

Institut für Parallele und Verteilte Systeme

Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Bachelorarbeit

Gradientenbasierte Approximation mit B-Splines auf dünnen Gittern

Felix Diez

Studiengang:	Informatik
Prüfer/in:	Jun.-Prof. Dr. rer. nat. Dirk Pflüger
Betreuer/in:	M. Sc. Julian Valentin
Beginn am:	19. Juni 2017
Beendet am:	19. Dezember 2017
CR-Nummer:	G.1.2.

Kurzfassung

Bei vielen numerischen Problemen werden für jede Auswertung einer Funktion komplexe Simulationen benötigt. Die Auswertung ist dadurch extrem zeitaufwendig. Herkömmliche Interpolationsmethoden verwenden dabei jedoch nur Funktionswerte. Verschiedene Löser ermöglichen jedoch mittlerweile auf die vollen Gradienteninformationen zuzugreifen. In dieser Arbeit werden verschiedene Ansätze dargestellt, die diese zusätzlichen Daten verwenden um den Zeitaufwand zu verringern bzw. die Interpolation zu verbessern. Im Speziellen werden, mithilfe von dünnen Gittern, Surrogate erstellt, die unter anderem Hermitebasisfunktionen oder B-Splines und diese zusätzlichen Informationen verwenden. Anschließend werden diese Verfahren mit den herkömmlichen Methoden verglichen.

Inhaltsverzeichnis

1	Einleitung	7
1.1	Problemstellung	7
1.2	Dünne Gitter	8
1.3	Gliederung	8
2	Dünne Gitter	9
2.1	Grundlagen der dünnen Gitter	9
2.2	Kombinationstechnik	12
2.3	Hermite-Basisfunktionen	14
2.4	B-Splines	15
3	Gradientenerweiterte Gitter	19
3.1	Kubische Hermiteinterpolation	19
3.2	Gradienten-erweitertes Gitter nach de Baar und Harding	20
3.3	Gradienten-erweitertes Hermite-Gitter	21
4	Evaluation	25
4.1	Methodik	25
4.2	Zweidimensionale Funktionen	25
4.3	Mehrdimensionale Funktionen	31
5	Zusammenfassung und Ausblick	41
	Literaturverzeichnis	43

1 Einleitung

In der numerischen Strömungsmechanik hat eine Auswertung oft die Berechnung einer komplexen Simulation zur Folge. Je nach Problem können jedoch auch sehr viele Auswertungen benötigt werden. Beispiele sind die Berechnung des Integrals einer Funktion oder die Suche des Minimums. Wenn eine Auswertung jedoch schon teuer ist, würde das in einer extrem langen Laufzeit resultieren.

Eine mögliche Verbesserung stellt die Ersetzung der Funktion durch ein sogenanntes Surrogat dar. Dieses erhält als Eingabe eine kleine Menge an Auswertungen der Funktion f und erstellt daraus eine Interpolation \tilde{f} . Im Allgemeinen werden dabei nur die Funktionswerte berücksichtigt. In bestimmten Fällen sind jedoch auch Gradienteninformationen vorhanden, beispielsweise bei sogenannten „adjoint solvern“. Bei diesen sind die vollen Gradienteninformationen mit nur wenig zusätzlichem Aufwand berechenbar. Bei vielen Lösern der numerischen Strömungsmechanik ist dies mittlerweile möglich.[BH15] Sollen auch diese für die Verbesserung der Interpolation mitverwendet werden, erhält man die im folgenden Abschnitt genannte Problemstellung.

1.1 Problemstellung

Gegeben sei eine Funktion $f: [0, 1]^d \rightarrow \mathbb{R}$, die auf dem Hypereinheitswürfel $[0, 1]^d$ definiert ist. Diese Funktion soll durch ein Surrogat \tilde{f} interpoliert werden. Um dieses Ziel zu erreichen, sind an bestimmten Stellen \vec{x}_i , im weiteren Verlauf Gitterpunkte genannt, die Funktionswerte $f(\vec{x})$ und die Gradienten $\nabla f(\vec{x})$ gegeben.

Für einen zweiten Ansatz sind an diesen Stellen zusätzlich noch die gemischten Ableitungen $\frac{\partial}{\partial x_{i_1} \dots \partial x_{i_m}} f(\vec{x})$ für alle paarweise verschiedenen $i_1, \dots, i_m \in \{1, \dots, d\}$ gegeben.

Die gegebenen Werte sollen nun verwendet werden, um eine möglichst gute Interpolation der Funktion f zu ermöglichen. Im Normalfall wird nur gefordert, dass an den Gitterpunkten $\tilde{f}(\vec{x}_i) = f(\vec{x}_i)$ gilt. In unserem Fall kommen noch die Gradienteninformationen hinzu. Auch diese sollen an den Gitterpunkten möglichst korrekt dargestellt werden. Daraus resultiert die zusätzliche Anforderung:

$$\forall \vec{x}_i : \nabla \tilde{f}(\vec{x}_i) = \nabla f(\vec{x}_i). \quad (1.1)$$

Werden die zur Verfügung stehenden Informationen wieder auf die gesamten gemischten Ableitungen erweitert, gilt

$$\forall x_i : \frac{\partial^{i_1 + \dots + i_m}}{\partial x_{i_1} \dots \partial x_{i_m}} \tilde{f}(\vec{x}_i) = \frac{\partial^{i_1 + \dots + i_m}}{\partial x_{i_1} \dots \partial x_{i_m}} f(\vec{x}_i) \quad (1.2)$$

für alle paarweise verschiedenen $i_1, \dots, i_m \in \{1, \dots, d\}$.

Die Intention dabei ist, die Ableitungen mitzuinterpoliern, um dadurch den globalen Fehler des Surrogats zu verringern. Um dies zu erreichen, sind passende Basisfunktionen und Ansätze gesucht, die diese mitverwenden. Das Ziel dieser Arbeit ist verschiedene Ansätze zu untersuchen, die die genannten Kriterien erfüllen, um sie anschließend mit herkömmlichen Methoden zu vergleichen.

1.2 Dünne Gitter

Eine spezielle Form der Surrogate stellen die dünnen Gitter dar. Bei normalen Gittern steigt die Anzahl der benötigten Auswertungen („Samples“) exponentiell mit der Anzahl der Dimensionen an. Dies ist auch als „Fluch der Dimensionalität“ bekannt. Um dieses Problem zu beheben, werden in dieser Arbeit dünne Gitter verwendet, die bei nur geringer Erhöhung des Fehlers viel weniger Punkte benötigen. Im Speziellen wird die Kombinationstechnik verwendet.

1.3 Gliederung

Zu Beginn der Arbeit werden in Kapitel 2 die Grundlagen der dünnen Gitter und der Kombinationstechnik erläutert, die für die weitere Arbeit benötigt werden. Anschließend werden die verwendeten Stützfunktionen vorgestellt. In Kapitel 3 werden die vorgestellten Methoden verwendet, um darauf aufbauend verschiedene Gitter zu erläutern, die den Anforderungen der Problemstellung entsprechen. In Kapitel 4 werden diese dann ausgewertet und mit herkömmlichen Methoden verglichen.

2 Dünne Gitter

2.1 Grundlagen der dünnen Gitter

Eine gute Möglichkeit Surrogate zu bilden, die den in Kapitel 1 bereits erwähnten Fluch der Dimensionalität entgegen wirken können, stellen dünne Gitter dar. Dünne Gitter wurden am Anfang hauptsächlich verwendet, um partielle Differentialgleichungen zu lösen. [Zen91] Mittlerweile wurden diese aber auch in vielen weiteren Anwendungsfeldern, unter anderem für die numerische Integration durch Ritter und Novak in [NR96], der Interpolation durch Klimke und Wohlmuth in [KW05] und auch dem Datamining in [Pfl10] und [GGT01], eingesetzt.

Die folgenden Definition basieren im Wesentlichen auf [Pfl10] und [Val14]. Detaillierte Informationen können bei [BG04] und [Pfl10] gefunden werden.

2.1.1 Eindimensionale Gitter

Wir definieren zunächst ein eindimensionales Gitter G_l auf dem Intervall $[0, 1]$ mit Level $l \in \mathbb{N}$ und Schrittweite $h_l := 2^{-l}$.

$$G_l := \begin{cases} \{ih_l | i = 0, \dots, 2^l\}, & \text{falls } l \geq 1, \\ \{\frac{1}{2}\}, & \text{falls } l = 0, \end{cases} \quad (2.1)$$

Dies unterteilt den Bereich $[0,1]$ in gleich große Stücke. Die Gitterdefinition ermöglicht jeden Punkt eindeutig durch $x_{l,i} := ih_l$ auszudrücken. Häufig wird die Hütchenbasis als Stützfunktion verwendet. Diese ist wie folgt definiert.

Definition 2.1.1 (Hütchenbasis)

$$\varphi : \mathbb{R} \rightarrow \mathbb{R} \quad \varphi(x) := \max(1 - |x|, 0) \quad (2.2)$$

Diese Standard-Hutfunktion wird skaliert, damit sie einen Träger der Breite 2^{1-l} besitzt, und um den Punkt $x_{l,i}$ zentriert. Dies führt zu der Definition:

$$\varphi_{l,i}(x) := \varphi(x/h_l - i). \quad (2.3)$$

Ordnet man jetzt jedem Knoten eines Gitters G_l eine Stelle $x_{l,i}$ zu, erhält man die

Definition 2.1.2 (nodale Basis)

$$V_l := \text{span}\{\phi_{l,i} | i = 0, \dots, 2^l\}. \quad (2.4)$$

2 Dünne Gitter

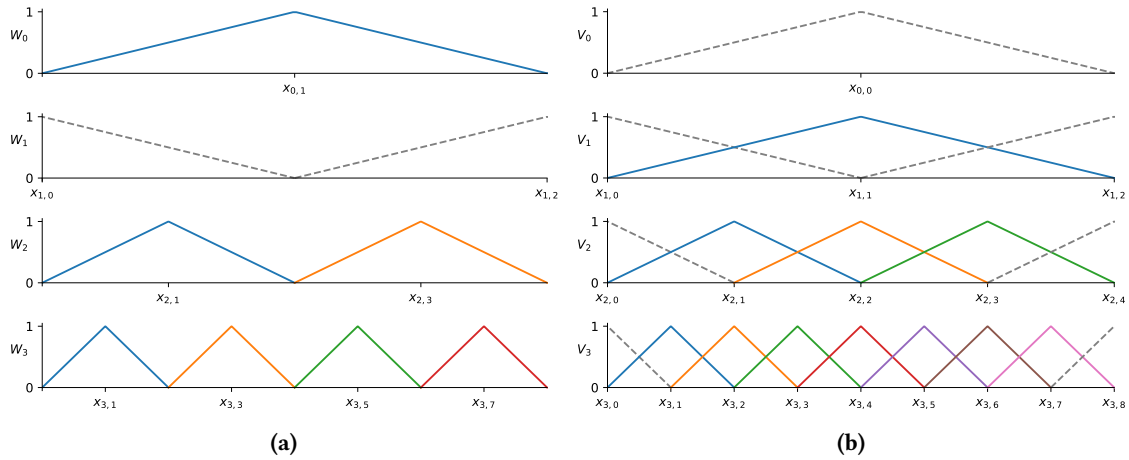


Abbildung 2.1: In (a) wird die hierarchische Basis von Level 0 bis 3 dargestellt. Die Abbildung (b) zeigt die nodale Basis. Die gestrichelten Linien sind für die Randbehandlung verantwortlich.

Betrachtet man nur jede zweite Basisfunktion erhält man die Indexmenge:

$$I_l := \begin{cases} \{i \in \mathbb{N} : 1 \leq i < 2^l, i \text{ ungerade}\}, & \text{falls } l \geq 2, \\ \{1\}, & \text{falls } l = 0, \\ \{0, 2\}, & \text{falls } l = 1, \end{cases} \quad (2.5)$$

Bildet man die Linearkombination dieser Basisfunktionen, kann der hierarchische Unterraum

$$W_l := \text{span}\{\varphi_{l,i} | i \in I_l\} \quad (2.6)$$

dargestellt werden. Eine wichtige Eigenschaft des hierarchischen Unterraums ist, dass der Schnitt zweier Träger von Basisfunktionen maximal einen Punkt enthält. Abbildung 2.1 stellt die hierarchische und nodale Basis bis $l = 3$ anschaulich dar.

Aus der Kombination von hierarchischen Untergittern kann man V_n auch wieder als direkte Summe ausdrücken:

$$V_{l_{max}} = \bigoplus_{l=0}^{l_{max}} W_l. \quad (2.7)$$

Daher lassen sich alle Funktionen $u \in V_{l_{max}}$ als Linearkombination

$$u(x) = \sum_{x_{l,i} \in G_l} \alpha_{l,i} \varphi_{l,i}(x) \quad (2.8)$$

darstellen. Die Koeffizienten $\alpha_{l,i}$ werden dabei „hierarchische Überschüsse“ genannt. Da wir jedoch in dieser Arbeit die Kombinationstechnik verwenden werden, ist die genaue Berechnung nicht relevant.

2.1.2 Mehrdimensionale dünne Gitter

Mithilfe eines Tensorprodukt-Ansatzes kann das eindimensionale Gitter auf mehrere Dimensionen verallgemeinert werden. Der Definitionsbereich wird dabei von $[0, 1]$ auf den Hyperwürfel $[0, 1]^d$ mit d Dimensionen erweitert. Dabei bezeichnet $\vec{l} = (l_1, \dots, l_d)$ das Level, wobei l_j für das jeweilige Level in Dimension j steht. Für den Index $\vec{i} = (i_1, \dots, i_d)$ gilt das Gleiche. Jede Dimension kann so ihr eigenes Level besitzen. Wird weiter die Hütchenbasis verwendet, resultiert das in folgender Definition:

Definition 2.1.3 (d -dimensionale Hütchenbasis)

$$\varphi_{\vec{l}, \vec{i}}(\vec{x}) := \prod_{j=1}^d \varphi_{l_j, i_j}(x_j) \quad (2.9)$$

Weiterhin werden auch die anderen Eigenschaften des 1D-Falls übernommen. So kann auch im mehrdimensionalen Fall wieder die Basis des Raums aller stückweise d -linearen und stetigen Funktionen mit den Hütchenfunktionen $\varphi_{l,i}$ gebildet werden:

$$V_{\vec{l}} := \text{span}\{\varphi_{\vec{l}, \vec{i}} | \forall t=1, \dots, d \quad i_t = 0, \dots, 2^{l_t}\}. \quad (2.10)$$

Das Gleiche gilt für die hierarchischen Unterräume:

$$W_{\vec{l}} := \text{span}\{\varphi_{\vec{l}, \vec{i}} | \vec{i} \in I_{\vec{l}}\} \quad \text{mit} \quad I_{\vec{l}} := I_{l_1} \times \dots \times I_{l_d} \quad (2.11)$$

Je nach Wahl der Unterräume entstehen dadurch verschiedene Gitterarten. Verwendet man für ein Maximal-Level l_{max} alle Unterräume $W_{\vec{l}}$, für die gilt $\forall j : l_j \leq l_{max}$, erhält man ein volles Gitter:

$$V_{l_{max}} = \bigoplus_{|\vec{l}|_{\infty} \leq l_{max}} W_{\vec{l}} \quad (2.12)$$

