Institut für Informatik
der
Technischen Universität München

# Dünne Gitter
# und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung

**Hans-Joachim Bungartz**

Vollständiger Abdruck der von der Fakultät für Mathematik und Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

**Doktors der Naturwissenschaften**

genehmigten Dissertation.

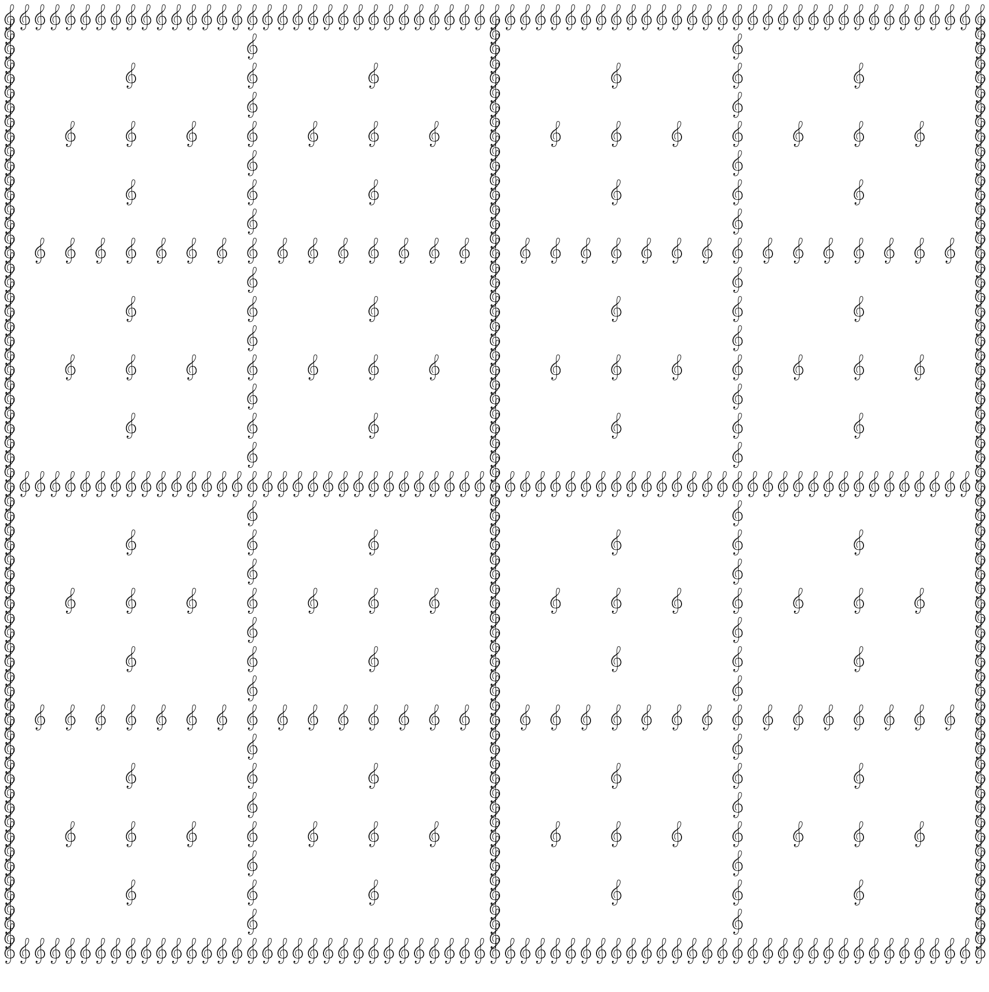
| | | |
|---|---|---|
| Vorsitzender: | | Univ.-Prof. Dr. J. Eickel |
| Prüfer der Dissertation: | | |
| | 1. | Univ.-Prof. Dr. Chr. Zenger |
| | 2. | Univ.-Prof. Dr. R. H. W. Hoppe |

Die Dissertation wurde am 7.2.1992 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik und Informatik am 3.6.1992 angenommen.

**Dünne Gitter**
**und deren Anwendung bei der adaptiven Lösung**
**der dreidimensionalen Poisson-Gleichung**

## Zusammenfassung


Zur Interpolation glatter Funktionen von $d$ Veränderlichen werden anstelle der regelmäßigen, $d$-linearen Finite-Elemente-Räume geeignete Teilräume betrachtet. Dies entspricht dem Übergang von regelmäßigen, *vollen* Gittern auf sogenannte *dünne* Gitter. Es wird gezeigt, daß dünne Gitter bei gleichen Genauigkeitsanforderungen wesentlich weniger Punkte benötigen als volle Gitter. Als Anwendung dieser Resultate wird dann ein adaptives Verfahren zur Lösung der dreidimensionalen Poisson-Gleichung auf dünnen Gittern vorgestellt. Die Methode stützt sich auf einen Finite-Elemente-Ansatz und auf hierarchische Basen. Ein Konvergenzbeweis und numerische Experimente schließen die Analyse des Verfahrens ab.



