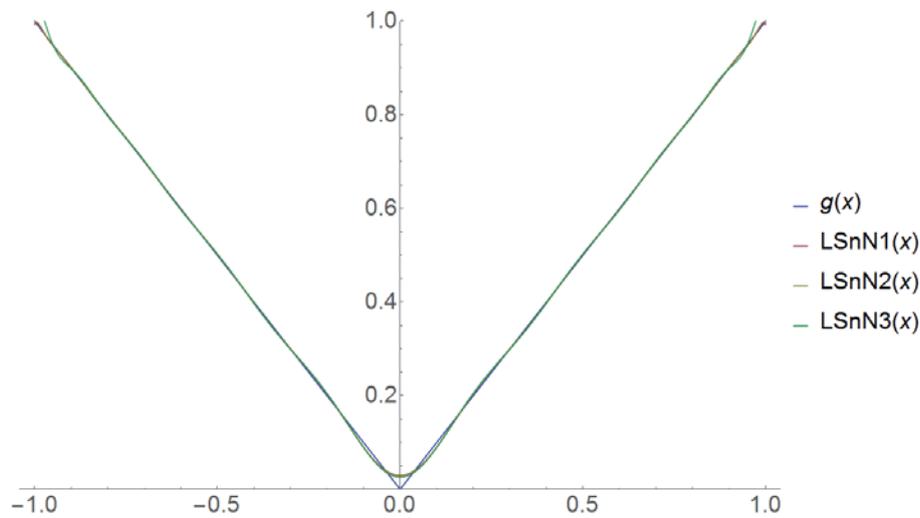


Abbildung 8.23

Funktionen:  $g, LS_n^{N_1}[g], LS_n^{N_2}[g], LS_n^{N_3}[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 20$ ,

$$N_1 = 2n(n+1) = 840, N_2 = n^{\frac{3}{2}} = 90, N_3 = 2n = 40.$$

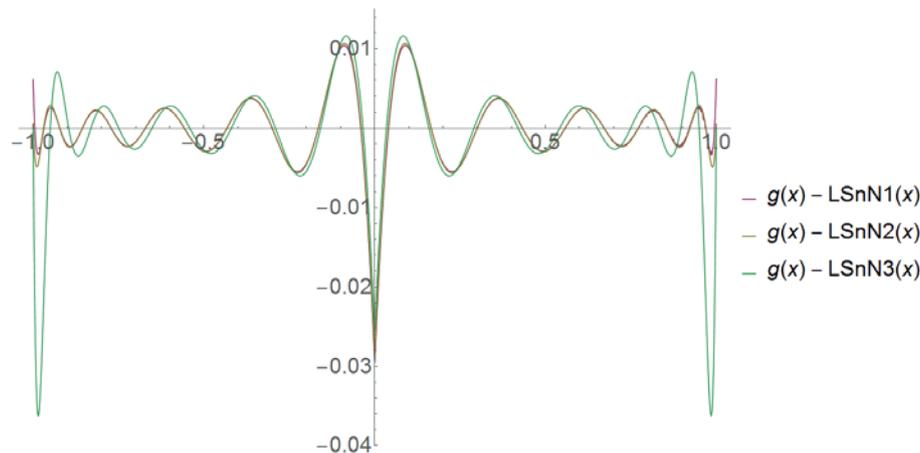


Abbildung 8.24

Funktionen:  $g - LS_n^{N_1}[g], g - LS_n^{N_2}[g], g - LS_n^{N_3}[g]$ .

Parameter:  $\alpha = 0, \beta = 0, n = 20$ ,

$$N_1 = 2n(n+1) = 840, N_2 = n^{\frac{3}{2}} = 90, N_3 = 2n = 40.$$

Fehler:  $\varepsilon_M^1 = 0.0296598, \varepsilon_D^1 = 0.00305392$ ,

$$\varepsilon_M^2 = 0.0291157, \varepsilon_D^2 = 0.00313687,$$

$$\varepsilon_M^3 = 0.0362119, \varepsilon_D^3 = 0.00434942.$$

Obwohl wir bei den Verhältnissen  $N_1 = 2n(n+1)$ ,  $N_2 = n^{\frac{3}{2}}$  und  $N_3 = 2n$  geblieben sind, kommt es nun im Fall von  $N_3 = 2n$  zu erheblichen Fehlern am Rand des Intervalls. Der Maximalfehler unterscheidet sich in diesem Fall bei der Funktion  $g$



um etwa 22% und bei der Funktion  $f$  sogar um etwa 360%. Auch der durchschnittliche Fehler unterscheidet sich jeweils um mindestens 36%. Zwischen den Fällen  $N_1 = 2n(n+1)$  und  $N_2 = n^{\frac{3}{2}}$  sind aber trotz der Verdopplung des Polynomgrades weiterhin keine großen Unterschiede sichtbar. Der Maximalfehler unterscheidet sich jeweils um höchstens 1.9%, der durchschnittliche Fehler jeweils um höchstens 2.1%. Insbesondere im Bereich  $(-0.9, 0.9) \subset [-1, 1]$  werden in allen Fällen sehr gute Ergebnisse bei der Approximation der Funktionen  $f$  und  $g$  erzielt.