Bei dünnen Gittern werden nur Unterräume gewählt, die besonders viel zum Ergebnis beitragen. Die Beitragsstärksten können dabei über ein kontinuierliches Rucksackproblem hergeleitet werden.[BG04] Bezüglich des L^2 -Fehlers¹ ergibt sich daraus

Definition 2.1.4 (Reguläres dünnes Gitter)

$$V_{l_{max}}^{(s)} := \bigoplus_{|\vec{l}|_1 \leq n+d-1} W_{\vec{l}} \quad (2.13)$$

¹Der L^2 -Fehler ist gegeben durch: $|u - f|_{L^2}$

Analog zu 2.8 können alle Funktionen $v \in V_{l_{max}}^s$ dargestellt werden als

$$\sum_{|\vec{l}|_1 \leq n+d-1} \sum_{\vec{i} \in I_{\vec{l}}} \alpha_{\vec{l}, \vec{i}} \varphi_{\vec{l}, \vec{i}}(\vec{x}). \quad (2.14)$$

Durch die Verringerung der Unterräume kann die Anzahl der Gitterpunkte von $\mathcal{O}(2^{dl_{max}})$ auf $\mathcal{O}(2^{l_{max}}) \cdot l_{max}^{d-1}$ reduziert werden. Der L^2 -Fehler wird dabei jedoch nur um den Faktor l_{max}^{d-1} schlechter als beim vollen Gitter und erhöht sich von $\mathcal{O}2^{-2l_{max}}$ auf $\mathcal{O}(2^{-2l_{max}} \cdot l_{max}^{d-1})$. [BG04]

2.1.3 Randbehandlung

Besitzt die zu interpolierende Funktion am Rand Werte ungleich 0, dann kann dies mit der Hütchenbasis ohne Randbehandlung nicht richtig interpoliert werden. Um dem entgegenzuwirken, werden wir wie in den Definitionen und Abbildung 2.1 dünne Gitter verwenden, die Punkte auf dem Rand besitzen. Alternativ können auch die Basisfunktionen modifiziert werden, um die Randwerte zu extrapolieren. [Pfl10]

2.2 Kombinationstechnik

Eine praktische Methode um dünne Gitter zu konstruieren stellt die Kombinationstechnik dar. Hier wird ausgenutzt, dass ein dünnes Gitter aus der linearen Kombination von gröberen, anisotropischen Vollgittern erstellt werden kann:

$$V_{\vec{l}} := \bigoplus_{\vec{k} \leq \vec{l}} \tilde{W}_{\vec{k}}. \quad (2.15)$$

Detailliertere Einführungen sind bei [GSZ92] oder [HGC07] zu finden. In Abbildung 2.2 wird die Technik verdeutlicht durch die Konstruktion eines dünnen Gitters mit Level 2. Die roten Gitter mit $|\vec{l}|_1 = 1$ werden subtrahiert und die grünen Gitter mit $|\vec{l}|_1 = 2$ werden addiert. Das Resultat wird in Abbildung 2.3 gezeigt.

Das Resultat des gesamten Surrogats lässt sich berechnen durch:

$$u(\vec{x}) := \sum_{k=0}^{d-1} (-1)^k \binom{d-1}{k} \sum_{|\vec{l}|_1 = n+(d-1)-k} u_{\vec{l}}(\vec{x}). \quad (2.16)$$

Dabei ist 2.15 der volle Gitterraum von einem Level \vec{l} und $\vec{u}_{\vec{l}} \in V_{\vec{l}}$ der Interpolant einer Funktion f . [Pfl10]

Da bei der Technik nur die nodale Basis verwendet wird, ist die Implementierung einfacher als Implementierungen mit der hierarchischen Basis. So fällt unter anderem die Berechnung der hierarchischen Überschüsse weg. Auch bei den Basisfunktionen, die wir verwenden werden, stellt dies eine Vereinfachung dar. Zusätzlich kann die Berechnung durch die Kombination der Gitter leicht parallelisiert werden, weil jedes Gitter unabhängig von den Anderen berechnet werden kann. Ein weiterer Vorteil der Technik ist, dass diese minimalinvasiv ist und wenig an bereits vorhandenen Strukturen von Programmen oder Lösern geändert werden muss. [HGC07]

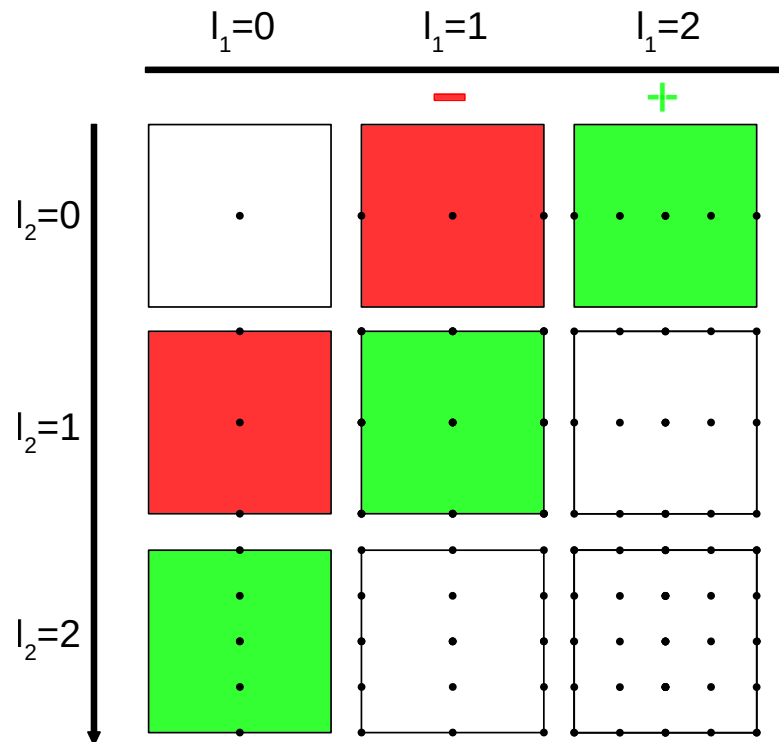


Abbildung 2.2: Anwendung der Kombinationstechnik um ein zweidimensionales Gitter mit Level 2 zu erhalten. Die zu addierenden Gitter sind Grün gefärbt und die zu subtrahierenden Rot.

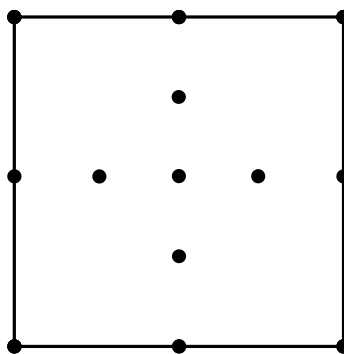
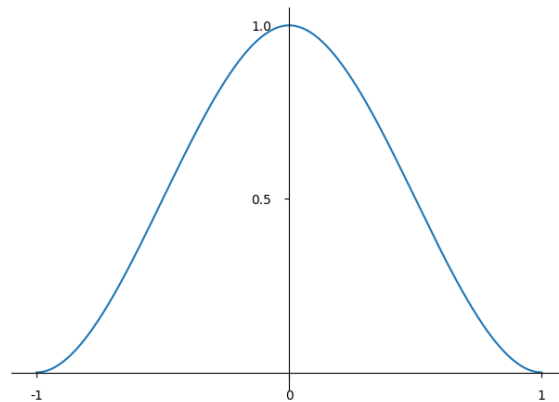


Abbildung 2.3: resultierendes dünnes Gitter bei Addition der grünen und Subtraktion der roten Gitter in Abbildung 2.2 für $l = 2$

Abbildung 2.4: ψ -Funktion

2.3 Hermite-Basisfunktionen

Ein Nachteil der Hütchenfunktionen ist, dass sie nicht stetig differenzierbar sind. Da unser Ziel auch die Interpolation der Ableitung ist, sind andere Basisfunktionen nötig. Die Basisfunktionen, die wir verwenden werden, entspringen der kubischen Hermiteinterpolation. Der Vorteil von diesen ist, dass die Funktionen den gleichen Träger wie die Hütchenfunktion haben.

2.3.1 ψ -Funktion

Die erste Funktion hat den Zweck, die Funktionswerte an den Gitterpunkten richtig darzustellen ohne die Ableitung zu verändern. Daraus folgen die Bedingungen $\psi(0) = 1$ und $\psi'(0) = 0$. Weiterhin soll der Träger gleich sein wie bei Hütchenfunktionen und eine stetige Differenzierbarkeit möglich sein. Daraus folgen weitere Bedingungen am Rand. Dort sollen die Ableitung und der Funktionswert jeweils 0 betragen. Sollen alle Anforderungen erfüllt werden, ist mindestens ein stückweises Polynom von Grad 3 erforderlich. Das einzige mit diesem Grad ist gegeben durch:

Definition 2.3.1 (ψ -Funktion)

$$\psi(x) = \begin{cases} -2x^3 + 3x^2 + 1, & x < 0, \\ 2x^3 + 3x^2 + 1, & x \geq 0, \end{cases} \quad (2.17)$$

Die resultierende Funktion ist in Abbildung 2.4 dargestellt. Analog zur Hütchenbasis kann die Zentrierung um einen Gridpunkt erfolgen:

$$\psi_{l,i}(x) := \psi(x/h_l - i). \quad (2.18)$$

2.3.2 ζ -Funktion

Das Ziel der zweiten Basisfunktion ist die korrekte Darstellung der Ableitung an den Gitterpunkten, ohne dabei jedoch Einfluss auf den Funktionswert zu nehmen. Daraus ergeben sich die Forderungen nach $\zeta(0) = 0$ und $\zeta'(0) = 1$. Genau wie bei ψ soll auch diese den gleichen Träger stetig differenzierbar sein und ähnlich der Hütchenfunktion eingesetzt werden können. Durch die Anzahl der Bedingungen gibt es wieder ein eindeutiges stückweises Polynom vom Grad 3:

Definition 2.3.2 (ζ -Funktion)

$$\zeta(x) = \begin{cases} x^3 + 2x^2 + x, & x < 0, \\ x^3 - 2x^2 + x, & x \geq 0, \end{cases} \quad (2.19)$$

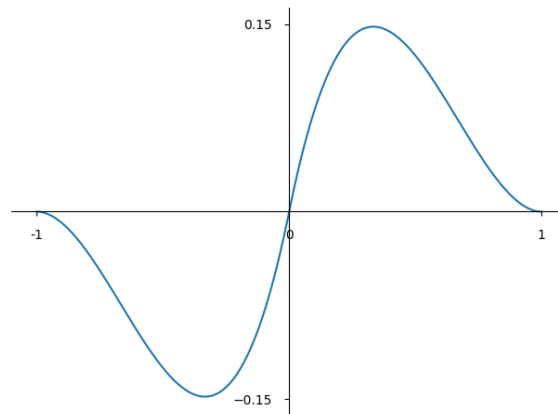


Abbildung 2.5: ζ -Funktion

Analog zu ψ und der Hütchenbasis kann die Skalierung und Verschiebung auf einen Gitterpunkt erfolgen.

2.4 B-Splines

Eine für unsere Zwecke interessante Verallgemeinerung der Hütchenfunktion stellen B-Splines dar. Ein Vorteil dieser ist, dass sie $p - 1$ -mal stetig differenzierbar sind. Weitere Eigenschaften sind die freie Wählbarkeit des Grades p und die Begrenztheit auf $p + 1$ lokale polynomielle Abschnitte.[Val14] Die bereits erwähnte Hütchenfunktion ist durch B-Splines mit Grad $p = 1$ gegeben. Bei diesem Grad ist jedoch keine stetige Differenzierung möglich. Deshalb werden wir auch höhere Grade evaluieren.

Mithilfe der Cox-de-Boor-Rekursion können B-Splines mit Grad p hergeleitet werden. [De 72]

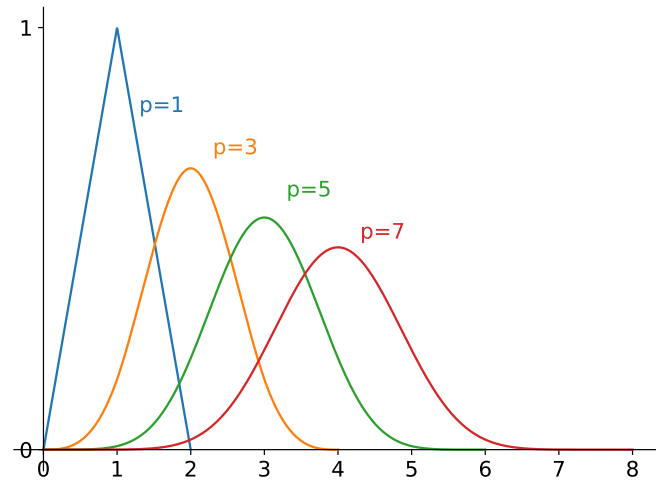


Abbildung 2.6: B-Splines mit Grad $p = 1, 3, 5, 7$

Definition 2.4.1 (uniformer B-Spline)

$$N^0(x) := \begin{cases} 1, & \text{falls } 0 \leq x \leq 1, \\ 0, & \text{sonst,} \end{cases} \quad (2.20)$$

$$N^p(x) := \left\{ \frac{x}{p} N^{p-1}(x) + \frac{p+1-x}{p} N^{p-1}(x-1), \right. \quad (2.21)$$

Der B-Spline eines Grades p hat den Träger $[0, p + 1]$. Der maximale Wert von 1 wird im Punkt $\frac{p+1}{2}$ erreicht. In Abbildung 2.6 werden die ungeraden B-Splines bis Grad 7 anschaulich dargestellt. Für die Anwendung auf dem Gitter wird, wie bei Grad 1, die gewählte B-Spline- Funktion durch Strecken und Verschieben auf die Gridpunkte $x_{l,i}$ zentriert. Hierdurch folgt

Definition 2.4.2 (uniforme B-Spline-Basisfunktion mit Level l und Index $i \in I_l$)

$$\varphi_{l,i}^p(x) := b^p \left(\frac{x}{h_l} + \frac{p+1}{2} - i \right) \quad (2.22)$$

Dies bedeutet, dass das Maximum bei $x_{l,i}$ liegt und der Träger $h_l \cdot [i - (p + 1)/2, i + (p + 1)/2]$ ist. Daraus folgt, dass für ungerade Grade p die B-Spline Knoten für $\varphi_{l,i}^p(x)$ genau auf den Gitterpunkten

$$x_{l,i-(p+1)/2}, \dots, x_{l,i}, \dots, x_{l,i+(p+1)/2} \quad (2.23)$$

liegen. Für gerade p liegen die Knoten genau zwischen zwei Gitterpunkten. Deshalb werden der Einfachheit halber nur noch B-Splines mit ungeradem Grad betrachtet. Durch den breiteren Träger ist die Bestimmung der richtigen Koeffizienten schwerer als bei den zuvor vorgestellten Basisfunktionen. Das hat zur Folge, dass ein Gleichungssystem aufgestellt und gelöst werden muss.[Val14]

Für die Verwendung von B-Splines in mehreren Dimensionen kann analog zu den bereits genannten Funktionen der Tensorprodukt-Ansatz verwendet werden.