## I n h a l t

**1. Einleitung.** Zahlreiche Aufgabenstellungen der Ingenieur- und Naturwissenschaften bedürfen zu ihrer genauen, d.h. wirklichkeitsnahen Modellierung mittels eines diskretisierten Problems einer extrem feinen Auflösung der jeweils verwendeten Gitter. Als Beispiel sei hier etwa die Strömungsmechanik genannt, wo zur Erfassung kleinster Wirbel entsprechend feine Netze gewählt werden müssen. Diese Notwendigkeit führt insbesondere bei drei- oder höherdimensionalen Problemen oftmals zu einer Anzahl von Unbekannten, die die Rechenkapazität selbst leistungsfähigster Computer auch auf absehbare Zeit übersteigt. Ähnliches gilt für den Bereich der Computergraphik und Animation, wo — beispielsweise bei realistischen Flugsimulatoren — riesige Datenmengen gespeichert und manipuliert werden müssen.

Eine Möglichkeit, solchen Schwierigkeiten bei der diskreten Darstellung von Funktionen beziehungsweise bei der Behandlung numerischer Probleme wie der Quadratur oder der Lösung partieller Differentialgleichungen beizukommen, stellt die Technik der adaptiven Verfeinerung dar. Hier werden regelmäßige Gitterstrukturen betrachtet, die jedoch lokal eine unterschiedlich hohe Auflösung aufweisen. Dadurch wird erreicht, daß nur dort eine große Gitterfeinheit gewählt wird, wo diese auch aufgrund der Problemstellung erforderlich ist.

In dieser Arbeit behandeln wir nun ein weiteres, mächtigeres Konzept zur Reduzierung der Datenmengen beziehungsweise zur Reduzierung der Anzahl von Unbekannten bei (annähernd) gleichem Informationsgehalt beziehungsweise gleicher Genauigkeit — die in [38] vorgestellten *dünnen Gitter*. Im Gegensatz zu herkömmlichen, *vollen Gittern* nutzt hier bereits die reguläre Gitterstruktur Glattheitseigenschaften der Daten aus, um diese in komprimierter Form darzustellen beziehungsweise das betreffende Problem mit einer deutlich geringeren Zahl an Unbekannten zu lösen. Darüber hinaus bieten auch dünne Gitter auf sehr bequeme Art und Weise die Möglichkeit adaptiver Verfeinerung. Sie erweisen sich als ein Konzept, das geradezu auf Adaptivität abzielt und für das adaptiv verfeinerte (dünne) Gitter der Normal- und reguläre (dünne) Gitter der Spezialfall sind.

Wesentlicher Bestandteil der Dünn-Gitter-Theorie ist das Prinzip der hierarchischen Basen (siehe [36]). Trotz ihrer Aktualität in unserem Zusammenhang hat die hierarchische Sehweise weit zurückreichende Wurzeln. So berichtet Yserentant in [37] über eine Arbeit von Faber aus dem Jahre 1909 ([11]), in der unter anderem die Idee der hierarchischen Darstellung von Funktionen vorgestellt wird.

Die vorliegende Arbeit gliedert sich in zwei Teile. Zunächst werden in den Abschnitten 2 bis 5 eine Reihe theoretischer Ergebnisse zu dünnen Gittern beliebiger Dimensionalität präsentiert, die vor allem die Darstellung glatter, d.h. gewissen Differenzierbarkeitsanforderungen genügender Funktionen sowie den damit verbundenen Speicheraufwand zum Thema haben. Anschließend wird in den Abschnitten 6 bis 8 mit der Lösung des zwei- beziehungsweise dreidimensionalen Poisson-Problems auf dem Einheitsquadrat beziehungsweise Einheitskubus eine erste über die Darstellung von Funktionen hinausgehende Anwendung dünner Gitter vorgestellt.

In Abschnitt 2 führen wir auf dem betrachteten Grundgebiet $\bar{\Omega}^{(d)} = [0,1]^d$ äquidistante, rechtwinklige Gitter (*volle Gitter*) sowie die Funktionenräume der darauf stückweise $d$-linearen Funktionen ein. Diese Räume werden als direkte Summen von Teilräumen dargestellt, welche eine hierarchische Sehweise gestatten. Ausgehend von den üblichen stückweise $d$-linearen Finite-Elemente-Basen auf den einzelnen Teilräumen gelangen wir so zu einer hierarchischen Basis unserer Funktionenräume. Die Wahl rechtwinkliger und produktorientierter Elemente gestattet dabei eine einfache Verallgemeinerung vom eindimensionalen Fall auf die Situation bei $d$ Raumdimensionen. Wir geben Schranken für den Interpolationsfehler bezüglich der $L_2$- sowie der Maximumsnorm an und schätzen die Beiträge der einzelnen Teilräume zum Interpolationsfehler ab.

Dieses Maß für die Bedeutung eines Teilraumes bei der Interpolation ist der Ausgangspunkt für eine erste Definition dünner Gitter und der entsprechenden Funktionenräume in Abschnitt 3. Teilräume, die im Verhältnis zur Zahl der Unbekannten, die sie in die Darstellung einer Funktion einbringen, nur einen geringen Beitrag zum Interpolationsfehler leisten, werden jetzt bei der Summendarstellung der Dünn-Gitter-Räume nicht mehr berücksichtigt. Betrachtet man die zugrundeliegenden Gitter, so enthalten diese im Vergleich zu vollen Gittern weniger Punkte — es entstehen sogenannte *dünne Gitter* (siehe [38]). Es wird gezeigt, daß die Dimension der resultierenden Dünn-Gitter-Räume von der Ordnung $O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^{d-1})$ ist, was gegenüber $O(h_n^{-d})$ bei Verwendung voller Gitter eine signifikante Ersparnis bedeutet. Für den Interpolationsfehler werden bezüglich $L_2$- und Maximumsnorm Schranken der Größenordnung $O(h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^{d-1})$ angegeben. Somit gestatten dünne Gitter bei nur leicht verschlechterter Approximationsgenauigkeit die Repräsentation glatter Funktionen mit deutlich weniger Unbekannten als volle Gitter.