Während in den Fällen  $N = n^3$  und  $N = n^4$  bei den vorgestellten Beispielen kein signifikanter Qualitätszuwachs im Vergleich zum Fall  $N = 2n(n+1)$  zu verzeichnen ist - und im Fall  $N = 2n$  ein deutlicher Qualitätsverlust auftritt - bleibt eine Beurteilung im Fall  $N = n^{\frac{3}{2}}$  unbefriedigend. Eine Vermutung, bei welchem asymptotisch konstanten Verhältnis  $n^\alpha/N$  mit  $\alpha \in (1, 2)$  sich noch gute Approximationen erzielen lassen, bedarf vermutlich erheblichen zusätzlichen Aufwand.

## 9 Ausblick

In diesem Kapitel geben wir einen Ausblick auf eine mögliche Anwendung der Methode der kleinsten Quadrate im diskreten Fall bei Spektralen-Differenzen-Verfahren zur Lösung hyperbolischer Erhaltungsgleichungen. Eine detaillierte Einführung in die Theorie hyperbolischer Erhaltungsgleichungen findet man in den Werken [5] und [51]. Zur ausführlichen Darstellung und Verwendung von Spektralen-Differenzen-Verfahren verweisen wir insbesondere auf die Doktorarbeiten [61] und [78] sowie die Standardliteratur [22].

Viele Phänomene der Natur lassen sich sehr gut durch Systeme von partiellen Differentialgleichungen beschreiben. Eine besondere Klasse sind dabei die sogenannten hyperbolischen Erhaltungsgleichungen. Mit ihnen lassen sich beispielsweise Strömungen gut modellieren. Da im Allgemeinen keine exakten Lösungen für sie bekannt sind, ist man an numerischen Lösungsmethoden interessiert. Eine solche Methode ist das Spektrale-Differenzen-Verfahren. Im Folgenden geben wir zunächst eine kurze Einführung in hyperbolische Erhaltungsgleichungen und anschließend beschreiben wir Spektrale-Differenzen-Verfahren.

Ein System von hyperbolischen Erhaltungsgleichungen ist gegeben durch

$$\frac{\partial}{\partial t} \mathbf{u}(\mathbf{x}, t) + \sum_{j=1}^d \frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{u}(\mathbf{x}, t)) = 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad t > 0. \quad (9.1)$$

Hierbei ist  $\mathbf{u} : \mathbb{R}^d \times \mathbb{R}_0^+ \rightarrow \mathbb{R}^m$  der  $m$ -dimensionale Vektor der Erhaltungsvariablen und die  $\mathbf{f}_j : \mathbb{R}^m \rightarrow \mathbb{R}^m$ ,  $j = 1, \dots, d$  sind die Flussfunktionen. Dabei bezeichnen  $\mathbf{x} \in \mathbb{R}^d$  die Variablen des Raumes und  $t \in \mathbb{R}^+$  die Variable der Zeit. Also ist  $d$  die Raumdimension und  $m$  ist die Anzahl der Erhaltungsvariablen.

Ein wichtiges Beispiel für ein System von hyperbolischen Erhaltungsgleichungen sind die Euler-Gleichungen. Mit ihnen werden in der Praxis gasdynamische Prozesse modelliert. Sie setzen sich physikalisch aus der Erhaltung von Masse, Impuls und Energie zusammen und sind ein Spezialfall der Navier-Stokes-Gleichung, bei der noch die Wärmeleitung und die Viskosität berücksichtigt werden. Im zweidimensionalen Fall  $d = 2$  ist das System durch die folgenden partiellen Differentialgleichungen gegeben

$$\frac{\partial}{\partial t} \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix} + \frac{\partial}{\partial x} \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ \rho uH \end{pmatrix} + \frac{\partial}{\partial y} \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ \rho vH \end{pmatrix} = 0. \quad (9.2)$$

Dabei sind die Dichte  $\rho$ , die Geschwindigkeiten  $u$  in  $x$ - sowie  $v$  in  $y$ -Richtung und der Druck  $p$  physikalische Größen. Der Vektor  $\mathbf{u} = (\rho, \rho u, \rho v, \rho E)^T$  ist der Vektor der Erhaltungsvariablen. Die Energie  $E = e + \frac{u^2+v^2}{2}$  setzt sich aus der kinetischen Energie  $\frac{u^2+v^2}{2}$  und der spezifischen inneren Energie  $e$  zusammen. Diese steht mit der Enthalpie  $H$  in der Beziehung  $H = E + \frac{p}{\rho}$ . Weiterhin gilt die Relation  $p = \rho(\gamma - 1)e$ .

Ein weiteres, für Testzwecke häufig verwendetes, sehr einfaches Beispiel einer nicht-linearen hyperbolischen Erhaltungsgleichung ist die sogenannte Burgers-Gleichung, welche in einer Raumdimension durch

$$\frac{\partial}{\partial t} u + \frac{\partial}{\partial x} \left( \frac{1}{2} u^2 \right) = 0 \quad (9.3)$$

gegeben ist.