2.4.1 Not-a-Knot-B-Splines

Bei den normalen B-Splines kann es selbst bei relativ einfach zu interpolierenden Funktionen zu unnatürlichen Überschüssen am Rand kommen. [VP16] Um diesem Problem vorzubeugen, werden wir speziell angepasste B-Splines, "Not-a-Knot-B-Splines"(NAK-B-Splines) verwenden. Bei diesen werden in jedem Level die B-Splines mit dem kleinsten und größten Index verändert und erfüllen dadurch die sogenannten NAK-Bedingungen. Detailliertere Informationen sind in [HH15] zu finden.

3 Gradientenerweiterte Gitter

3.1 Kubische Hermiteinterpolation

Da die weiteren Ansätze alle auf den Hermitebasisfunktionen ψ und ζ basieren, wird nun zuerst die standardmäßige Hermite-Interpolation mit einer Dimension erläutert, um anschließend verschiedene Ansätze zur Erweiterung auf mehrere Dimensionen durchzuführen.

In einer Dimension kann aus den zwei vorgestellten Basisfunktionen die kubische Hermiteinterpolation gebildet werden. Hierbei verwenden wir die in Kapitel 2 definierte Gitterstruktur mit den Gitterpunkten $x_{l,i}$. Jedoch wird nicht die Hütchenbasis als Stützfunktion verwendet. Anstelle dieser wird jeweils ein Gitter mit ψ und eines mit ζ als Stützfunktion gebildet. Im weiteren Verlauf bezeichnet G^ψ das Gitter mit ψ und das andere wird analog dazu G^ζ genannt. Hierbei wird $G^\psi(f)$ verwendet um die Funktionswerte an den Stützstellen richtig darzustellen. Die zu interpolierende Funktion wird in den Klammern genannt, in diesem Fall f . G^ζ wiederum erhält als zu interpolierende Funktion die Ableitung von f und erzeugt dadurch richtige Ableitungswerte an den Gitterpunkten. Durch Addition erhält man die stückweise kubische Hermiteinterpolation:

$$u(x) = \sum_k \left(f(x_k) \psi_k(x) + f'(x_k) \zeta_k(x) \right). \quad (3.1)$$

In Abbildung 3.1 wird die Kombination der zwei Gitter anhand der Interpolation einer Parabel verdeutlicht. Wie in Abbildung 3.1 zu sehen ist, stellt das G_ψ die Funktionswerte an den Gitterpunkten bereits richtig dar. Jedoch ist die Ableitung 0, wodurch Hochpunkte entstehen. Das zusätzliche G_ζ hat an diesen Stellen die richtige Steigung, jedoch mit den Funktionswerten 0. Durch die Kombination entsteht dann die in Abbildung 3.2 gezeigte, gewünschte Parabelform.

Das Prinzip mit den zwei Gittern kann jedoch nicht einfach so auf den mehrdimensionalen Fall übertragen werden. Bei d Dimensionen kann die Funktion nach d verschiedenen Dimensionen abgeleitet werden, beispielsweise ergibt dies für $d = 2$ eine Ableitung nach x_1 und x_2 . Um die Gradienteninformationen sinnvoll zu verwenden, sind deshalb andere Methoden nötig. Einen Ansatz für die Erweiterung auf mehrere Dimensionen wird von de Baar und Harding beschrieben.[BH15] Im Folgenden wird dieser näher erläutert.

3 Gradientenerweiterte Gitter

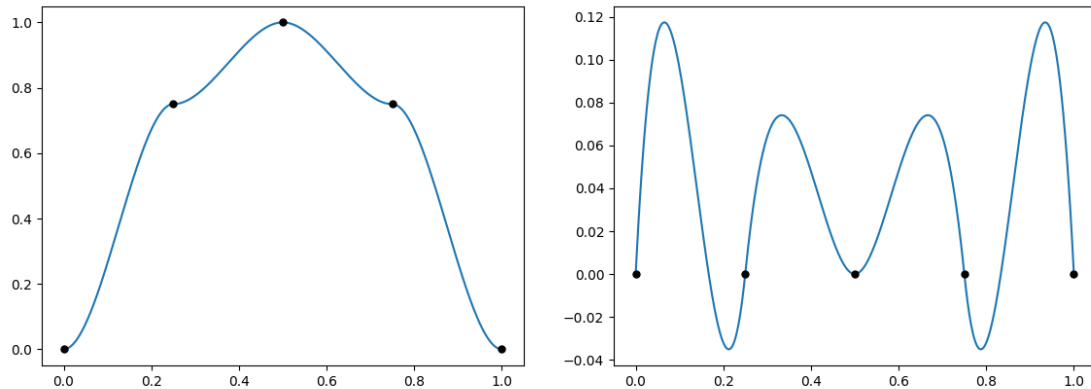


Abbildung 3.1: Teilgitter der stückweisen kubischen Hermiteinterpolation der Funktion $f(x) := -4((x - 0.5)^2) + 1$ mit Level $l = 2$. Auf der linken Seite ist G^ψ und rechts G^ζ abgebildet. Zu beachten ist die unterschiedliche Achsenskalierung.

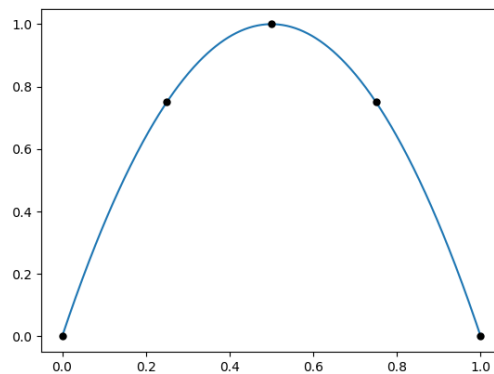


Abbildung 3.2: Resultierende Interpolation aus der Kombination der Gitter in Abbildung 3.1

3.2 Gradienten-erweitertes Gitter nach de Baar und Harding

3.2.1 Gitter nach de Baar und Harding mit Hütchenfunktionen

Das „Gradient-enhanced Grid“ nach de Baar und Harding stellt eine Möglichkeit dar, die zusätzlichen Informationen der Gradienten zu verwenden. Die Idee des Gitters ist weiterhin das Kombinieren von verschiedenen Gittern. Im Speziellen werden für jede Dimension i jeweils zwei Gitter erstellt. Wird nur diese allein betrachtet ist der Ansatz ähnlich zum eindimensionalen Fall. In diese Dimension wird jeweils ein Gitter mit ψ und ein zweites mit ζ als Stützfunktion verwendet. Für die restlichen $d - 1$ Dimensionen wird bei beiden die lineare Basis benutzt. Als Kurzform werden diese im weiteren Verlauf als $G_{\psi,i}^{BH}$ bzw. $G_{\zeta,i}^{BH}$ bezeichnet, wobei i die erwähnte spezielle Dimension darstellt. Falls eine bestimmte Funktion interpoliert werden soll, wird diese in Klammern dahinter angegeben, wie in Abschnitt 3.1. Alle verwendeten Gitter basieren auf der Definition aus Kapitel 2 und die dort definierte Kombinationstechnik. Diese Konstruktion hat zur Folge, dass für Richtung i der Funktionswert und die Ableitung in Richtung i an den

Gitterpunkten durch die Addition der zwei Gitter korrekt interpoliert wird. Werden nun jedoch alle $2d$ Gitter addiert, dann ist der Funktionswert um den Faktor $d - 1$ zu hoch. Auch bei der Ableitung kommt es durch die Verwendung der linearen Stützfunktionen zu Fehlern. Um diese möglichst gut „auszugleichen“, wird nun noch $d - 1$ -mal das normale stückweise lineare Gitter G^{linear} subtrahiert. Formal folgt daraus die Definition:

Definition 3.2.1 (mehrdimensionales Gitter nach de Baar und Harding)

$$\tilde{f}(\vec{x}) := -(d - 1) \cdot G^{linear}(f) + \sum_i G_{\psi,i}^{BH}(f) + \sum_i G_{\zeta,i}^{BH} \left(\frac{\partial f}{\partial x_i} \right) \quad (3.2)$$

Insgesamt werden dadurch $2d + 1$ Gitter benötigt. Zusätzlich zu erwähnen ist, dass durch die linearen Stützfunktionen keine stetige Differenzierbarkeit gegeben ist. Ebenfalls ist dies keine richtige Hermiteinterpolation, da hierfür noch die gemischten Ableitungen nötig wären.

Um das Konstrukt zu verdeutlichen, betrachten wir nun den zweidimensionalen Fall. Hier werden fünf verschiedene Gitter benötigt. In Abbildung 3.3 sind diese für das Level $l = 2$ dargestellt. Diese werden mithilfe der Kombinationstechnik erstellt und dann nochmals, mithilfe der Definition 3.2.1, kombiniert. In diesem Fall folgt daraus die Addition aller Gitter bis auf G^{linear} , welches einmal subtrahiert wird.

3.2.2 Gitter nach de Baar und Harding mit B-Splines

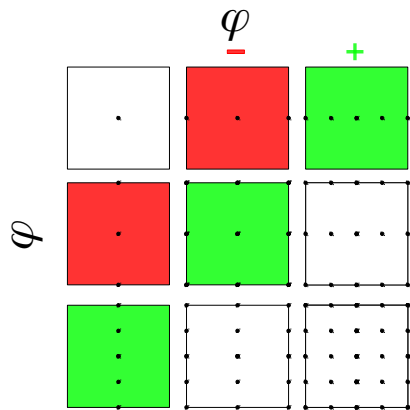
Wie bereits erwähnt, ist das Gitter nach de Baar und Harding nicht stetig differenzierbar. Ist die Berechnung der Ableitung bzw. die stetige Differenzierbarkeit aber auch ein gewünschter Aspekt, dann stellt dies ein Hindernis dar. In Abschnitt 2.4 wurden Basisfunktionen vorgestellt, die ab Grad 2 mindestens einmal stetig differenzierbar sind. Um die Interpolation der Ableitung gleich zu belassen, bietet es sich an nur die linearen Stützfunktionen zu ersetzen. Im Grunde stellt dies nur eine Verallgemeinerung von Grad 1 auf einen beliebigen dar. Hierfür ändert sich die Gitterdefinition. Das reine B-Spline-Gitter wird jetzt dargestellt durch $G^{bspl(p)}$, wobei p den Grad angibt. Bei den anderen Gittern wird ebenfalls jeweils die linearen Stützfunktionen durch B-Splines vom Grad p ersetzt.

Definition 3.2.2

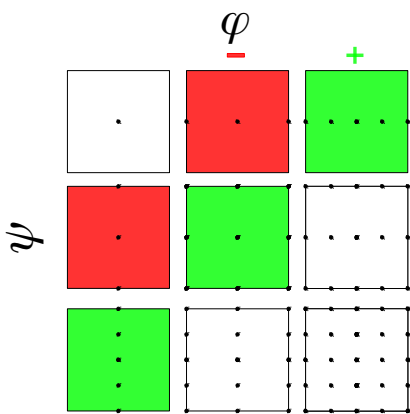
$$u(\vec{x}) := -(d - 1) \cdot G^{bspl(p)}(f) + \sum_i G_{\psi,i}^{BH,bspl(p)}(f) + \sum_i G_{\zeta,i}^{BH,bspl(p)} \left(\frac{\partial f}{\partial x_i} \right) \quad (3.3)$$

3.3 Gradienten-erweitertes Hermite-Gitter

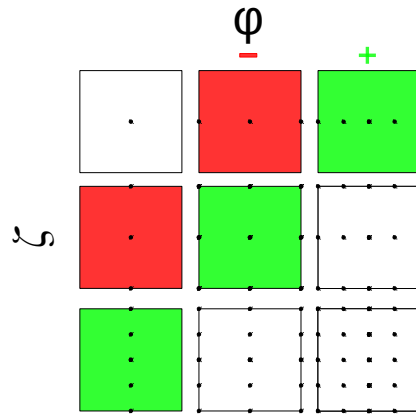
Eine andere Möglichkeit Ableitungsinformationen mitzuverwenden stellt die „richtige“ mehrdimensionale kubische Hermiteinterpolation dar. Hierfür werden auch keine linearen Stützfunktionen benötigt, wodurch die stetige Differenzierbarkeit möglich ist. Im Gegensatz zum Gitter nach de Baar und Harding sind aber zusätzliche Informationen, die gemischten Ableitungen, notwendig. Die Idee dabei ist ein Gitter G_{ψ} zu erstellen, das die Funktionswerte korrekt darstellt,



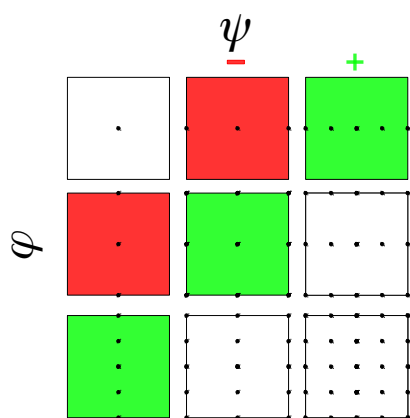
(a) G^φ



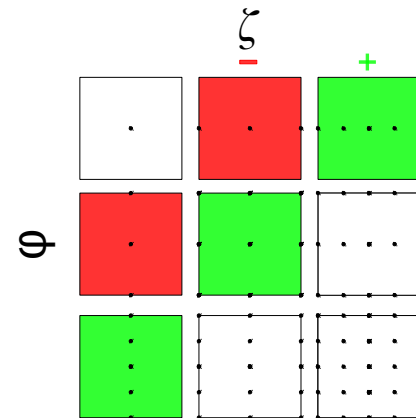
(b) $G_{\psi,1}^{BH,\varphi}$



(c) $G_{\zeta,1}^{BH,\varphi}$



(d) $G_{\psi,2}^{BH,\varphi}$



(e) $G_{\zeta,2}^{BH,\varphi}$

Abbildung 3.3: Komponenten des Gitters nach de Baar und Harding mit Level 2. Die Dimensionen mit φ stehen dabei für die Hütchenfunktionen oder B-Splines.

dabei jedoch die Ableitungswerte an den Gitterpunkten nicht beeinflusst. Zusätzlich wird für jede gegebene partielle Ableitung ein weiteres Gitter erzeugt.

Dies funktioniert, da das Gitter G_{ψ}^{He} nur die Funktionswerte interpoliert, die Ableitungen jedoch 0 sind. Um alle partiellen Ableitungen richtig darzustellen, kann, wie beim vorigen Gitter, ζ verwendet werden. Allerdings kombinieren wir dieses nicht mit den linearen Stützfunktionen, sondern ψ . Dadurch ist die Interpolation weiterhin stetig differenzierbar und die Ableitungen an den Gridpunkten in andere Dimensionen werden nicht beeinflusst bzw. sind genau 0. Für partielle Ableitungen in mehr als eine Dimension wird einfach in alle betroffenen Dimensionen ζ als Stützfunktion verwendet. Die gemischten Ableitungen können auch weggelassen werden, wodurch diese einfach 0 wären und vorraussichtlich keine guten Ergebnisse erzielt werden.