In Abschnitt 4 stellen wir dann ein auf der Energienorm basierendes Konzept zur Definition dünner Gitter und der entsprechenden Funktionenräume vor. Es zeigt sich, daß diese energieorientierte Vorgehensweise in verschiedener Hinsicht vorteilhaft ist. So gestattet sie beispielsweise Aussagen zur Genauigkeit des im zweiten Teil dieser Arbeit beschriebenen Finite-Elemente-Verfahrens zur Lösung von Poisson-Problemen auf dünnen Gittern. Außerdem zeigen wir, daß die Dimension der resultierenden Dünn-Gitter-Räume nur von der Ordnung $O(h_n^{-1})$, der in der Energienorm gemessene Interpolationsfehler jedoch wie beim vollen Gitter von der Ordnung $O(h_n)$ ist. Für diese neue Klasse dünner Gitter ist somit sowohl die Anzahl der Gitterpunkte als auch der Interpolationsfehler größenordnungsmäßig unabhängig von der Dimensionalität $d$.

Als Abschluß des ersten Teils erweitern wir in Abschnitt 5 die bislang eingeführten *regulären* dünnen Gitter um die Möglichkeit adaptiver Verfeinerung. Wir geben geeignete Adaptionsstrategien an und stellen so für jede der beiden Klassen dünner Gitter aus Abschnitt 3 und Abschnitt 4 ein adaptives Analogon bereit. Es zeigt sich, daß dünne Gitter nicht einfach als adaptiv verfeinerte volle Gitter angesehen werden dürfen. Vielmehr ist die Unterscheidung zwischen vollen Gittern einerseits und dünnen Gittern andererseits von grundsätzlich anderer Qualität als die Unterscheidung zwischen regulären und adaptiv verfeinerten Gittern. Schließlich stellen wir noch die in unserer Implementierung verwendete Binärbaum-Datenstruktur vor, die auf natürliche Weise die Realisierung adaptiver Verfeinerung ermöglicht.

Den zweiten Teil dieser Arbeit beginnen wir in Abschnitt 6 mit der Beschreibung unseres adaptiven Verfahrens zur Lösung zweidimensionaler Poisson-Probleme auf dem Einheitsquadrat. Wir zeigen dabei, daß dieses Verfahren in verschiedener Hinsicht als Verallgemeinerung der HB-MG-Methode von Bank, Dupont und Yserentant angesehen werden kann (siehe [5]). Ausgehend von der hierarchischen Basis aus Abschnitt 2 wählen wir einen Ansatz gemäß der Methode der Finiten Elemente. Für die Berechnung der bei der Diskretisierung der rechten Seite entstehenden Integrale stellen wir eine Technik zur adaptiven numerischen Quadratur auf dünnen Gittern vor. Das resultierende lineare Gleichungssystem lösen wir mit Hilfe des Gauß-Seidel-Verfahrens. Dabei muß die Steifigkeitsmatrix nicht explizit angegeben werden. Dadurch läßt sich der Algorithmus mit konstantem Rechenaufwand pro Iterationsschritt und pro Unbekannter realisieren. Es wird gezeigt, daß unser Verfahren eine Näherungslösung liefert, die mit einem in der Energienorm gemessenen Fehler der Ordnung $O(h_n)$ behaftet ist. Somit leisten dünne Gitter hinsichtlich der Energienorm größenordnungsmäßig gleich viel wie volle Gitter, obwohl sie mit signifikant weniger Unbekannten auskommen. Darüber hinaus zeigen numerische Experimente, daß die $L_2$- beziehungsweise Maximumsnorm des Fehlers wie bei der Interpolation von der Größenordnung $O(h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^{d-1})$ ist. In diesem Sinne kann unser Verfahren auch als Verfahren höherer Ordnung angesehen werden, da es mit $O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^{d-1})$ Unbekannten zu einem Fehler bezüglich der $L_2$- beziehungsweise Maximumsnorm der Ordnung $O(h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^{d-1})$ führt, wozu auf vollen Gittern bereits ein Verfahren der Ordnung $O(h_n^{2d})$ erforderlich ist. Als Maß für die (algebraische) Konvergenzgeschwindigkeit unseres Verfahrens geben wir eine Schranke der Ordnung $O(1 - h_n^{5/4})$ für den Spektralradius der Dünn-Gitter-Iterationsmatrix an. Dieser vermeintliche Nachteil einer langsamen Konvergenz unserer Methode kann jedoch durch den Übergang vom reinen Gauß-Seidel-Löser zu einem Mehrgitter-V-Zyklus mit der dort bekannten, von der Anzahl der Variablen unabhängigen Konvergenzrate behoben werden (siehe [16], [17]).