In der Regel werden Anfangsbedingungen  $\mathbf{u}_0 : \mathbb{R}^d \rightarrow \mathbb{R}^m$  an die Erhaltungsvariablen gestellt. Das Anfangswertproblem für hyperbolische Erhaltungsgleichungen ist dann gegeben durch

$$\begin{aligned} \frac{\partial}{\partial t} \mathbf{u}(\mathbf{x}, t) + \sum_{j=1}^d \frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{u}(\mathbf{x}, t)) &= 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad t > 0, \\ \mathbf{u}(\mathbf{x}, 0) &= \mathbf{u}_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d. \end{aligned} \quad (9.4)$$

Die Fragen nach der Existenz und der Eindeutigkeit von Lösungen  $\mathbf{u} : \mathbb{R}^d \times \mathbb{R}_0^+ \rightarrow \mathbb{R}^m$  dieses Anfangswertproblems werden an dieser Stelle nicht weiter behandelt. Wir verweisen auf [51] und [61].

Im Folgenden beschäftigen wir uns mit der Frage nach einer numerischen Lösung dieses Anfangswertproblems. Es gibt verschiedene Möglichkeiten hyperbolische Erhaltungsgleichungen numerisch zu lösen. Siehe dazu beispielsweise [29] und [75]. Wir konzentrieren uns jedoch auf Spektrale-Differenzen-Verfahren. Hierbei soll folgendes System von Erhaltungsgleichungen

$$\frac{\partial}{\partial t} \mathbf{u}(\mathbf{x}, t) + \sum_{j=1}^d \frac{\partial}{\partial x_j} \mathbf{f}_j(\mathbf{u}(\mathbf{x}, t)) = 0, \quad (9.5)$$

für  $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ ,  $t > t_0 \geq 0$  mit den Flussfunktionen  $\mathbf{f}_j : \mathbb{R}^m \rightarrow \mathbb{R}^m$ ,  $j = 1, \dots, d$  numerisch gelöst werden. Dabei diskretisiert man sowohl in der Zeit als auch im Raum.

Wir betrachten zunächst die Zeitdiskretisierung. Hierbei betrachtet man Gleichung (9.5) für festes  $\mathbf{x}$  und erhält dadurch eine gewöhnliche Differentialgleichung in der Zeit  $t$ . Es gibt verschiedene Ansätze gewöhnliche Differentialgleichungen numerisch zu lösen, siehe beispielsweise [40]. Die Entwicklung und Analyse solcher Verfahren ist ein eigenes Forschungsgebiet. Dabei unterscheidet man zwischen expliziten und impliziten Verfahren. Werden zur numerischen Lösung  $U^{k+1}$  im  $(k+1)$ -ten

Zeitschritt ausschließlich Werte  $U^i$  von vorhergehenden Zeitschritten  $0, \dots, k$  verwendet, so spricht man von expliziten Verfahren. Andernfalls spricht man von impliziten Verfahren. Sehr bekannte Verfahren sind das explizite sowie das implizite Eulerverfahren, welche durch

$$U^{k+1}(\mathbf{x}) = U^k(\mathbf{x}) - \int_{t_k}^{t_{k+1}} \left( \sum_{j=1}^d \frac{\partial}{\partial x_j} \mathbf{f}_j(U^k(\mathbf{x})) \right), \quad (9.6)$$

$$U^{k+1}(\mathbf{x}) = U^k(\mathbf{x}) - \int_{t_k}^{t_{k+1}} \left( \sum_{j=1}^d \frac{\partial}{\partial x_j} \mathbf{f}_j(U^{k+1}(\mathbf{x})) \right), \quad (9.7)$$

mit der Anfangsbedingung  $U^0(\mathbf{x}) = \mathbf{u}(\mathbf{x}, t_0)$  definiert sind.

Ein weiteres, sehr bekanntes, explizites Zeitschrittverfahren ist das sogenannte Runge-Kutta-Verfahren. Ein ausgewähltes Runge-Kutta-Verfahren vierter Ordnung aus [23] wurde in den Arbeiten [61] und [78] zur Zeitintegration verwendet. Solche Verfahren wurden beispielsweise in [39] genauer untersucht.

Im Folgenden betrachten wir die räumliche Diskretisierung. Bei der Spektralendifferenzen-Methode wird die Lösung  $\mathbf{u}(\mathbf{x}, t)$  in jedem Zeitschritt durch eine abgeschnittene Reihenentwicklung der Form

$$\mathbf{u}(\mathbf{x}) \approx \mathbf{u}_n(\mathbf{x}) = \sum_{l=0}^n \hat{\mathbf{u}}_l \phi_l(\mathbf{x}) \quad (9.8)$$

approximiert. Hierbei sind  $\phi_0, \dots, \phi_n$  linear unabhängige Funktionen auf  $\Omega$  und  $\hat{\mathbf{u}}_0, \dots, \hat{\mathbf{u}}_n$  die  $m$ -Vektoren der Koeffizienten. In den Arbeiten [61] beziehungsweise [78] wurde der Fall  $m = 2$  behandelt und das zugrunde liegende Gebiet  $\Omega$  trianguliert. Dabei wurden die einzelnen Dreiecke der Triangulierung auf das Einheitsdreieck