Die so erstellten Gitter werden im Folgenden $G_{\zeta, I}^{He}$ genannt. I stellt dabei alle Dimensionen dar, in die abgeleitet wird und dementsprechend ζ verwendet wird. Daraus ergibt sich die Definition:

Definition 3.3.1 (mehrdimensionale Gradienten-erweitete Gitter Hermite)

$$\tilde{f}(\vec{x}) := G_{\psi}^{He}(f) + \sum_{I \in \mathcal{P}(1\dots d) \setminus \{\emptyset\}} G_{\zeta, I}^{He}\left(\frac{\partial f}{\partial x_I}\right). \quad (3.4)$$

An der Definition lässt sich bereits ablesen, dass für höhere Dimensionen viele zusätzliche Gitter und Informationen benötigt werden. Im Vergleich benötigt dieses 2^d Gitter und der Ansatz nach de Baar und Harding nur $2d + 1$. Schon bei 6 Dimensionen würde man so 64 Gitter benötigen, im Gegensatz zu 13 beim Anderen. Das führt erneut zum sogenannten Fluch der Dimensionalität, der durch die dünnen Gitter eigentlich verhindert werden soll.

In zwei Dimensionen gibt es dementsprechend ein Gitter G_{ψ} , zwei Gitter für die Ableitungen x_1 , x_2 und ein weiteres für die gemischte Ableitung nach x_1 und x_2 . Durch die Addition von allen entsteht die resultierende Interpolation $u(\vec{x})$. In Abbildung 3.4 wird die Zusammensetzung für $l = 2$ verdeutlicht.

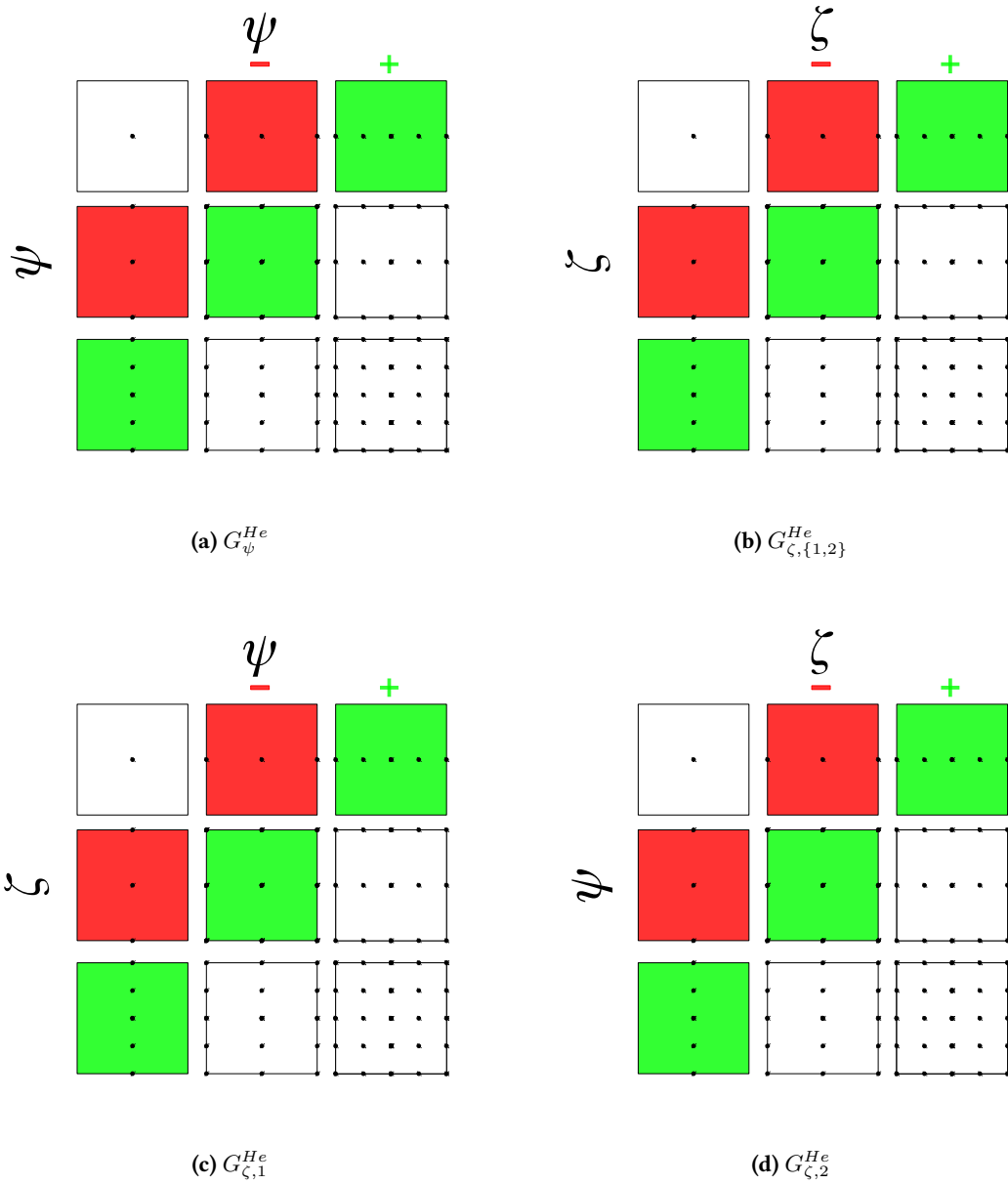


Abbildung 3.4: Zusammensetzung des Hermite-Gitters mit Level 2

4 Evaluation

4.1 Methodik

Bevor die Evaluation beginnt, werden zuerst die Notationen und Eigenheiten der Implementierung erläutert. Wenn Funktionen dargestellt werden, ist immer ein Definitionsbereich mit angegeben. Dieser wird auf $[0, 1]^d$ skaliert, um unsere Definition der dünnen Gitter zu verwenden. Weitergehend werden L^2 -Fehlerplots dargestellt. Dabei wird der Fehler mit mit $n = 10000$ uniform verteilten Punkten x_i durch $\sqrt{\frac{1}{n} \sum_{i=1}^n (u(\vec{x}_i) - f(\vec{x}_i))^2}$ bestimmt.

Die verwendeten Gitter basieren alle auf den in Kapitel 2 definierten dünnen Gittern mit Punkten auf dem Rand und verwenden die Kombinationstechnik. In diesem Zusammenhang stellt eine Besonderheit die Implementierung der Not-A-Knot-B-Splines dar. Diese werden erst ab Level 3 verwendet, unabhängig davon ob der Grad 3 oder 5 beträgt. Stattdessen werden Lagrange-Polynome verwendet, wodurch Polynome bis Grad 8 dargestellt werden können.

Eine weitere Einschränkung bei der Implementierung stellt dar, dass bei den Gittern mit Hütchen keine stetige Differenzierbarkeit gegeben ist. Deswegen werden bei allen Ableitungen der Gitter Finite Differenzen verwendet, die einen zusätzlichen Fehler enthalten.

Funktionen, die einen Betrag enthalten, wurden nicht berücksichtigt, da bei diesen keine stetige Differenzierbarkeit gegeben ist. Dies ist aber in unserem Fall gewünscht bzw. erforderlich, um die Gitter zu erzeugen.

4.2 Zweidimensionale Funktionen

Die erste Funktion ist eine zweidimensionale Parabel, gegeben durch:

$$f(\vec{x}) := 4x_1(1 - x_1)4x_2(1 - x_2). \quad (4.1)$$

Besonderes Augenmerk liegt hierbei auf dem visuellen Vergleich der Gittertypen. In der Abbildung 4.1 fällt direkt auf, dass das Gitter nach de Baar und Harding viel glatter ist als das andere gezeigte mit stückweisen linearen Basisfunktionen. Dies ist überraschend, da G^{BH} weiterhin zum Teil ebendiese Hütchen verwendet. Bei den anderen beiden Gittern ist die Glattheit weniger überraschend, da nur stetig differenzierbare Polynome mit Grad ≥ 2 verwendet werden. Damit sind glattere Funktionen möglich, weswegen diese hier nicht extra gezeigt werden. Betrachten wir nun nur ein spezielles Teilgitter, zum Beispiel das in Abbildung 4.2 dargestellte jeweilige Gitter, das die Ableitung nach x_1 zur Interpolation erhält, sind deutliche Unterschiede zwischen den Gittertypen sichtbar. Hier ist die Verwendung der Hütchenfunktionen wieder klar an den Spitzen zu sehen. Das Hermite-Gitter und die BH-Variante mit B-Splines sind auch hier wenig überraschend glatter.

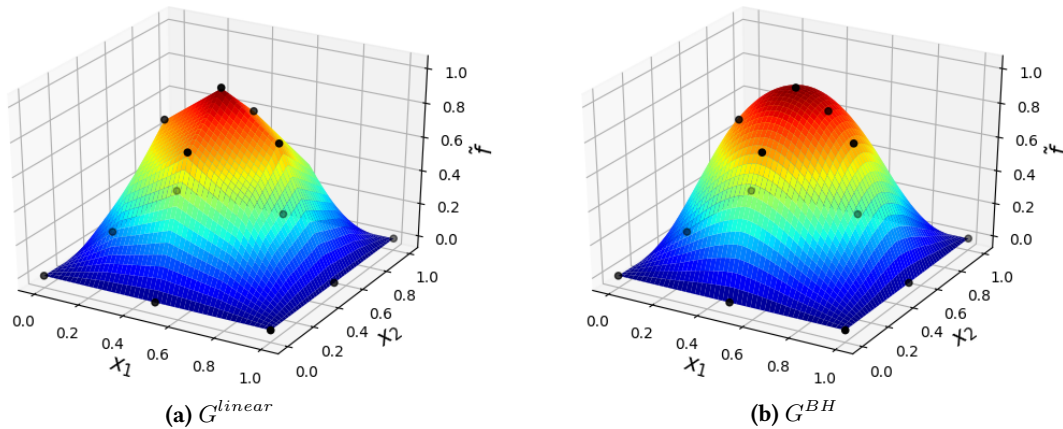


Abbildung 4.1: Beide Abbildungen sind Interpolationen mit $l = 3$ von 4.1. Bei (a) wird ein dünnes Gitter mit linearen Stützfunktionen verwendet, bei (b) das Gitter nach de Baar und Harding

Himmelblau-Funktion

Als nächste Testfunktion, die wir genauer betrachten wollen, betrachten wir die Himmelblau-Funktion:

$$f(\vec{x}) := (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2, \quad \vec{x} \in [-5, 5]^2. \quad (4.2)$$

Diese ist ein Polynom vom Grad 4 und enthält keine trigonometrischen Funktionen. Ausgewählt wurde sie stellvertretend für Funktionen, die mit B-Splines oder einer Polynombasis gut zu interpolieren sind. Im Speziellen können Not-a-Knot-B-Splines vom Grad p alle Polynome vom Grad $\leq p$ korrekt abbilden. Abbildung 4.3 zeigt die visuelle Form.

In Abbildung 4.4 wird der zugehörige L^2 -Fehler der Interpolation mit den verschiedenen Gittertypen dargestellt. Vergleichen wir nun zuerst das lineare Gitter mit der Variante nach de Baar und Harding. Hier fällt auf, dass die Konvergenzordnung nahezu identisch ist, jedoch das Gitter nach de Baar und Harding immer um einen konstanten Faktor besser ist. Mögliche Gründe könnten einerseits die zusätzlich verwendeten Gradienteninformationen, aber auch die erhöhte Glattheit, beispielsweise zu sehen in Abbildung 4.1, durch die Verwendung der Hermite-Basisfunktionen sein.

Bei den restlichen Gittern gibt es noch zwei, die besonders auffallen. Das Erste ist das Hermite-Gitter ohne die gemischten Ableitungen, welches am schlechtesten abschneidet. Das Zweite sind die B-Splines vom Grad 5, die bereits ab Level 2 die Funktion richtig darstellen, unter Berücksichtigung der Maschinengenauigkeit. Dies entspricht den bereits erwähnten Erwartungen. Dass Level 2 der B-Splines vom Grad 3 ebenfalls die Funktion korrekt interpoliert, ist auf die spezifische Implementierung zurückzuführen.

Nimmt man dieses Level aus, sind die übrigen Gitter alle im gleichen Bereich bzw. konvergieren ungefähr gleich schnell. Ein Grund hierfür könnte sein, dass alle mindestens zu großen Anteilen Polynome vom Grad 3 verwenden. Obwohl dem Hermite-Gitter mehr Informationen zur Verfügung stehen, schneidet es sogar etwas schlechter ab als das reine B-Spline-Gitter mit Grad 3. Die

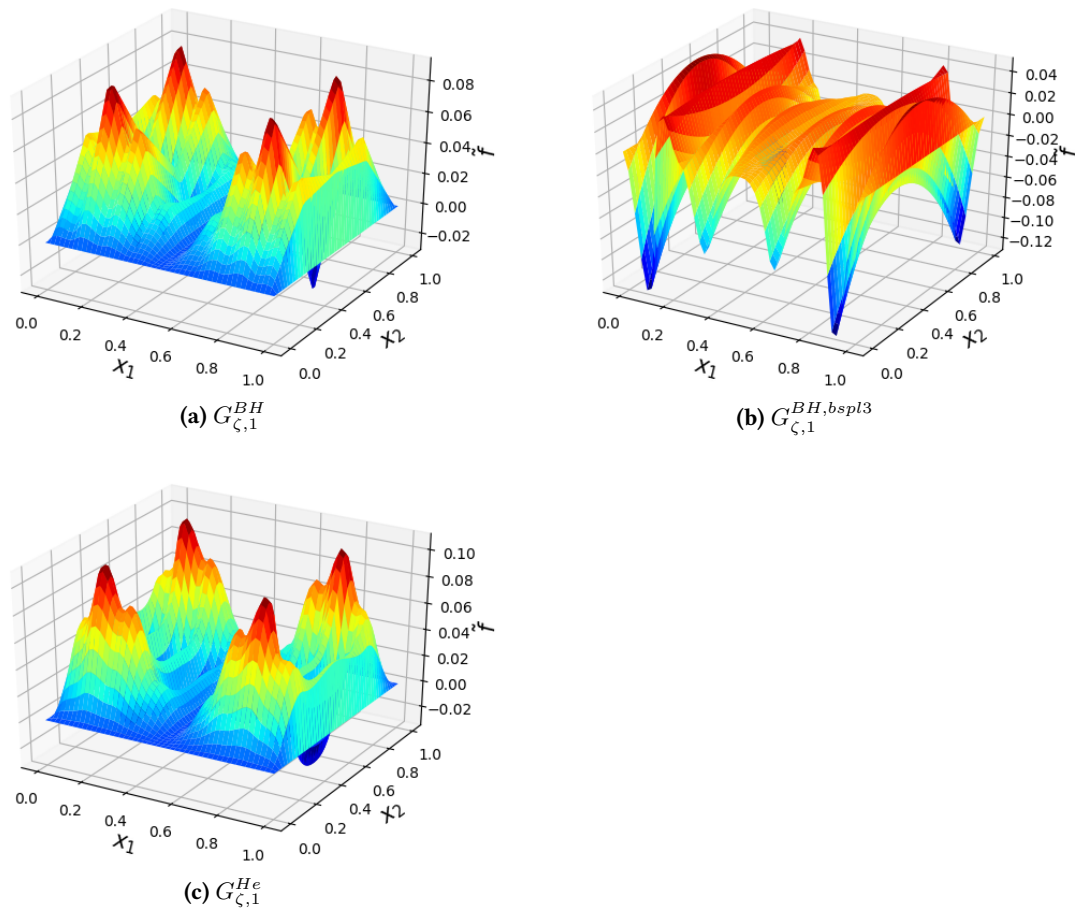


Abbildung 4.2: Das jeweilige Gitter, welches die Ableitung nach x_1 erhält. In dieser Dimension verwenden alle drei Varianten die ζ -Basisfunktion. Die andere unterscheidet sich je nach Definition.