In Abschnitt 7 erweitern wir unser Verfahren auf dreidimensionale Poisson-Probleme auf dem Einheitskubus.

Abschnitt 8 stellt dann zwei Varianten des zweidimensionalen Algorithmus' vor. Die erste ermöglicht die Lösung der Poisson-Gleichung mit einem nochmals reduzierten Speicheraufwand. Anstelle der in der ursprünglichen Form erforderlichen sechs Variablen benötigt sie nur zwei Variablen pro Gitterpunkt. Zusammen mit den Zeigern zur Realisierung der Binärbaum-Datenstruktur ergibt sich bei Rechnung in doppelter Genauigkeit ein Gesamt-Speicheraufwand von 40 Byte pro Gitterpunkt. Eine zweite Modifikation führt schließlich bezüglich der $L_2$- und Maximumsnorm in numerischen Experimenten zu Näherungslösungen höherer Genauigkeit, nämlich $O(h_n^3 \cdot \mathrm{ld}(h_n^{-1}))$ beziehungsweise $O(h_n^4 \cdot \mathrm{ld}(h_n^{-1}))$ anstelle von $O(h_n^2 \cdot \mathrm{ld}(h_n^{-1}))$. Somit haben wir ein Verfahren, das — bei Vernachlässigung der logarithmischen Faktoren — Fehler der Ordnung $O(h_n^4)$ mit $O(h_n^{-1})$ Gitterpunkten erzielt. Dies entspricht auf vollen Gittern bereits einem Verfahren der Ordnung $O(h_n^8)$.

Schließlich präsentieren wir in Abschnitt 9 numerische Ergebnisse zu den vorgestellten Techniken. Alle numerischen Berechnungen in dieser Arbeit wurden am Institut für

Informatik der Technischen Universität München auf einer mit 64 MByte Hauptspeicher ausgestatteten DECstation 5000/200 von Digital Equipment durchgeführt. Sämtliche Programme wurden in C geschrieben.

An dieser Stelle möchte ich mich bei der Abteilung DV 182 im Unternehmensbereich Kommunikations- und Datentechnik der Firma Siemens (jetzt Siemens Nixdorf), namentlich bei Herrn Dr. Horst Gietl, für die mir während der Anfertigung dieser Arbeit zugekommene materielle Unterstützung bedanken.

Dank schulde ich ferner meinen Kollegen Dr. Ulrich Rüde und Dr. Michael Griebel für die gewissenhafte Durchsicht des Manuskripts dieser Arbeit und für zahlreiche sich daraus ergebende Hinweise sowie Verbesserungsvorschläge. Bei Michael Griebel möchte ich mich überdies für viele hilfreiche Anregungen sowie den stets fruchtbaren Gedankenaustausch bedanken, von dem ich hoffe, daß er nicht nur für mich von großem Nutzen war und ist. Weiterhin bedanke ich mich herzlich bei meiner Schwester Christine, die mich vor manchem sprachlichen Lapsus bewahrt hat.

Mein besonderer und herzlicher Dank gilt hier freilich meinem Doktorvater Prof. Dr. Christoph Zenger für die Überlassung des Themas sowie die intensive Betreuung meiner Arbeit. Sein Dünn-Gitter-Enthusiasmus, seine geduldigen Erklärungen und nicht zuletzt sein Ideenreichtum haben mir bei der Erstellung dieser Arbeit in ganz entscheidendem Maße geholfen.

## 2. Hierarchische Basen zur Darstellung glatter Funktionen.

**2.1. Hierarchische-Basis-Zerlegung des Voll-Gitter-Raumes $\mathbf{S_n^{(d)}}$.** Der erste Teil der vorliegenden Arbeit befaßt sich mit dünnen Gittern im Hinblick auf die Möglichkeit effizienter Speicherung und Darstellung glatter, auf $\bar{\Omega}^{(d)} := [0,1]^d$ definierter Funktionen. Dabei bedarf zunächst der Begriff der Glattheit einer Präzisierung. Wir wollen im folgenden unter einer glatten Funktion von $d$ Veränderlichen ($d = 1, 2, 3, ...$) eine Funktion $u^{(d)}$ verstehen, deren gemischte Ableitung $\frac{\partial^{2d} u^{(d)}}{\partial x_1^2 ... \partial x_d^2}$ auf $\bar{\Omega}^{(d)}$ existiert und stetig ist. Die Gesamtheit aller in diesem Sinne glatten Funktionen bezeichnen wir mit $X^{(d)}$. Die Menge derjenigen Funktionen $u^{(d)} \in X^{(d)}$, die auf dem Rand $\delta\bar{\Omega}^{(d)}$ verschwinden, heiße $X_0^{(d)}$:

$$
\begin{aligned}
X^{(d)} &:= \left\{ u^{(d)} : \ \bar{\Omega}^{(d)} \subset D^{(d)} \to \mathbb{R}, \ \frac{\partial^{2d} u^{(d)}}{\partial x_1^2 \dots \partial x_d^2} \ \in \mathcal{C}^0(\bar{\Omega}^{(d)}) \right\}, \\
X_0^{(d)} &:= \left\{ u^{(d)} \in X^{(d)} : \ u^{(d)}|_{\delta\bar{\Omega}^{(d)}} = 0 \right\}.
\end{aligned} \tag{1}
$$