$$\mathbb{T} = \left\{ (\xi, \eta) \in \mathbb{R}^2 : 0 \leq \xi, \eta \leq 1, \xi + \eta \leq 1 \right\} \quad (9.9)$$

beziehungsweise

$$\mathbb{T}^2 = \left\{ (r, s) \in \mathbb{R}^2 : -1 \leq r, s \leq 1, r + s \leq 0 \right\} \quad (9.10)$$

transformiert. Die im Folgenden definierten PKD-Polynome  $\varphi_{l,m}$  beziehungsweise APK-Polynome  $A_{l,m}$  wurden für die Funktionen  $\phi_0, \dots, \phi_n$  aus Gleichung (9.8) in [78] beziehungsweise in [61] verwendet. Für  $\alpha, \beta, \gamma \in \mathbb{N}$  mit  $\gamma > \alpha + \beta - 1$  sind die Appell-Proriol-Koornwinder-Polynome, oder kurz APK-Polynome, für  $m, l \in \mathbb{N}_0$  auf  $\mathbb{T}$  definiert durch

$$A_{m,l}(x, y) := P_m^{\alpha-1, a_l} (1 - 2x) P_l^{p, \beta-1} \left( \frac{2y}{1-x} - 1 \right) (1-x)^l. \quad (9.11)$$

Hierbei sind der Kürze halber  $p := \gamma - \alpha - \beta$  und  $a_l := p + \beta + 2l$ . Die APK-Polynome  $A_{l,m}$  bilden ein System von stetigen orthogonalen Polynomen auf  $\mathbb{T}$  bezüglich des Skalarprodukts

$$(f, g)_h := \int_{\mathbb{T}} f(\mathbf{x}) g(\mathbf{x}) h(\mathbf{x}) d\mathbf{x}, \quad (9.12)$$

wobei die Gewichtsfunktion  $h$  gegeben ist durch

$$h(x, y) := x^{\alpha-1} y^{\beta-1} (1 - x - y)^{\gamma-\alpha-\beta}. \quad (9.13)$$

Transformiert man die APK-Polynome  $A_{l,m}$  auf das Dreieck  $\mathbb{T}^2$ , so ergeben sich unter Verwendung der Parameter  $\alpha = 1$ ,  $\beta = 1$  und  $\gamma = 2$  die Proriol-Koornwinder-Dubiner-Polynome, oder kurz PKD-Polynome, durch

$$\varphi_{m,l}(r, s) := P_l^{0,0} \left( \frac{2(1+r)}{1-s} - 1 \right) \left( \frac{1-s}{2} \right)^l P_m^{2l+1,0}(s). \quad (9.14)$$

Diese bilden ein System von stetigen orthogonalen Polynomen auf  $\mathbb{T}^2$  bezüglich der Gewichtsfunktion 1.

Zur Berechnung der Koeffizienten  $\hat{\mathbf{u}}_l$  in Gleichung (9.8) wurde in den Arbeiten [61] und [78] in den sogenannten 2D-Gauß-Lobatto-Punkten aus den Arbeiten [15] und [16] interpoliert. Demnach sind in einer Dimension die Gauß-Lobatto-Punkte im Intervall  $[0, 1]$  folgendermaßen definiert: Zunächst sei für  $n \in \mathbb{N}_0$  das Lobatto-Polynom  $n$ -ten Grades definiert durch  $\text{Lo}_n(x) := P'_{n+1}(x)$ . Dabei ist  $P_{n+1}$  das Legendre-Polynom  $(n+1)$ -ten Grades gemäß Definition 3.9. Für ein  $n \in \mathbb{N}$  sind die Gauß-Lobatto-Punkte  $v_i$ ,  $i = 0, \dots, n$  definiert durch  $v_0 := 0$ ,  $v_n := 1$  und

$$v_i := \frac{1}{2} (1 + t_i) \quad \text{für } i = 1, \dots, n-1, \quad (9.15)$$

wobei  $t_i$ ,  $i = 1, \dots, n-1$  die Nullstellen des Lobatto-Polynoms  $\text{Lo}_{n-1}$  sind. Damit lassen sich die 2D-Gauß-Lobatto-Punkte nun folgendermaßen definieren: Zu einem  $n \in \mathbb{N}$  seien  $v_i$ ,  $i = 0, \dots, n$  die Gauß-Lobatto-Punkte. Die 2D-Gauß-Lobatto-Punkte sind auf  $\mathbb{T}$  gegeben durch

$$(\xi_i, \eta_j) := \left( \frac{1}{3} (1 + 2v_i - v_j - v_k), \frac{1}{3} (1 + 2v_j - v_i - v_k) \right), \quad (9.16)$$

für  $i = 0, \dots, n$ ,  $j = 0, \dots, n-i$  und  $k = n-i-j$ .