Vermutung liegt nahe, dass B-Splines bereits eine hinreichende bzw. sogar bessere Approximation der Ableitung und gemischten Ableitung erreichen. So verschlechtern sich auch jeweils die Gitter nach de Baar und Harding mit B-Splines. Dass die B-Splines bereits eine gute Approximation der Ableitung erreicht, kann in Abbildung 4.5 und 4.6 gesehen werden. Diese zeigen den L^2 -Fehler der Ableitung nach x_1 und die gemischte Ableitung nach x_1 und x_2 . Bei den Ableitungen ist zu beachten, dass durch die Finiten Differenzen Fehler entstehen, wodurch bei 10^{-6} bereits die numerische Null erreicht wird. Bei der gemischten Ableitung verschlechtert sich dies noch einmal, doch die zugrundeliegende Tendenz ist erkennbar. Bereits die B-Splines mit Grad 3 interpolieren die gemischte Ableitung richtig und bei der Ableitung nach x_1 sind die B-Splines auch besser als die Hermite-Variante. Dies stellt auch eine mögliche Erklärung dar, warum die Gitter nach de Baar und Harding mit B-Splines schlechter abschneiden als die reinen B-Spline-Gitter. Hier ist es interessant zu sehen, ob dies nur gilt, weil die Funktion leicht durch B-Splines darstellbar ist und keine trigonometrischen Funktionen enthält, oder ob B-Splines auch dort besser abschneiden.

Eine weitere interessante Frage stellt dar, ob sich der Fehler an bestimmten Stellen besonders häuft. Um dieser Frage nachzugehen, betrachten wir als Nächstes den Betrag des punktwisen Fehlers

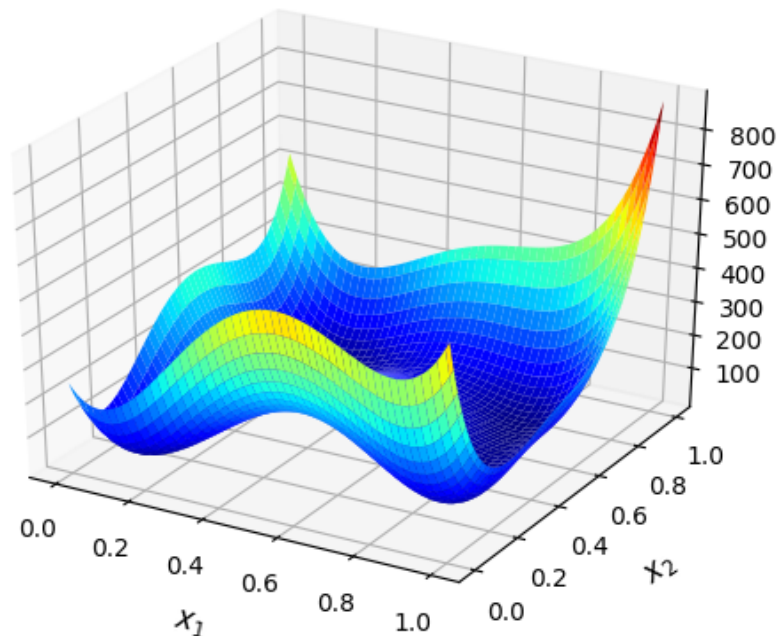


Abbildung 4.3: Abbildung der Himmelblau-Funktion

der Himmelblau-Funktion für die verschiedenen Gitterarten mit Level 6, gezeigt in Abbildung 4.7. Es fällt auf, dass alle Gittertypen besonders zu den Eckpunkten hin hohe Fehlerwerte besitzen. Dies könnte unter anderem auf die Verwendung der dünnen Gitter zurückzuführen sein. Wie anhand den grauen Gitterpunkten zu beobachten ist, ist der Großteil der Punkte auf der Hauptachse verteilt. Nach außen hin nimmt die Anzahl der Punkte stark ab, auch wenn einige Punkte auf dem Rand durch die Verwendung des Boundary-Gitters hinzukommen. Beim Hermite-Gitter und zu kleineren Teilen auch bei dem Gitter nach de Baar und Harding verteilt sich der Fehler etwas stärker über die gesamte Fläche. Die B-Splines vom Grad 5 wurden ausgenommen, da diese sowieso global gleich der Zielfunktion sind.

Branin-Funktion

Als Nächstes betrachten wir die Branin-Funktion, gegeben durch:

$$f(\vec{x}) := (x_2 - 5, 1x_1^2/(4\pi^2) + 5x_1/\pi - 6)^2 + 10(1 - 1/(8\pi)) \cos x_1 + 10, \quad \vec{x} \in [-5, 5]^2 \quad (4.3)$$

Diese Funktion beinhaltet eine trigonometrische Funktion (Cosinus) und sollte dadurch nicht so leicht durch B-Splines dargestellt werden können.

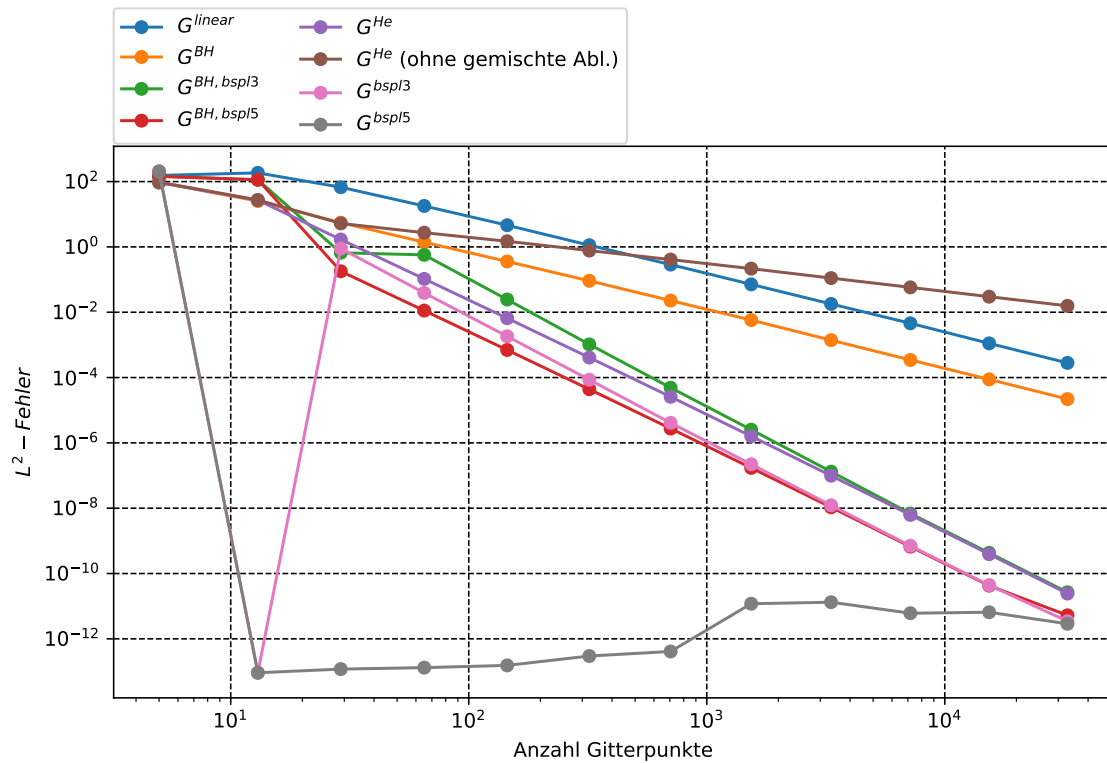


Abbildung 4.4: L^2 -Fehler der Himmelblau-Funktion von verschiedenen Gittertypen mit Level 1 bis 12

Wird nun wieder der L^2 -Fehler betrachtet, der in Abbildung 4.9 gezeigt wird, bietet sich im Großen und Ganzen ein ähnliches Bild wie bei der Himmelblau-Funktion. Ein größerer Unterschied ist jedoch, dass die B-Splines vom Grad 5 nicht bereits ab Level 3 praktisch keinen Fehler mehr besitzen. Der Grund hierfür kann in dem verwendeten Cosinus gefunden werden.

Ein interessanteres Bild ergibt sich auch bei Betrachtung des L^2 -Fehlers der Ableitungen in Abbildung 4.10 und 4.11. Hier wird die bereits zuvor geäußerte Vermutung noch einmal bestätigt, dass B-Splines bereits bessere bzw. ähnliche Approximationen der Ableitungen im Vergleich mit den Hermitebasisfunktionen ermöglichen. Die Ableitung nach x_2 wird sogar richtig interpoliert. Weiterhin ist interessant, dass auch das Gitter mit B-Splines von Grad 3 die Ableitung korrekt interpoliert. Bei der näheren Betrachtung der Ableitung nach x_2 ist zu sehen, dass der Term mit Cosinus wegfällt und dadurch ein Polynom entsteht. Werden die B-Splines im Rahmen des Gitter nach de Baar und Harding kombiniert, ist dies jedoch nicht mehr der Fall.

Bei der Ableitung nach x_1 interpoliert keine Variante bereits mit niedrigem Level die Funktion. Der beobachtete Trend ist aber abgesehen davon wieder gleich.

Durch diese Funktion bestätigt sich wieder die Vermutung, dass B-Splines bereits eine bessere Approximation der Ableitung, im Vergleich zu Hermitebasen, bei Polynomen haben. Aber auch bei Verwendung von trigonometrischen Termen ist die Verbesserung nicht groß bzw. nicht vorhanden. Dadurch stellen die zusätzlichen Informationen der Ableitung keinen Mehrwert dar und die

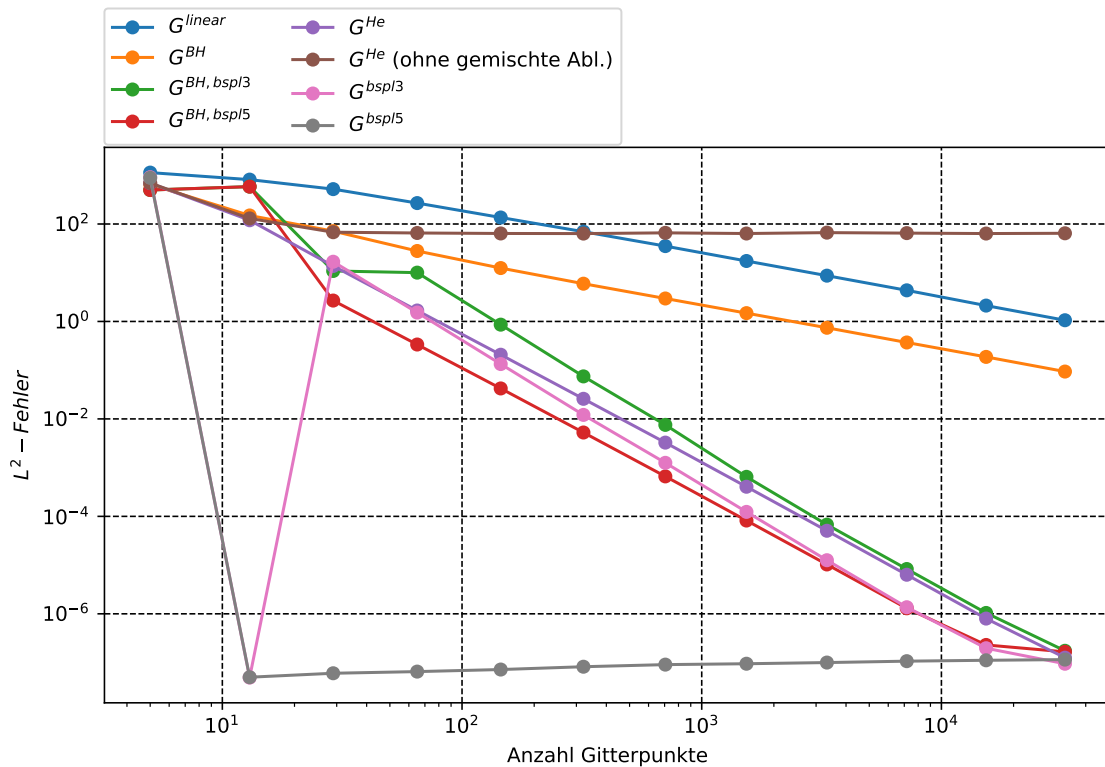


Abbildung 4.5: L^2 -Fehler der Ableitung nach x_1 der Himmelblau-Funktion von verschiedenen Gittertypen mit Level 1 bis 12

Approximation der Ableitung wird durch die Kombination mit bzw. durch die Benutzung von Hermitebasisfunktionen sogar verschlechtert.

Vergleichen wir als nächsten Schritt wieder die punktweisen Fehler in Abbildung 4.12, fällt zuerst auf, dass bei allen Gittern die Interpolation der x_1 Richtung deutlich schwerer zu fallen scheint. Dies korreliert mit der zuvor gemachten Beobachtung der L^2 -Fehlerplots. Weiterhin ist zu beobachten, dass beim Hermite-Gitter der Fehler wieder stärker verteilt ist. Bei den Gittern mit B-Splines zeigt sich ein anderes Fehlerermuster. Dieses unterscheidet sich jedoch, je nachdem ob reine B-Splines oder eine Variante nach de Baar und Harding verwendet wird.

Ackley-Funktion $d = 2$

Als letzte zweidimensionale Funktion wird die Ackley-Funktion betrachtet. Diese ist gegeben durch:

$$f(\vec{x}) := 20 + e - 20 \exp\left(\frac{|\vec{x}|_2}{5\sqrt{d}}\right) - \exp\left(\frac{1}{d} \sum_{t=1}^d \cos(2\pi x_t)\right), \quad \vec{x} \in [-1, 9]^d. \quad (4.4)$$

Wie bereits aus Abbildung 4.13 deutlich wird, ist diese deutlich schwerer und hat viele lokale Minima. Darum wird diese oft als Testfunktion bei Optimierungsproblemen verwendet. Da hier

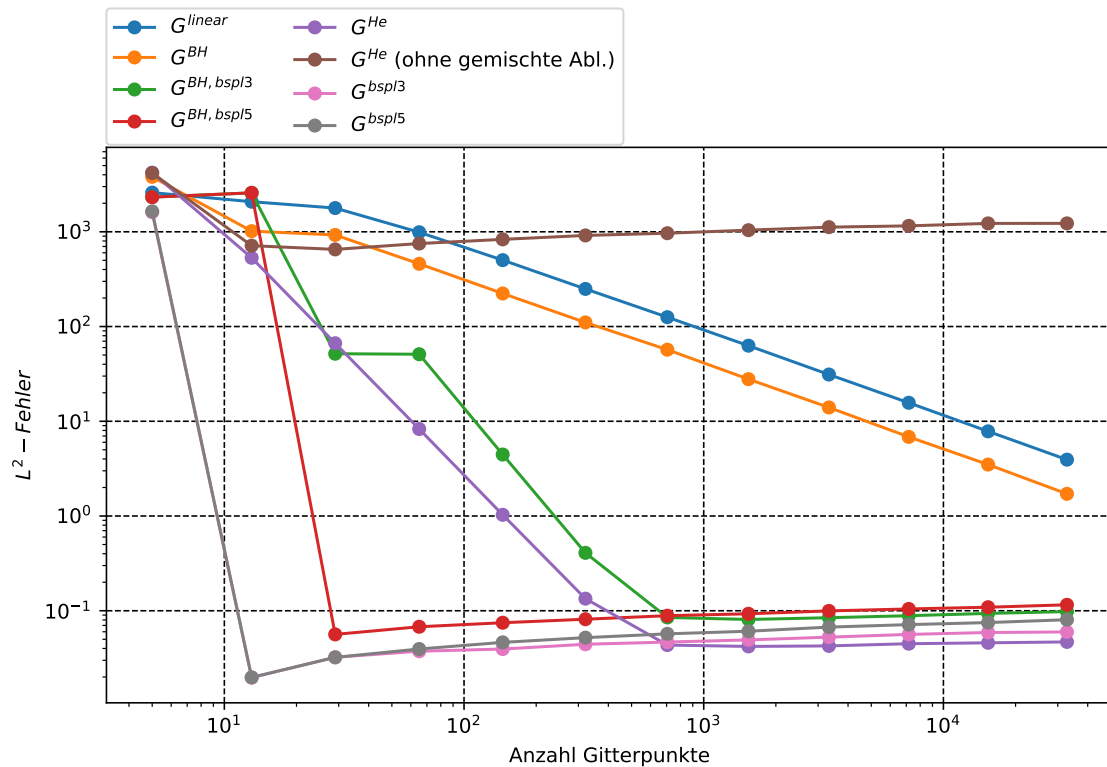


Abbildung 4.6: L^2 -Fehler der Ableitung nach x_1, x_2 der Himmelblau-Funktion von verschiedenen Gittertypen mit Level 1 bis 12

viele trigonometrische Funktionen und e-Funktionen verwendet werden ist die Erwartung, dass B-Splines vom Grad 5 nicht mehr so überlegen sind.