Soll nun eine auf $\bar{\Omega}^{(d)}$ definierte Funktion $u^{(d)}$ im Rechner gespeichert werden, so führt man üblicherweise auf $\bar{\Omega}^{(d)}$ ein Gitter ein. Gespeichert werden dann lediglich die Werte von $u^{(d)}$ an den Gitterpunkten, und zur Gewinnung der Funktionswerte an Zwischenstellen können nun verschiedene Interpolationstechniken eingesetzt werden. Als einfaches und doch zumeist ausreichendes Interpolationsverfahren bietet sich hier im Falle einer Unbekannten die lineare beziehungsweise — in der Verallgemeinerung auf $d$ Unabhängige — die $d$-lineare Interpolation an, bei der in jeder Koordinatenrichtung linear interpoliert wird. Man erhält so einen stückweise $d$-linearen Interpolanten von $u^{(d)}$, der längs der Gitterlinien zwischen benachbarten Gitterpunkten linear verläuft.

Sei nun $G_{n_1,...,n_d}$ ein äquidistantes, rechtwinkliges Gitter auf $\bar{\Omega}^{(d)}$ mit zugehörigen Maschenweiten $h_{n_j} = 2^{-n_j}, \ j = 1, ..., d$. Wir nennen $G_{i_1,...,i_d}$ Verfeinerung von $G_{k_1,...,k_d}$ oder kurz feiner als $G_{k_1,...,k_d}$, wenn gilt: $k_j \leq i_j \ \forall j, \ \sum_{j=1}^d k_j < \sum_{j=1}^d i_j$. Damit erhalten wir auf natürliche Weise eine Hierarchie von Gittern: Ist $G_{i_1,...,i_d}$ Verfeinerung von $G_{k_1,...,k_d}$, so nennen wir $G_{k_1,...,k_d}$ *hierarchisch höher* als $G_{i_1,...,i_d}$. Insbesondere nimmt in dieser Sprechweise das gröbste Gitter $G_{1,...,1}$ allein die höchste Hierarchiestufe ein.

Ausgehend von dieser Definition von Gittern bezeichnen wir mit $S_{n_1,...,n_d}$ den $(2^{n_1} - 1) \cdot ... \cdot (2^{n_d} - 1)$-dimensionalen Raum aller auf $G_{n_1,...,n_d}$ stückweise $d$-linearen Funktionen, die auf dem Rand $\delta\bar{\Omega}^{(d)}$ verschwinden. Letztere Eigenschaft stellt keine wesentliche Einschränkung dar. Sie soll vielmehr der Übersichtlichkeit der Darstellung dienen. Ferner führen wir Unterräume $T_{i_1,...,i_d}$ von $S_{n_1,...,n_d}$ ein ($1 \leq i_j \leq n_j, \ 1 \leq j \leq d$). $T_{i_1,...,i_d}$ enthält dabei alle auf $G_{i_1,...,i_d}$ stückweise $d$-linearen Funktionen, die an den Gitterpunkten sämtlicher gröberer Gitter $G_{k_1,...,k_d}$ verschwinden, wodurch die Hierarchie von Gittern auf natürliche Weise eine Hierarchie von Teilräumen induziert. Wir definieren auf der Menge der Teilräume $T_{i_1,...,i_d}$ eine irreflexive und transitive Ordnungsrelation (partielle Ordnung)

$$
T_{k_1,...,k_d} < T_{i_1,...,i_d} \quad :\Leftrightarrow \quad \left( \forall j : \ k_j \leq i_j \right) \ \wedge \ \left( \sum_{j=1}^d k_j < \sum_{j=1}^d i_j \right) \tag{2}
$$

und nennen in diesem Fall $T_{k_1,\ldots,k_d}$ *hierarchisch höher* als $T_{i_1,\ldots,i_d}$.

Mit Hilfe der Teilräume $T_{i_1,\ldots,i_d}$ können wir $S_{n_1,\ldots,n_d}$ nun schreiben als

$$S_{n_1,\ldots,n_d} \;=\; \bigoplus_{i_1=1}^{n_1} \ldots \bigoplus_{i_d=1}^{n_d} T_{i_1,\ldots,i_d}. \tag{3}$$

Im Spezialfall $n := n_1 = \ldots = n_d$ schreiben wir kurz

$$G_n^{(d)} := G_{n,\ldots,n}, \qquad S_n^{(d)} := S_{n,\ldots,n}. \tag{4}$$

Abbildung 1 veranschaulicht die Zerlegung von $S_{3,3}$ in Teilräume $T_{i,j}$ im Fall $d = 2$. (3) stellt nun eine Hierarchische-Basis-Zerlegung des Raumes $S_{n_1,\ldots,n_d}$ dar, wobei in den einzelnen Teilräumen $T_{i_1,\ldots,i_d}$ stückweise $d$-lineare Finite-Elemente-Basen verwendet werden. Diese sollen im folgenden genauer betrachtet werden.