Eine Entwicklung gemäß (9.8) mit einem Orthogonalsystem  $\{\phi_0, \dots, \phi_n\}$  legt jedoch nahe die Koeffizienten  $\hat{\mathbf{u}}_l$  entsprechend des zugehörigen Skalarprodukts zu bestimmen. Im obigen Fall der APK-Polynome  $A_{l,m}$  könnte man also die Koeffizienten folgendermaßen ermitteln:

$$\hat{\mathbf{u}}_{l,m} = \frac{(\mathbf{u}, A_{l,m})_h}{(A_{l,m}, A_{l,m})_h}. \quad (9.17)$$

Um die rechte Seite von (9.17) zu berechnen, müsste jedoch  $\mathbf{u}$  auf  $\mathbb{T}$  bekannt sein. Da dies nicht der Fall ist, bliebe nur die Möglichkeit die entsprechenden Integrale mit Hilfe von Quadraturverfahren anzunähern. Dieses Problem kann aber vollständig umgangen werden, wenn man ein Orthogonalsystem bezüglich eines diskreten Skalarprodukts verwendet.

Die Hahn-Polynome  $Q_n$  lassen sich ebenfalls auf mehrere Dimensionen erweitern. In [44] wurden die Hahn-Polynome folgendermaßen auf zwei Dimensionen erweitert, man vergleiche auch die Ausführungen in [80]. Wir betrachten dazu das Dreieck

$$\mathbb{T}_N = \left\{ \mathbf{x} = (x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x + y \leq N \right\} \quad (9.18)$$

und das darauf definierte Gitter

$$V = \left\{ \mathbf{x} = (x, y) \in \mathbb{Z}^2 : x \geq 0, y \geq 0, x + y \leq N \right\}. \quad (9.19)$$

Seien  $\sigma_1, \sigma_2, \sigma_3 > -1$  und  $N \in \mathbb{N}_0$ . Die Hahn-Polynome in zwei Variablen  $\phi_{l,m}$  sind definiert durch

$$\begin{aligned} & \phi_{l,m}(\mathbf{x}; \sigma_1, \sigma_2, \sigma_3, N) \\ & := Q_l(x; \sigma_1, \sigma_2 + \sigma_3 + 2m + 1, N - l) (-N + x)_m Q_m(y; \sigma_2, \sigma_3, N - x), \end{aligned} \quad (9.20)$$

für jedes  $n = 0, \dots, N$  und  $l \in \mathbb{N}_0$  mit  $l \leq n$  und  $m = n - l$ .

Die Hahn-Polynome in zwei Variablen sind orthogonal auf  $\mathbb{T}_N$  bezüglich des Skalarprodukts

$$\langle f, g \rangle_W := \sum_{\mathbf{x} \in V} f(\mathbf{x}) g(\mathbf{x}) W(\mathbf{x}), \quad (9.21)$$

wobei die Gewichtsfunktion  $W$  gegeben ist durch

$$W(\mathbf{x}) := \binom{x + \sigma_1}{\sigma_1} \binom{y + \sigma_2}{\sigma_2} \binom{N - x - y + \sigma_3}{\sigma_3}. \quad (9.22)$$

Verwendet man bei der Zeitintegration die Punkte aus  $V$ , so kann die Rekonstruktion im Raum gemäß (9.8) unter Verwendung der Hahn-Polynome in zwei Variablen  $\phi_{l,m}$  durch

$$\hat{\mathbf{u}}_{l,m} = \frac{\langle \mathbf{u}, \phi_{l,m} \rangle_W}{\langle \phi_{l,m}, \phi_{l,m} \rangle_W} \quad (9.23)$$

erfolgen.

Da Lösungen zeitabhängiger hyperbolischer Erhaltungsgleichungen trotz stetiger Anfangsdaten Unstetigkeiten nach endlicher Zeit entwickeln können, muss dies im numerischen Lösungsverfahren berücksichtigt werden. Denn bei der Rekonstruktion einer unstetigen Funktion durch eine abgebrochene Reihentwicklung gemäß (9.8) können an den Unstetigkeitsstellen Oszillationen auftreten. Dieses Verhalten wird als *Gibbs Phänomen* bezeichnet. Eine ausführliche Darstellung dieser Problematik findet man beispielsweise in [43]. Zur Reduktion dieser Oszillationen werden sogenannte Filtertechniken angewendet, verschiedene Ansätze findet man etwa in [22] und [45]. Die in den Arbeiten [78] beziehungsweise [61] verwendeten Filtertechniken basieren auf den Orthogonalsystemen bestehend aus den PKD-Polynomen beziehungsweise APK-Polynomen. Die Konstruktion der Filter geht dabei auf die zugehörigen Differentialgleichungen der Orthogonalsysteme im zweidimensionalen Fall entsprechend

Definition 3.5 im eindimensionalen Fall zurück. Die Hahn-Polynome erfüllen eine Differenzgleichung gemäß Definition 3.14. Die Hahn-Polynome in zwei Variablen erfüllen ebenfalls eine Differenzgleichung. Dazu benötigen wir die folgenden Differenzenoperatoren:

$$\Delta_i f(\mathbf{x}) = f(\mathbf{x} + \mathbf{e}_i) - f(\mathbf{x}), \quad (9.24)$$

$$\nabla_i f(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x} - \mathbf{e}_i), \quad (9.25)$$

für  $i = 1, 2$  mit den Einheitsvektoren  $\mathbf{e}_1 = (1, 0)$  und  $\mathbf{e}_2 = (0, 1)$ . Nach [80] erfüllen die Hahn-Polynome in zwei Variablen  $\phi_{l,m}$  die Differenzgleichung:

$$\begin{aligned} & x(N - x + \sigma_2 + \sigma_3 + 2) \Delta_1 \nabla_1 u - y(x + \sigma_1 + 1) \Delta_1 \nabla_2 u \\ & - x(y + \sigma_2 + 1) \Delta_2 \nabla_1 u + y(N - y + \sigma_1 + \sigma_3 + 2) \Delta_2 \nabla_2 u \\ & + [(N - x)(\sigma_1 + 1) - x(\sigma_2 + \sigma_3 + 2)] \Delta_1 u \\ & + [(N - y)(\sigma_2 + 1) - y(\sigma_1 + \sigma_3 + 2)] \Delta_2 u \\ & = -n(n + \sigma_1 + \sigma_2 + \sigma_3 + 2)u. \end{aligned} \quad (9.26)$$

Auf dieser Grundlage lassen sich die entsprechenden Filtertechniken eventuell auf den diskreten Fall übertragen.

Der hier vorgestellte Weg zur „Umstellung“ der bisherigen klassischen (stetigen) Methode auf die von Anfang an diskrete Methode erscheint durch die in Kapitel 6 und 7 präsentierten Ergebnisse vielversprechend zu sein. Es ist beabsichtigt Untersuchungen hierzu an anderer Stelle fortzusetzen.



# Literaturverzeichnis

- [1] Abdulle, A.; Wanner, G.: »200 years of least squares method«. *Elemente der Mathematik* Vol. 57, Iss. 2 (2002), S. 45–60.
- [2] Abramowitz, M.; Stegun, I. A.: *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Publications, Inc., 1964.
- [3] Agakhanov, S. A.; Natanson, G. I.: »Approximation of functions by Fourier-Jacobi sums«. *Doklady Akademii Nauk SSSR* 166, No. 1 (1966), S. 9–10.
- [4] Alexits, G.: *Konvergenzprobleme der Orthogonalreihen*. VEB Deutscher Verlag der Wissenschaften, 1960.
- [5] Ansorge, R.; Sonar, T.: *Mathematical Models of Fluid Dynamics*. Wiley-VCH Verlag, 2009.
- [6] Arens, T.; Hettlich, F.; Karpfinger, C.; Kockelkorn, U.; Lichtenegger, K.; Stachel, H.: *Mathematik*. Springer-Verlag, 2015.
- [7] Atkinson, K. E.: *An Introduction to Numerical Analysis*. Second Edition. Wiley, 1989.
- [8] Badkov, V. M.: »Estimates of Lebesgue functions and remainders of Fourier-Jacobi series«. *Siberian Mathematical Journal* Vol. 9, Iss. 6 (1968), S. 947–962.
- [9] Barnard, R. W.; Dahlquist, G.; Pearce, K.; Reichel, L.; Richards, K. C.: »Gram polynomials and the Kummer function«. *Journal of Approximation Theory* Vol. 94, Iss. 1 (1998), S. 128–143.
- [10] Belen'kii, A. M.: »Uniform convergence of the Fourier-Jacobi series on the orthogonality segment«. *Mathematical notes of the Academy of Sciences of the USSR* Vol. 46, Iss. 6 (1989), S. 901–906.
- [11] Berman, D. L.: »The impossibility of constructing a linear polynomial operator furnishing an approximation within the order of the best approximation«. *Doklady Akademii Nauk SSSR* Vol. 120, No. 6 (1958), S. 1175–1177.
- [12] Bernstein, S. N.: »Démonstration du théoreme de Weierstrass fondée sur le calcul des probabilités«. *Comm. Soc. Math. Kharkov* Vol. 13 (1912), S. 1–2.
- [13] Bernstein, S. N.: »On the best approximation of continuous functions by polynomials of a given degree«. *Comm. Soc. Math. Kharkow* Vol. 2, No. 13 (1912), S. 49–194.
- [14] Björck, Å.: *Numerical Methods for Least Squares Problems*. SIAM, 1996.