Betrachten wir wieder den L^2 -Fehler in Abbildung 4.14 fällt auf, dass der konvergente Bereich, verglichen mit den anderen Funktionen, erst mit viel mehr Gitterpunkten erreicht wird. Dies lässt sich leicht durch die Form erklären. Wie erwartet ist die Interpolation der B-Splines vom Grad 5 nicht mehr so gut. Sonst sind wieder die zwei Gruppen erkennbar. Das Gitter nach de Baar und Harding ist wieder leicht besser als die linearen Basisfunktionen. Die zweite Gruppe stellen die restlichen Funktionen dar, ausgenommen der Hermite-Gitter ohne gemischte Ableitung. Dabei ist jedoch noch auffällig, dass das Hermite-Gitter hier um einen konstanten Faktor besser ist und die Unterschiede der Gitter nach de Baar und Harding mit den normalen B-Splines nicht so stark sind. Eine mögliche Erklärung könnte sein, dass hier die Hermitebasisfunktionen gut abschneiden und diese deswegen nicht zu einer Verschlechterung führen.

4.3 Mehrdimensionale Funktionen

Da wir bisher nur Testfunktionen im zweidimensionalen Fall angesehen haben, ist es noch sinnvoll mehrdimensionale Funktionen zu betrachten. Hier fällt besonders beim Hermite-Gitter eine Vielzahl an zusätzlich benötigten Gittern und Informationen ins Gewicht.

4 Evaluation

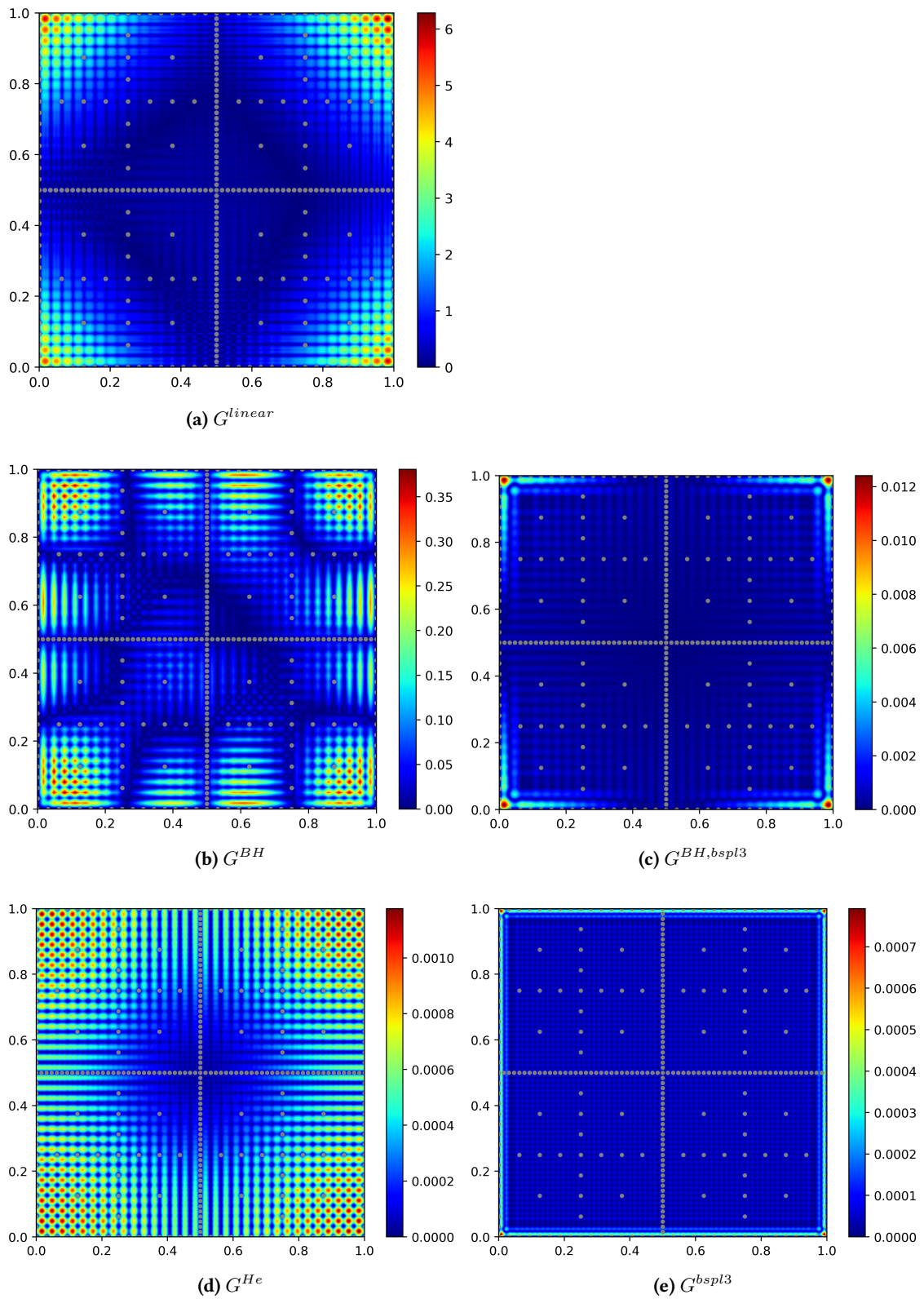


Abbildung 4.7: Punktwiser Fehler der Himmelblau-Funktion mit Level 6. In Grau sind die Gitterpunkte eingezeichnet. Zu beachten sind die verschiedenen Skalierungen.

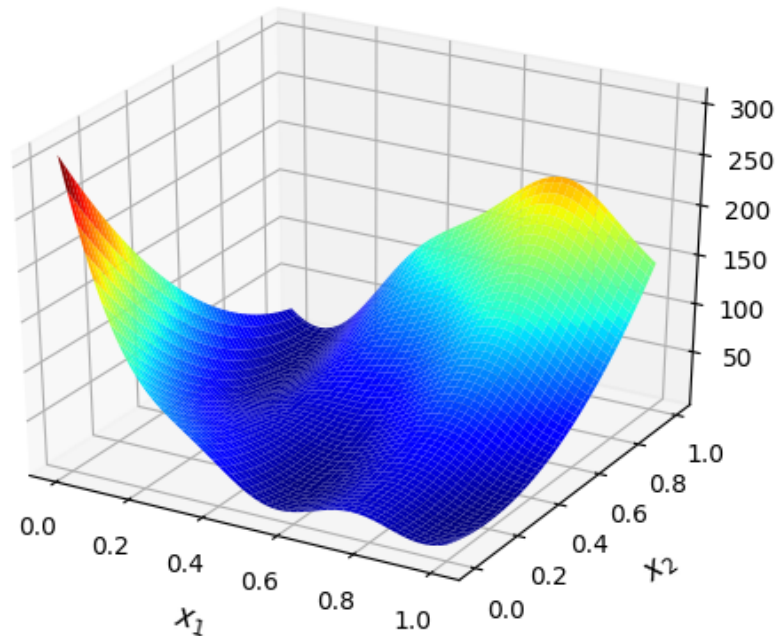


Abbildung 4.8: Abbildung der Branin-Funktion

Die höchste verwendete Testfunktion hat nur sechs Dimensionen, weil das Hermite-Gitter aus sehr vielen einzelnen Gittern besteht und auf der anderen Seite bei den B-Splines die Koeffizienten berechnet werden müssen. Hierdurch steigt die Rechenzeit und der Speicherbedarf stark an.

Rosenbrock-Funktion

Die Rosenbrock-Funktion ist gegeben durch:

$$f(\vec{x}) := \sum_{t=1}^{d-1} (100(x_{t+1} - x_t^2)^2 + (1 - x_t)^2), \quad \vec{x} \in [-5, 10]^d. \quad (4.5)$$

Hierbei handelt es sich wieder um ein reines Polynom ohne trigonometrische Terme. Dadurch ist die Erwartung, dass B-Splines eine besonders gute Interpolation ermöglichen. In Abbildung 4.15 wird der L2-Fehler der Rosenbrock-Funktion in vier Dimensionen gezeigt. Zu sehen ist, dass die Resultate im Großen und Ganzen ähnlich ausfallen wie die Himmelblau-Funktion. Da diese Funktion ebenfalls vom Grad 4 ist, ist dies nicht verwunderlich. Die lineare Basis und das Gitter nach de Baar und Harding haben wieder ungefähr die gleiche Steigung, wobei letzteres um einen konstanten Faktor besser abschneidet. Noch auffällig ist der große Abstand zwischen dem Hermite-Gitter und den Anderen. Ebenso ist das Gitter nach de Baar und Harding mit B-Splines

4 Evaluation

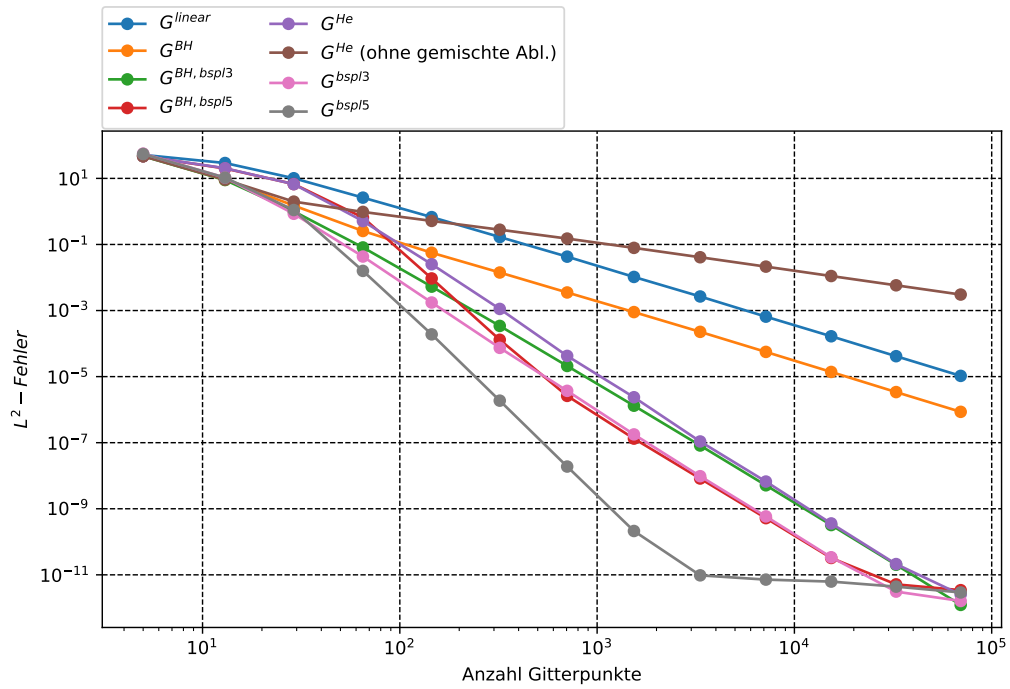


Abbildung 4.9: L^2 -Fehler der Branin-Funktion von verschiedenen Gittertypen mit Level 1 bis 13

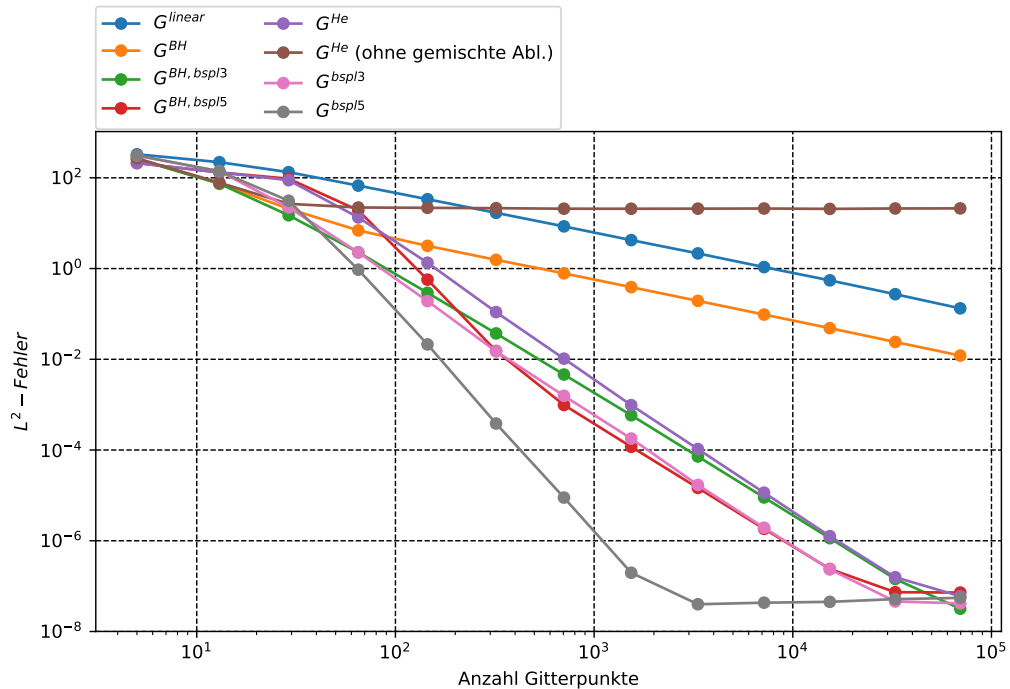


Abbildung 4.10: L^2 -Fehler der Ableitung nach x_1 der Branin-Funktion von verschiedenen Gittertypen mit Level 1 bis 13