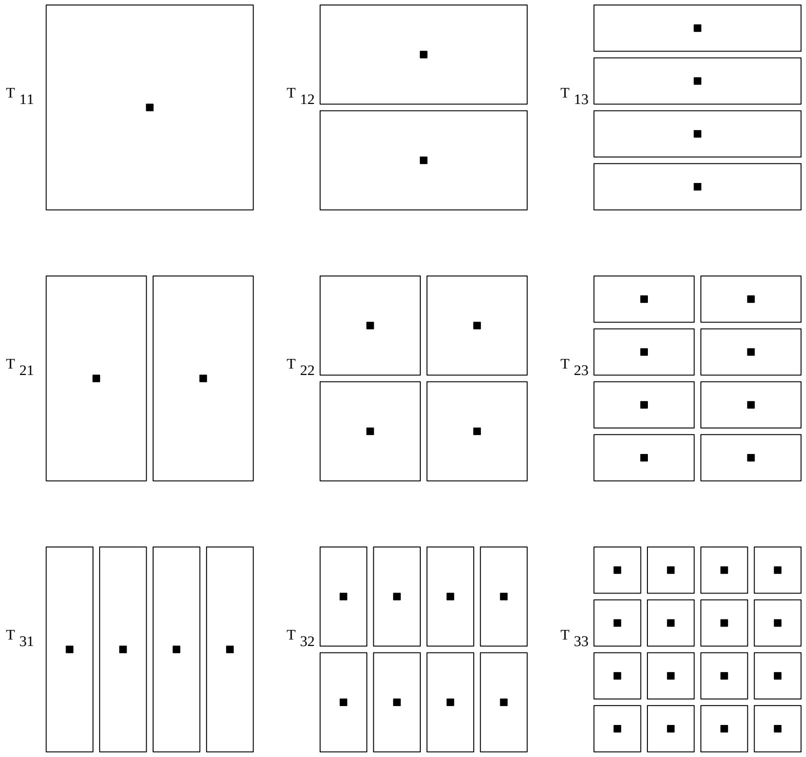



ABB. 1. $d = 2$*: Der Raum* $S_{3,3}$ *und seine Teilräume.*

Sei $T_{i_1,\ldots,i_d}$, $i_1 \le n, \ldots, i_d \le n$, ein Teilraum des Funktionenraumes $S_n^{(d)}$ gemäß obiger Festlegung. Da $T_{i_1,\ldots,i_d}$ per definitionem auch Teilraum von $S_{i_1,\ldots,i_d}$ ist und deshalb auf dem Gitter $G_{i_1,\ldots,i_d}$ dargestellt werden kann und da sämtliche Funktionen aus $T_{i_1,\ldots,i_d}$ auf allen in einer beliebigen Koordinatenrichtung gegenüber $G_{i_1,\ldots,i_d}$ vergröberten Gittern verschwinden, folgt für die Dimension von $T_{i_1,\ldots,i_d}$

$$dim(T_{i_1,\ldots,i_d}) \;=\; 2^{i_1-1} \cdot \ldots \cdot 2^{i_d-1}. \tag{5}$$

Mit Hilfe der Funktionen $w_l$, $l \in \mathbb{N}$,

$$w_l(x) \ := \ \begin{cases} \frac{h_l+x}{h_l}\ , & \text{falls} \quad -h_l \leq x \leq 0, \\ \frac{h_l-x}{h_l}\ , & \text{falls} \quad 0 \leq x \leq h_l, \\ 0 & \text{sonst,} \end{cases}$$

können wir nun die stückweise $d$-linearen Basisfunktionen des betrachteten Teilraumes $T_{i_1,...,i_d}$ definieren:

Definition 2.1. *Als Basisfunktionen des Teilraumes $T_{i_1,...,i_d}$ von $S_n^{(d)}$ definieren wir*

$$(6) \qquad g_{k_1,...,k_d}^{(i_1,...,i_d)}(x_1,...,x_d) \ := \ \prod_{j=1}^{d} w_{i_j}(x_j - m_{k_j}^{(i_j)}), \qquad 1 \leq k_j \leq 2^{i_j-1},$$

*wobei* $m_{k_j}^{(i_j)} := (2k_j - 1) \cdot h_{i_j}$.
Somit ist das kartesische Produkt von Intervallen

$$\prod_{j=1}^{d} \Big[ (2k_j - 2) \cdot h_{i_j}, 2k_j \cdot h_{i_j} \Big]$$

der Träger der Basisfunktion $g_{k_1,...,k_d}^{(i_1,...,i_d)}$. Der Gitterpunkt $(m_{k_1}^{(i_1)}, ..., m_{k_d}^{(i_d)})$ ist der Mittelpunkt des Trägers. Wie man leicht sieht, sind die Träger der einzelnen Basisfunktionen von $T_{i_1,...,i_d}$ gleich groß. Ferner ist der Durchschnitt der Träger zweier beliebiger verschiedener Basisfunktionen von $T_{i_1,...,i_d}$ entweder leer oder eine Nullmenge im $d$-dimensionalen Raum. Wir werden deshalb fortan zwei solche Träger vereinfachend als disjunkt bezeichnen.

Abbildung 2 zeigt die beiden Basisfunktionen des Teilraumes $T_2$ mit ihrem jeweiligen Träger im Spezialfall $d = 1$. Abbildung 3 veranschaulicht die Situation im Fall $d = 2$. Hier werden die vier Basisfunktionen des Teilraumes $T_{2,2}$ angegeben.