- [15] Blyth, M. G.; Luo, H.; Pozrikidis, C.: »A comparison of interpolation grids over the triangle or the tetrahedron«. *Journal of Engineering Mathematics* Vol. 56, Iss. 3 (2006), S. 263–272.
- [16] Blyth, M. G.; Pozrikidis, C.: »A Lobatto interpolation grid over the triangle«. *IMA Journal of Applied Mathematics* Vol. 71, Iss. 1 (2006), S. 153–169.
- [17] Bojanic, R.; Vuilleumier, M.: »On the rate of convergence of Fourier-Legendre series of functions of bounded variation«. *Journal of Approximation Theory* Vol. 31, Iss. 1 (1981), S. 67–79.
- [18] Brass, H.: *Quadraturverfahren*. Vandenhoeck & Ruprecht, 1977.
- [19] Brass, H.: »Approximation durch Teilsummen von Orthogonalpolynomreihen«. In: *Numerische Methoden der Approximationstheorie*. Hrsg. von Collatz, L.; Meinardus, G.; Werner, H. Bd. 52. Springer-Verlag, 1980, S. 69–83.
- [20] Brass, H.: »Error estimates for least squares approximation by polynomials«. *Journal of Approximation Theory* Vol. 41, Iss. 4 (1984), S. 345–349.
- [21] Brass, H.; Petras, K.: *Quadrature Theory; The Theory of Numerical Integration on a Compact Interval*. Bd. 178. Mathematical Surveys and Monographs. American Mathematical Society, 2011.
- [22] Canuto, C.; Hussaini, M. Y.; Quarteroni, A.; Zang, T. A.: *Spectral Methods in Fluid Dynamics*. Springer-Verlag, 1988.
- [23] Carpenter, M. H.; Kennedy, C. A.: *Fourth-order 2N-storage Runge-Kutta schemes*. Techn. Ber. 109112. NASA Technical Memorandum, 1994.
- [24] Chihara, T. S.: *An Introduction to Orthogonal Polynomials*. Gordon and Breach, 1978.
- [25] Davis, P. J.: *Interpolation and Approximation*. Dover Publications, Inc., 1975.
- [26] Dette, H.: »New bounds for Hahn and Krawtchouk polynomials«. *SIAM Journal on Mathematical Analysis* Vol. 26, Iss. 6 (1995), S. 1647–1659.
- [27] Durand, L.: »Nicholson-type integrals for products of Gegenbauer functions and related topics«. In: *Theory and Application of Special Functions*. Hrsg. von Askey, R. A. Academic Press, Inc., 1975, S. 353–374.
- [28] Eisinberg, A.; Fedele, G.: »Discrete orthogonal polynomials on Gauss–Lobatto Chebyshev nodes«. *Journal of Approximation Theory* Vol. 144, Iss. 2 (2007), S. 238–246.
- [29] Ekaterinaris, J. A.: »High-order accurate, low numerical diffusion methods for aerodynamics«. *Progress in Aerospace Sciences* Vol. 41, Iss. 3-4 (2005), S. 192–300.
- [30] Epperson, J. F.: »On the Runge example«. *American Mathematical Monthly* Vol. 94 (1987), S. 329–341.
- [31] Feller, W.: *An Introduction to Probability Theory and Its Applications*. 3rd Edition. Bd. 1. John Wiley & Sons, 1968.

- [32] Förster, K.-J.: »Inequalities for ultraspherical polynomials and application to quadrature«. *Journal of Computational and Applied Mathematics* Vol. 49, Iss. 1-3 (1993), S. 59–70.
- [33] Freitag, E.; Busam, R.: *Funktionentheorie 1*. 4., korr. u. erw. Aufl. Springer-Verlag, 2006.
- [34] Gautschi, W.: *Orthogonal Polynomials: Computation and Approximation*. Numerical Mathematics and Scientific Computation. Oxford University Press, 2004.
- [35] Gautschi, W.: *Numerical Analysis*. Second Edition. Springer Science & Business Media, 2011.
- [36] Goertz, R.: »On Hahn polynomial expansion of a continuous function of bounded variation«. *PAMM* Vol. 16, Iss. 1 (2016), S. 655–656.
- [37] Goertz, R.; Öffner, P.: »On Hahn polynomial expansion of a continuous function of bounded variation«. *arXiv preprint arXiv:1610.06748* (2016).
- [38] Goertz, R.; Öffner, P.: »Spectral accuracy for the Hahn polynomials«. *arXiv preprint arXiv:1609.07291* (2016).
- [39] Gottlieb, S.; Shu, C.-W.: »Total variation diminishing Runge-Kutta schemes«. *Mathematics of Computation of the American Mathematical Society* Vol. 67, Iss. 221 (1998), S. 73–85.
- [40] Hairer, E.; Lubich, C.; Wanner, G.: *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Bd. 31. Springer Science & Business Media, 2013.
- [41] Heuser, H.: *Lehrbuch der Analysis Teil 1*. 17. Auflage. Vieweg+Teubner, 2009.
- [42] Jackson, D.: *The Theory of Approximation*. Bd. 11. American Mathematical Society, 1930.
- [43] Jerri, A. J.: *The Gibbs Phenomenon in Fourier Analysis, Splines and Wavelet Approximations*. Bd. 446. Springer Science & Business Media, 2013.
- [44] Karlin, S.; McGregor, J.: »Linear growth models with many types and multi-dimensional Hahn polynomials«. In: *Theory and Application of Special Functions*. Hrsg. von Askey, R. A. Academic Press, Inc., 1975, S. 261–288.
- [45] Karniadakis, G.; Sherwin, S.: *Spectral/hp Element Methods for Computational Fluid Dynamics*. Second Edition. Oxford University Press, 2013.
- [46] Kashin, B. S.; Saakyan, A. A.: *Orthogonal Series*. Bd. 75. Translations of Mathematical Monographs. American Mathematical Society, 1989.
- [47] Koekoek, R.; Lesky, P. A.; Swarttouw, R. F.: *Hypergeometric Orthogonal Polynomials and Their  $q$ -Analogues*. Springer Science & Business Media, 2010.
- [48] Kolmogorov, A. N.; Yushkevich, A. P.: *Mathematics of the 19th Century: Function Theory According to Chebyshev; Ordinary Differential Equations; Calculus of Variations; Theory of Finite Differences*. Bd. 3. Birkhäuser, 1998.