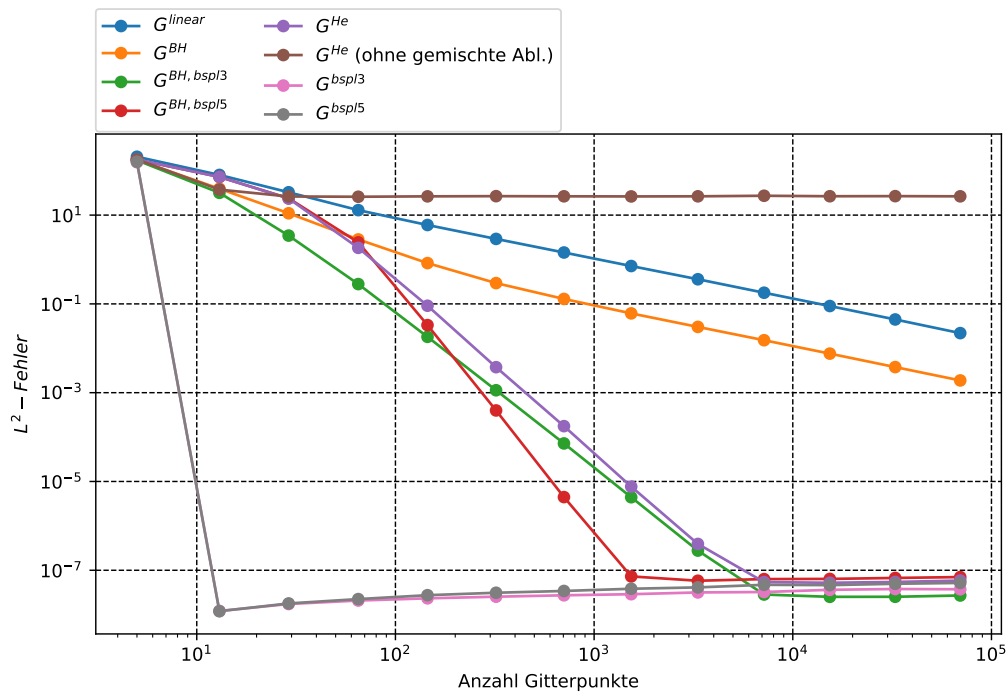


Abbildung 4.11: L^2 -Fehler der Ableitung nach x_2 der Branin-Funktion von verschiedenen Gittertypen mit Level 1 bis 13

vom Grad 3 deutlich schlechter als die reine B-Spline-Variante. Dies deutet stark darauf hin, dass Hermitebasisfunktionen, bei dieser Funktion auch zu einer Verschlechterung des Endresultats führen.

Die Verschlechterung des Gitters mit B-Splines vom Grad 5 ist mit großer Wahrscheinlichkeit auf die hohen Werte der Rosenbrock-Funktion zurückzuführen. Durch mehr Gitterpunkte erhöhen sich so die Rundungsfehler.

Bei Betrachtung der beispielhaften Ableitung nach x_2 , dargestellt in Abbildung 4.16, sind keine weiteren Auffälligkeiten zu erkennen.

Erhöhen wir als nächstes die Dimension der Funktion auf 6, sind in Abbildung 4.17 aus den bereits erwähnten Platz- und Zeitgründen nur die ersten Level abgebildet. Bereits hier zeigt sich eine Tendenz, die mit den vorherigen Resultaten ungefähr übereinstimmt. So gibt es auch hier wieder die selben Gruppierungen. Aufgrund der vielen Zusatzinformationen des Hermite-Gitters ist diese relativ schlechte Approximation jedoch trotzdem überraschend.

Sin-Kuppel-Funktion

Die nächste Funktion, gegeben durch:

$$f(\vec{x}) := \sum_{j=1}^d \sin(\pi x_j), \quad \vec{x} \in [0, 1]^d, \quad (4.6)$$

4 Evaluation

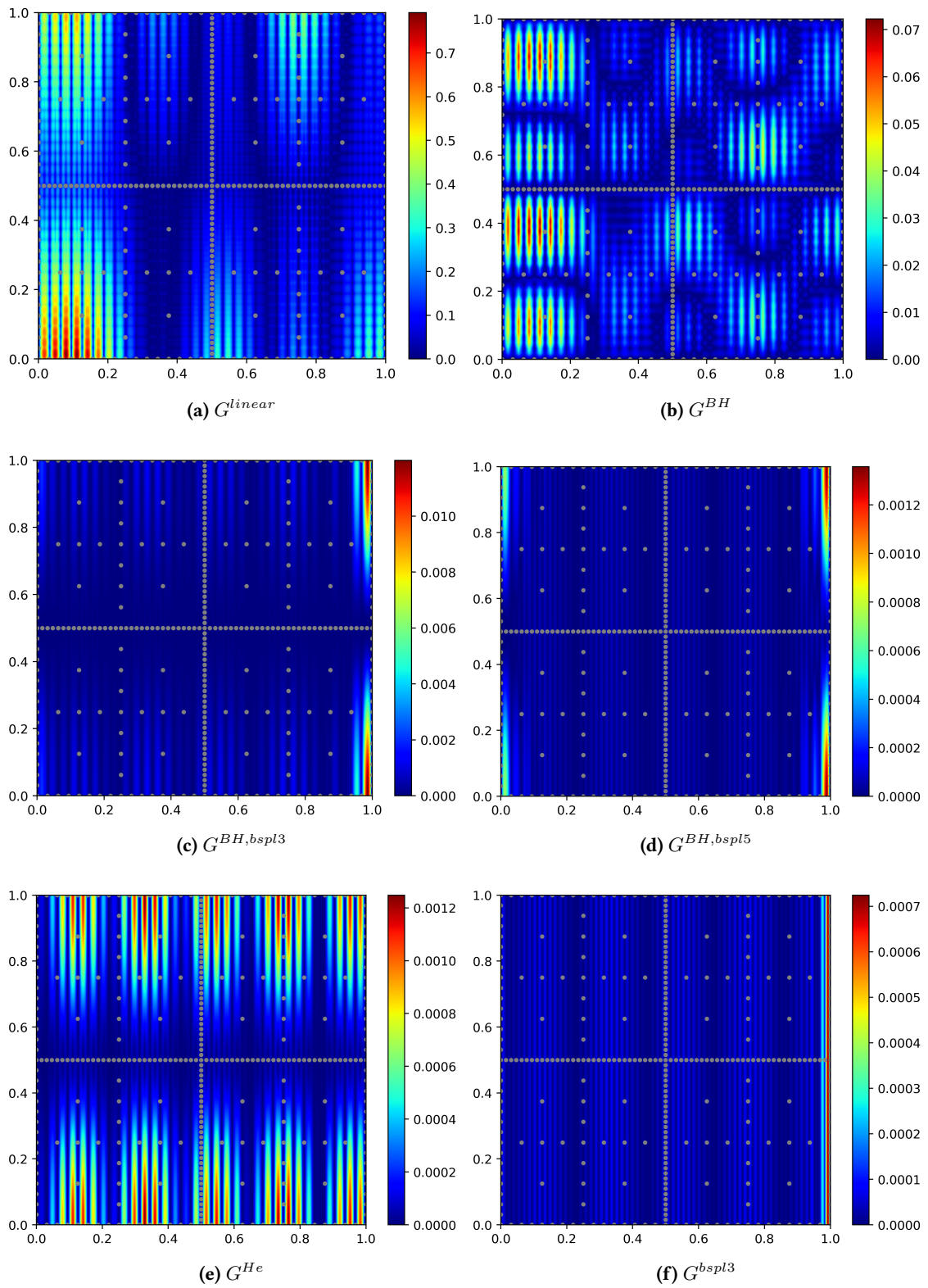


Abbildung 4.12: Punktwiser Fehler, berechnet durch $|f - u|$, der Branin-Funktion mit Level 6. Zu beachten sind die verschiedenen Skalierungen.

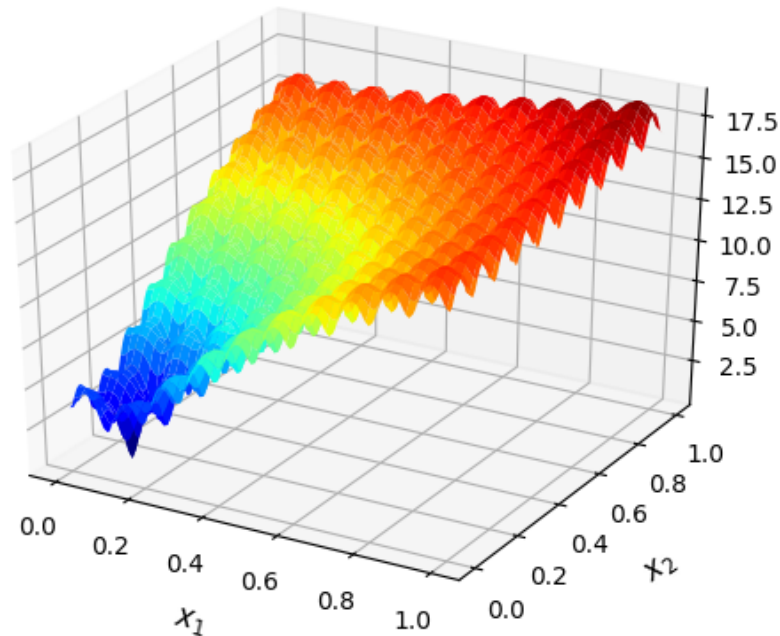


Abbildung 4.13: Abbildung der Ackley-Funktion mit $d = 2$

besteht nur aus einer trigonometrischen Funktion und kann deshalb wieder nicht so leicht durch B-Splines dargestellt werden. Abbildung 4.18 stellt für diese den L^2 -Fehler in drei Dimensionen dar. Hier ist zu sehen, dass beide B-Spline-Varianten nicht gleich die richtige Interpolation liefern. Die restlichen Ergebnisse sind jedoch vergleichbar mit den Vorhergegangenen. Einzig die Tatsache, dass das Hermite-Gitter um einen konstanten Faktor besser ist, als die anderen Gitter dieser Gruppe, stellt einen Unterschied dar.

In Abbildung 4.19 wird der L^2 -Fehler bei Erweiterung auf sechs Dimensionen dargestellt. Auch hier ist wieder die Anzahl der Level relativ niedrig, wodurch nur eine erste Tendenz zu erkennen ist und noch nicht auf die Konvergenzordnung geschlossen werden kann. Im Vergleich zum vierdimensionalen Fall stellt sich im Anfangsbereich schon ein relativ ähnliches Bild. Dadurch kann vermutet werden, dass sich auch in höheren Leveln nicht viel ändert.

4 Evaluation

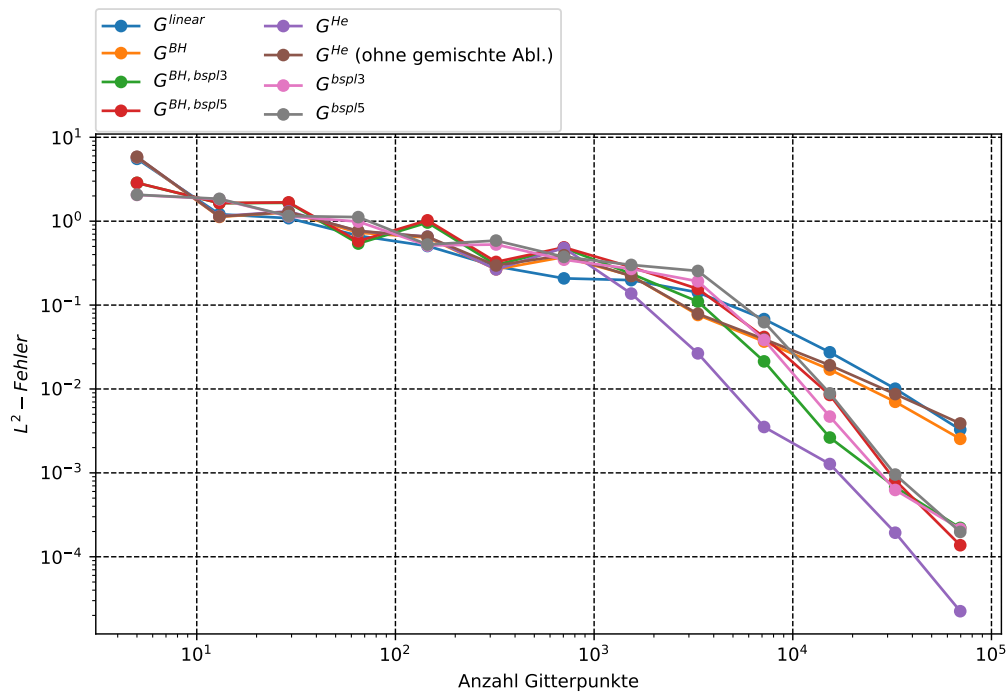


Abbildung 4.14: L^2 -Fehler der Ackley-Funktion von verschiedenen Gittertypen mit Level 1 bis 14

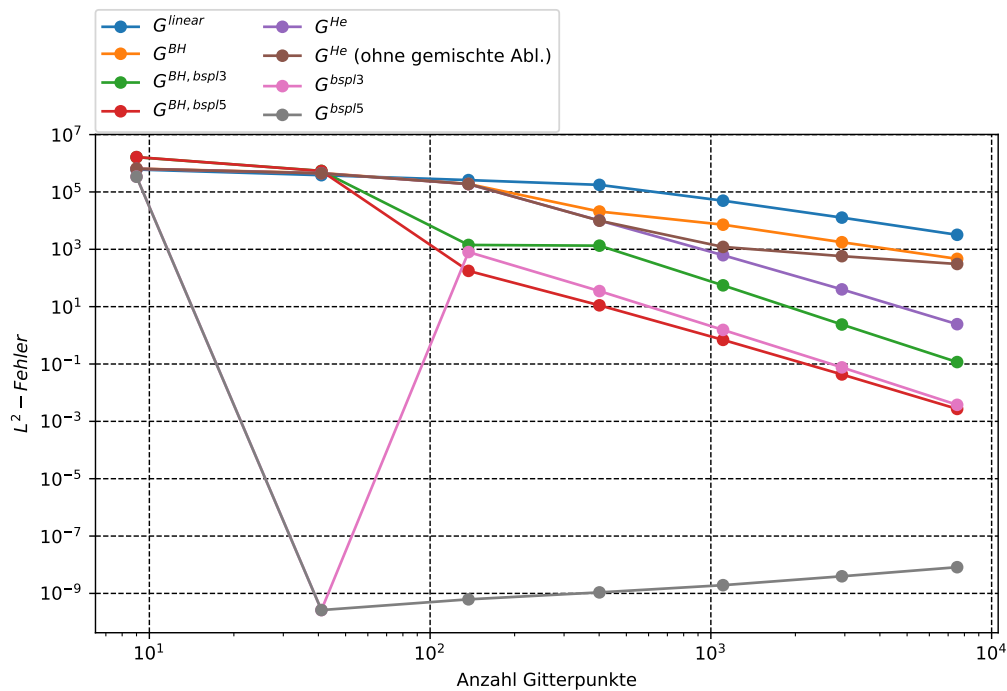


Abbildung 4.15: L^2 -Fehler der Rosenbrock-Funktion mit $d = 4$

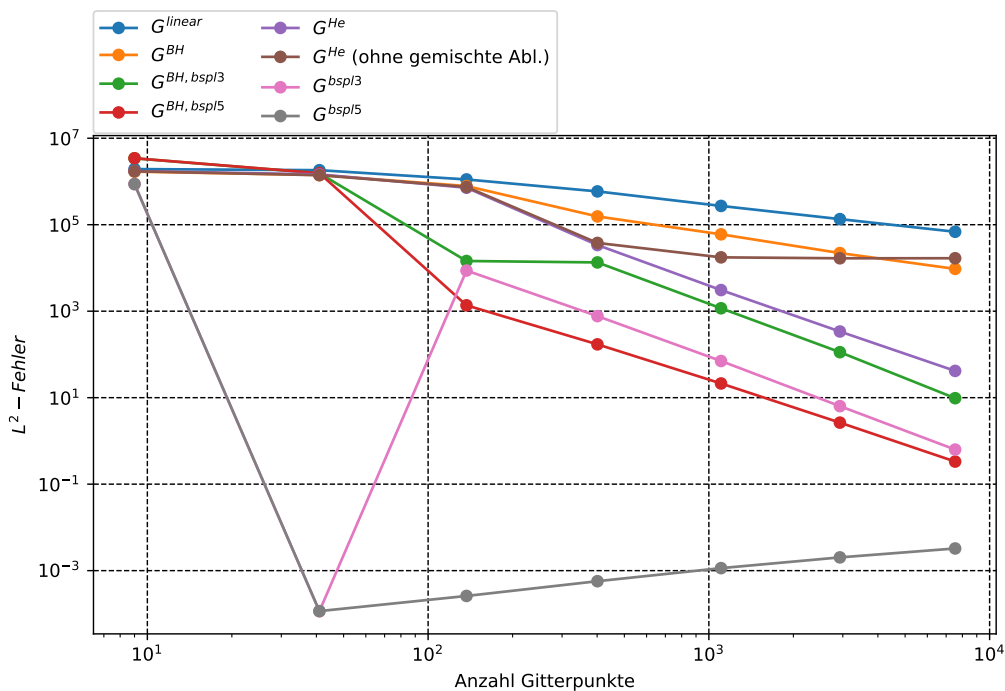


Abbildung 4.16: L^2 -Fehler der Ableitung der Rosenbrock-Funktion nach x_2 mit $d = 4$