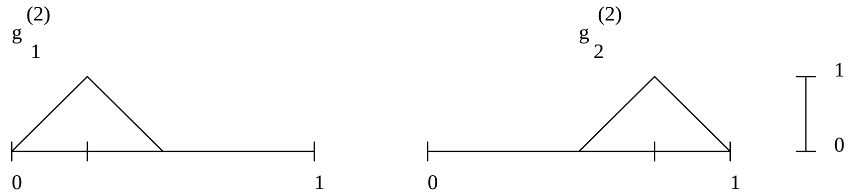



Abb. 2. $d = 1$*: Basisfunktionen des Teilraumes* $T_2$.

**2.2. Der Interpolationsfehler im Raum $\mathbf{S_n^{(d)}}$.** Wir wollen nun untersuchen, mit welcher Genauigkeit der Raum $S_n^{(d)}$ und damit das zugrundeliegende äquidistante, rechtwinklige Gitter $G_n^{(d)}$ eine Funktion $u^{(d)} \in X_0^{(d)}$ interpolieren beziehungsweise darstellen kann. Dazu schreiben wir zunächst den Interpolanten $u_{n,I}^{(d)} \in S_n^{(d)}$ von $u^{(d)} \in X_0^{(d)}$ gemäß (3) sowie $u^{(d)}$ selbst als Summen von Funktionen $u_{i_1,...,i_d} \in T_{i_1,...,i_d}$:

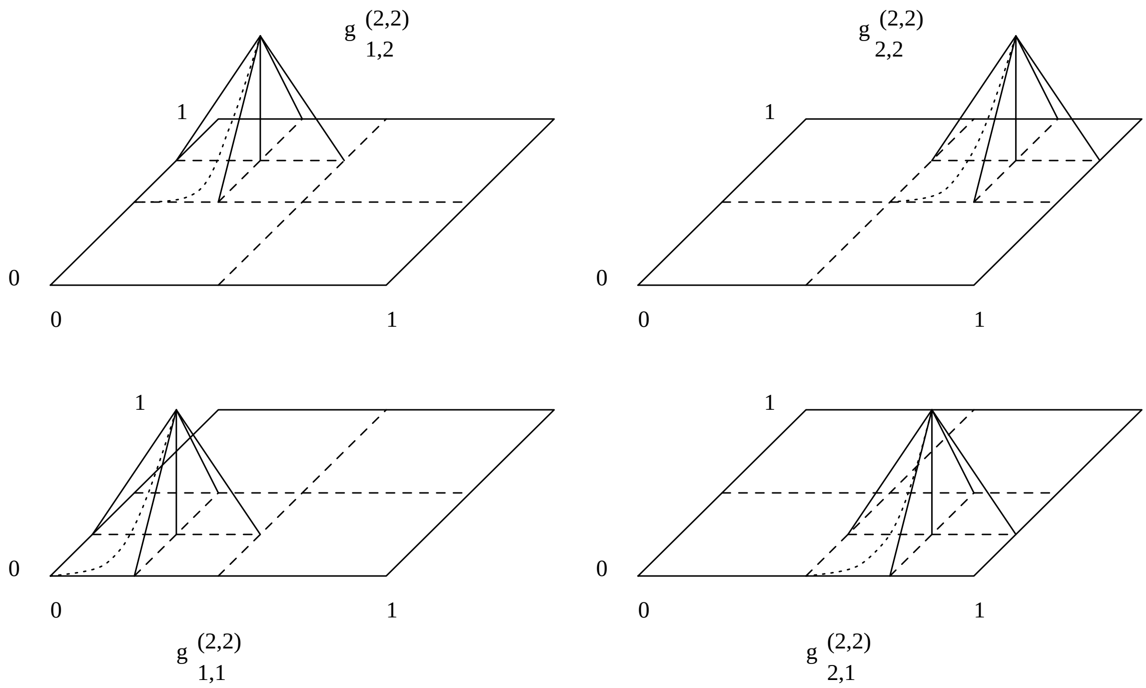



ABB. 3. $d = 2$: *Basisfunktionen des Teilraumes* $T_{2,2}$.

$$(7)\qquad \begin{aligned} u_{n,I}^{(d)} &= \sum_{i_1=1}^{n} \cdots \sum_{i_d=1}^{n} u_{i_1,\ldots,i_d}, \\ u^{(d)} &= \lim_{n\to\infty} u_{n,I}^{(d)} = \sum_{i_1=1}^{\infty} \cdots \sum_{i_d=1}^{\infty} u_{i_1,\ldots,i_d}. \end{aligned}$$

Jede der Funktionen $u_{i_1,\ldots,i_d} \in T_{i_1,\ldots,i_d}$ ist dabei ihrerseits wieder eine Linearkombination der Basisfunktionen $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$ des Teilraumes $T_{i_1,\ldots,i_d}$:

$$(8)\qquad u_{i_1,\ldots,i_d} = \sum_{k_1=1}^{2^{i_1-1}} \cdots \sum_{k_d=1}^{2^{i_d-1}} c_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \cdot g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}.$$