- [49] Koornwinder, T. H.: »Orthogonal polynomials with weight function  $(1-x)^\alpha(1+x)^\beta + M\delta(x+1) + N\delta(x-1)$ «. *Canadian Mathematical Bulletin* Vol. 27, Iss. 2 (1984), S. 205–214.
- [50] Kvernadze, G.: »Uniform convergence of Fourier-Jacobi series«. *Journal of Approximation Theory* Vol. 117, Iss. 2 (2002), S. 207–228.
- [51] Lax, P. D.: *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*. SIAM, 1973.
- [52] Lebesgue, H. L.: *Sur les intégrales singulières*. Gauthier-Villars, 1910.
- [53] Lorentz, G. G.: *Approximation of Functions*. Second Edition. American Mathematical Society, 2005.
- [54] Lorentz, G. G.: *Bernstein Polynomials*. American Mathematical Society, 2012.
- [55] Meaney, C.: »Divergent Jacobi polynomial series«. *Proceedings of the American Mathematical Society* Vol. 87, No. 3 (1983), S. 459–462.
- [56] Meinardus, G.: *Approximation of Functions: Theory and Numerical Methods*. Springer-Verlag, 1967.
- [57] Merriman, M.: »On the history of the method of least squares«. *The Analyst* Vol. 4, No. 2 (1877), S. 33–36.
- [58] Milne, W. E.: *Numerical Calculus*. Princeton University Press, 1949.
- [59] Müller, M. W.: *Approximationstheorie: Studienbuch für Studierende der Mathematik, Physik und Informatik ab 3. Semester*. Akademische Verlagsgesellschaft, 1978.
- [60] Nikiforov, A. F.; Uvarov, V. B.; Suslov, S. K.: *Classical Orthogonal Polynomials of a Discrete Variable*. Springer-Verlag, 1991.
- [61] Öffner, P. R.: »Zweidimensionale klassische und diskrete orthogonale Polynome und ihre Anwendung auf spektrale Methoden zur Lösung von hyperbolischen Erhaltungsgleichungen«. Diss. Technische Universität Braunschweig, 2015.
- [62] Petras, K.: »On the minimal norms of polynomial projections«. *Journal of Approximation Theory* Vol. 62, Iss. 2 (1990), S. 206–212.
- [63] Phillips, G. M.: *Interpolation and Approximation by Polynomials*. Bd. 14. Springer Science & Business Media, 2003.
- [64] Powierska, M.: »On the rate of convergence of some orthogonal polynomial expansions«. *Journal of Inequalities in Pure and Applied Mathematics* Vol. 8, Iss. 3 (2007).
- [65] Quarteroni, A.; Sacco, R.; Saleri, F.: *Numerische Mathematik 2*. Springer-Verlag, 2002.
- [66] Rivlin, T. J.: *The Chebyshev Polynomials*. Wiley, 1974.
- [67] Robbins, H.: »A remark on Stirling's formula«. *The American Mathematical Monthly* Vol. 62, No. 1 (1955), S. 26–29.

- [68] Runge, C.: »Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten«. *Zeitschrift für Mathematik und Physik* Vol. 46 (1901), S. 224–243.
- [69] Schatzman, M.: *Numerical Analysis: A Mathematical Introduction*. Oxford University Press, 2002.
- [70] Schönhage, A.: *Approximationstheorie*. Walter de Gruyter & Co, 1971.
- [71] Schwarz, H. R.; Köckler, N.: *Numerische Mathematik*. 8. Auflage. Vieweg+Teubner, 2011.
- [72] Slater, L. J.: *Generalized Hypergeometric Functions*. Cambridge University Press, 1966.
- [73] Szegő, G.: *Orthogonal Polynomials*. Bd. 23. American Mathematical Society, 1939.
- [74] Timan, A. F.: *Theory of Approximation of Functions of a Real Variable*. The Macmillan Company, 1963.
- [75] Wang, Z. J.: »High-order methods for the Euler and Navier–Stokes equations on unstructured grids«. *Progress in Aerospace Sciences* Vol. 43, Iss. 1-3 (2007), S. 1–41.
- [76] Weierstrass, K.: »Über die analytische Darstellbarkeit sogenannter willkürlicher Functionen einer reellen Veränderlichen«. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* (1885), S. 633–639, 789–805.
- [77] Werner, D.: *Funktionalanalysis*. 5., erweiterte Auflage. Springer-Verlag, 2004.
- [78] Wirz, M.: »Ein Spektrale-Differenzen-Verfahren mit modaler Filterung und zweidimensionaler Kantendetektierung mithilfe konjugierter Fourierreihen«. Diss. Technische Universität Braunschweig, 2012.
- [79] Xu, Y.: »On discrete orthogonal polynomials of several variables«. *Advances in Applied Mathematics* Vol. 33, Iss. 3 (2004), S. 615–632.
- [80] Xu, Y.: »Second-order difference equations and discrete orthogonal polynomials of two variables«. *International Mathematics Research Notices* Vol. 2005, Iss. 8 (2005), S. 449–475.