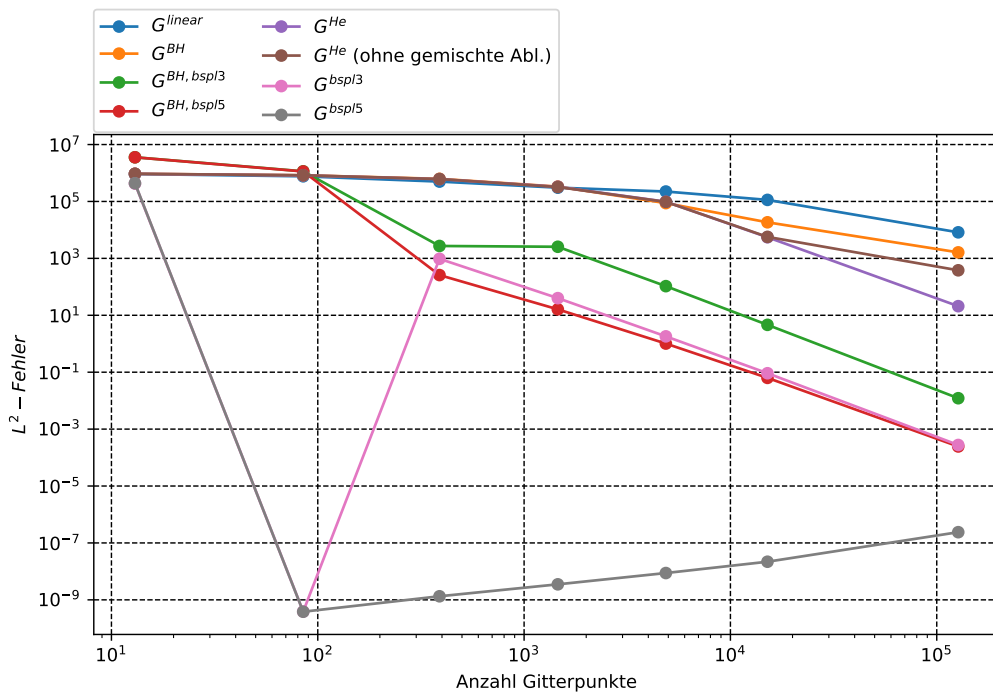


Abbildung 4.17: L^2 -Fehler der Rosenbrock-Funktion mit $d = 6$

4 Evaluation

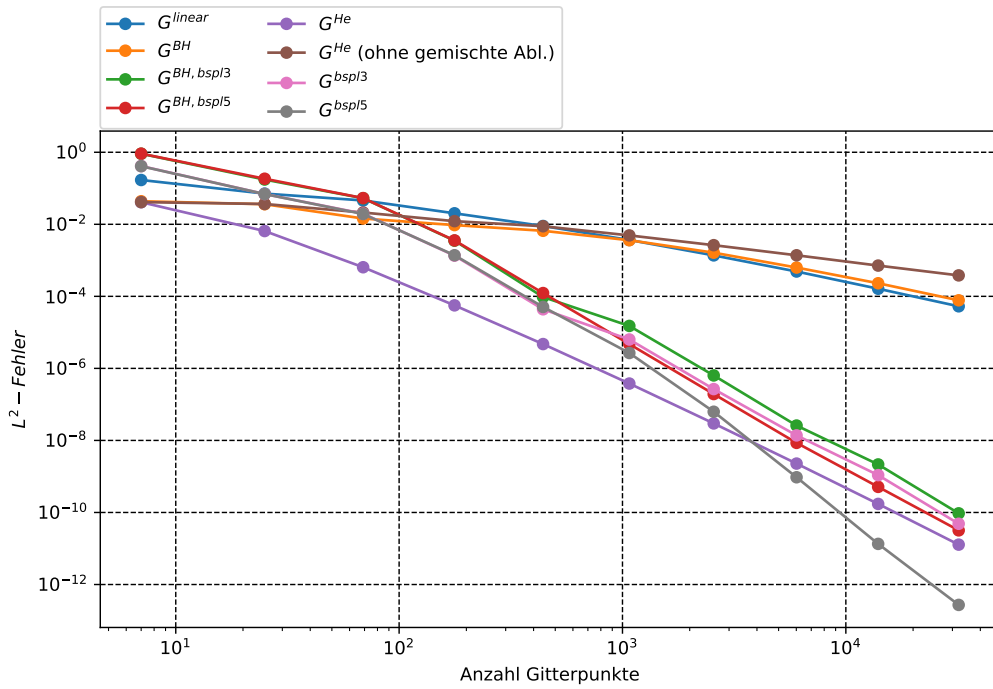


Abbildung 4.18: L^2 -Fehler der Sin-Kuppel-Funktion mit $d = 3$

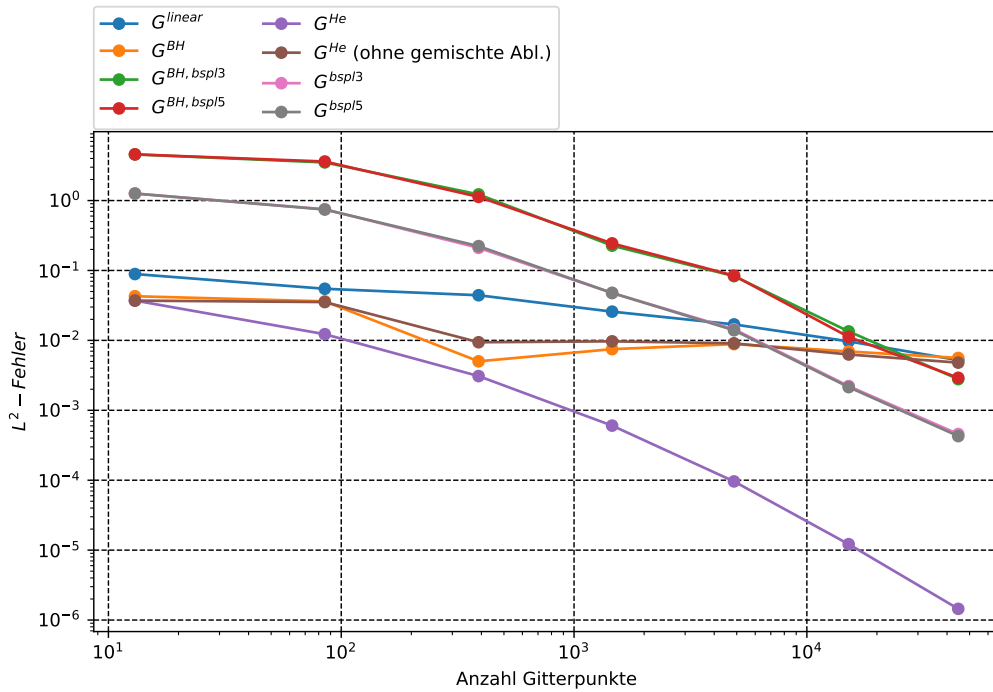


Abbildung 4.19: L^2 -Fehler der Sin-Kuppel-Funktion mit $d = 6$

5 Zusammenfassung und Ausblick

In der Arbeit wurden verschiedene Ansätze vorgestellt, um Ableitungsinformationen mitzuverwenden und diese anschließend mit herkömmlichen Interpolationsmethoden zu vergleichen. Dafür wurden zunächst die dünnen Gitter mit Kombinationstechnik eingeführt. Anschließend wurden verschiedene Basisfunktionen vorgestellt, die die Dünngitter-Struktur verwenden und für die darauffolgenden Ansätze nötig waren. Im nächsten Schritt wurden die verschiedenen gradienten-erweiterten Gitter vorgestellt. Um die Qualität zu überprüfen, wurden diese mit herkömmlichen Methoden, die keine Ableitungen verwenden, verglichen.

Die Ergebnisse zeigen, dass zumindest bei den getesteten Funktionen im Großen und Ganzen B-Splines am besten abschneiden. Die meiste Zeit waren B-Splines und die Varianten, die Hermitebasisfunktionen verwendet haben, ungefähr gleich gut. Eine Ausnahme stellen die B-Splines vom Grad 5 dar, die bei den richtigen Bedingungen (Polynome vom Grad ≤ 5) die Funktion richtig interpolieren. Die Zweite Ausnahme stellte das Gitter nach de Baar und Harding dar, welches Hütchenfunktionen verwendet. Wird jedoch der Aufwand für die Berechnung der gradientenerweiterten Gitter miteinbezogen, schneiden reine B-Spline-Gitter klar am Besten ab. So erzielt das Hermite-Gitter bei manchen Funktionen leicht bessere Ergebnisse, aber diese stehen in keinem Verhältnis zu den zusätzlich benötigten Daten und Gitter. Die Gitter nach de Baar und Harding mit B-Splines waren nie in relevanter Weise besser im Vergleich mit den reinen B-Spline-Gittern. Im Gegenteil wurde dadurch sogar eine teilweise eine korrekte Interpolation verhindert und das Resultat verschlechtert. Auch hier ist außerdem ein zusätzlicher Berechnungsaufwand notwendig und es werden, wenn auch nicht ganz so viele, zusätzliche Daten benötigt. Betrachtet man das lineare Gitter und das Gitter nach de Baar und Harding isoliert, stellt es eine Verbesserung dar. Wird es jedoch ins Verhältnis mit den B-Splines gesetzt, sind diese, zumindest bei den getesteten Funktionen, immer besser.

In zukünftigen Arbeiten könnten für eine bessere Vergleichbarkeit noch weitere Funktionen und auch praktische Anwendungen untersucht werden. Weiterhin wurden nur Boundary-Gitter verwendet und keine anderen Gitterarten betrachtet, die eventuell nochmals zu anderen Resultaten führen könnten. Die vorgeschlagenen Methoden könnten auch noch für Ableitungen höherer Ordnung verallgemeinert werden.

Literaturverzeichnis

- [BG04] H.-J. Bungartz, M. Griebel. „Sparse grids“. In: *Acta numerica* 13 (2004), S. 147–269 (zitiert auf S. 9, 11, 12).
- [BH15] J. H. de Baar, B. Harding. „A gradient-enhanced sparse grid algorithm for uncertainty quantification“. In: *International Journal for Uncertainty Quantification* 5.5 (2015) (zitiert auf S. 7, 19).
- [De 72] C. De Boor. „On calculating with B-splines“. In: *Journal of Approximation theory* 6.1 (1972), S. 50–62 (zitiert auf S. 15).
- [GGT01] J. Garcke, M. Griebel, M. Thess. „Data mining with sparse grids“. In: *Computing* 67.3 (2001), S. 225–253 (zitiert auf S. 9).
- [GSZ92] M. Griebel, M. Schneider, C. Zenger. *A Combination Technique For The Solution Of Sparse Grid Problems*. 1992 (zitiert auf S. 12).
- [HGC07] M. Hegland, J. Garcke, V. Challis. „The combination technique and some generalisations“. In: *Linear Algebra and its Applications* 420.2 (2007), S. 249–275. ISSN: 0024-3795. DOI: <https://doi.org/10.1016/j.laa.2006.07.014>. URL: <http://www.sciencedirect.com/science/article/pii/S002437950600334X> (zitiert auf S. 12).
- [HH15] K. Höllig, J. Hörner. *Approximation and Modeling with B-Splines*. Other Titles in Applied Mathematics. Society for Industrial und Applied Mathematics, 2015. ISBN: 9781611972948. URL: <https://books.google.de/books?id=ynYXAgAAQBAJ> (zitiert auf S. 17).
- [KW05] A. Klimke, B. Wohlmuth. „Algorithm 847: spinterp: Piecewise multilinear hierarchical sparse grid interpolation in MATLAB“. In: *ACM Transactions on Mathematical Software (TOMS)* 31.4 (2005), S. 561–579 (zitiert auf S. 9).
- [NR96] E. Novak, K. Ritter. „High dimensional integration of smooth functions over cubes“. In: *Numerische Mathematik* 75.1 (1996), S. 79–97 (zitiert auf S. 9).
- [Pfl10] D. Pflüger. *Spatially adaptive sparse grids for high-dimensional problems*. Verlag Dr. Hut, 2010 (zitiert auf S. 9, 12).
- [Val14] J. Valentin. „Hierarchische Optimierung mit Gradientenverfahren auf Dünngitterfunktionen“. Masterarbeit. IPVS, Fachbereich Informatik, Universität Stuttgart, 2014. DOI: [10.18419/opus-3462](https://doi.org/10.18419/opus-3462) (zitiert auf S. 9, 15, 16).
- [VP16] J. Valentin, D. Pflüger. „Hierarchical Gradient-Based Optimization with B-Splines on Sparse Grids“. In: *Sparse Grids and Applications – Stuttgart 2014*. Hrsg. von J. Garcke, D. Pflüger. Bd. 109. Lecture Notes in Computational Science and Engineering. Springer, 2016, S. 315–336. DOI: [10.1007/978-3-319-28262-6_13](https://doi.org/10.1007/978-3-319-28262-6_13) (zitiert auf S. 17).
- [Zen91] C. Zenger. „Sparse Grids“. In: *Notes on Numerical Fluid Mechanics* 31 (1991). Hrsg. von W. Hackbusch, S. 241–251 (zitiert auf S. 9).

Alle URLs wurden zuletzt am 17. 12. 2017 geprüft.

Erklärung

Ich versichere, diese Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommene Aussagen als solche gekennzeichnet. Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens. Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht. Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

Ort, Datum, Unterschrift