Die Grenzwertbildung in (7) ist zulässig, da $S_i^{(d)} \subset S_{i+1}^{(d)}$ gilt und weil $\bigcup_{i=1}^{\infty} S_i^{(d)}$ in $X_0^{(d)}$ dicht liegt. Im folgenden sollen nun Schranken für die Maximums- und die $L_2$-Norm des Interpolationsfehlers $u^{(d)} - u_{n,I}^{(d)}$ bestimmt werden. Die Ordnung $O(h_n^2)$ solcher Fehlernormschranken ist aus der Finite-Elemente-Theorie bekannt. Wir wollen hier jedoch unter Ausnutzung unserer Glattheitsforderungen eine andere Beweistechnik als die dort gebräuchliche verwenden — vor allem im Hinblick auf entsprechende Beweise für dünne Gitter. Zu diesem Zweck führen wir auf $X_0^{(d)}$ zunächst zwei Halbnormen ein und zeigen drei vorbereitende Lemmata:

DEFINITION 2.2. *Sei* $u^{(d)} \in X_0^{(d)}$. *Dann werden die beiden Halbnormen* $|u^{(d)}|_\infty$ *und* $|u^{(d)}|_2$ *wie folgt definiert:*

$$
(9) \qquad
\begin{aligned}
|u^{(d)}|_\infty &:= \left\| \frac{\partial^{2d} u^{(d)}}{\partial x_1^2 \, ... \, \partial x_d^2} \right\|_\infty , \\
|u^{(d)}|_2 &:= \left\| \frac{\partial^{2d} u^{(d)}}{\partial x_1^2 \, ... \, \partial x_d^2} \right\|_2 = \left( \int_0^1 ... \int_0^1 \left| \frac{\partial^{2d} u^{(d)}}{\partial x_1^2 \ldots \partial x_d^2} \right|^2 \, dx_d \, ... \, dx_1 \right)^{\frac{1}{2}} .
\end{aligned}
$$

LEMMA 2.3. *Für jede Basisfunktion $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \in T_{i_1,\ldots,i_d}$ gilt:*

$$
(10) \qquad
\begin{aligned}
\left\| g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \right\|_\infty &= 1, \\
\left\| g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \right\|_2 &= \left( \frac{2}{3} \right)^{\frac{d}{2}} \cdot h_{i_1}^{\frac{1}{2}} \cdot \ldots \cdot h_{i_d}^{\frac{1}{2}} .
\end{aligned}
$$

*Beweis.* Die Aussage für die Maximumsnorm folgt unmittelbar aus der Definition der Basisfunktionen $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$ (siehe (6)). Ferner gilt

$$
\begin{aligned}
\left\| g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \right\|_2 &= \left( \int_{m_{k_1}^{(i_1)} - h_{i_1}}^{m_{k_1}^{(i_1)} + h_{i_1}} ... \int_{m_{k_d}^{(i_d)} - h_{i_d}}^{m_{k_d}^{(i_d)} + h_{i_d}} \prod_{j=1}^{d} w_{i_j}^2 (x_j - m_{k_j}^{(i_j)}) \, dx_d \, ... \, dx_1 \right)^{\frac{1}{2}} \\
&= \prod_{j=1}^{d} \left( \int_{m_{k_j}^{(i_j)} - h_{i_j}}^{m_{k_j}^{(i_j)} + h_{i_j}} w_{i_j}^2 (x_j - m_{k_j}^{(i_j)}) \, dx_j \right)^{\frac{1}{2}} \\
&= \left( \frac{2}{3} \right)^{\frac{d}{2}} \cdot h_{i_1}^{\frac{1}{2}} \cdot \ldots \cdot h_{i_d}^{\frac{1}{2}} .
\end{aligned}
$$

□

Um nun daraus Aussagen über $\|u_{i_1,\ldots,i_d}\|_\infty$ beziehungsweise $\|u_{i_1,\ldots,i_d}\|_2$ gewinnen zu können, müssen die Beiträge der einzelnen Basisfunktionen $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$ zu $u_{i_1,\ldots,i_d}$, d.h. die Faktoren $c_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$ in (8), bestimmt werden. Hierzu ist es wichtig, sich die Bedeutung dieser Faktoren klarzumachen. $c_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$ gibt im zugehörigen Gitterpunkt — dem Mittelpunkt des Trägers von $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$ — gerade an, um welchen Betrag der aus Nachbarwerten gebildete interpolierte Wert noch modifiziert werden muß , um zum tatsächlichen Funktionswert zu gelangen. Im eindimensionalen Fall ($d = 1$, lineare Interpolation) gilt also

$$
c_{k_1}^{(i_1)} = u^{(1)}(m_{k_1}^{(i_1)}) - \frac{u^{(1)}(m_{k_1}^{(i_1)} - h_{i_1}) + u^{(1)}(m_{k_1}^{(i_1)} + h_{i_1})}{2}
$$

beziehungsweise in Operatorschreibweise

$$
c_{k_1}^{(i_1)} = \left[ \begin{array}{ccc} -\frac{1}{2} & 1 & -\frac{1}{2} \end{array} \right]_{m_{k_1}^{(i_1)}, h_{i_1}} u^{(1)} =: I_{m_{k_1}^{(i_1)}, h_{i_1}} u^{(1)} .
$$

Im zweidimensionalen Fall ($d = 2$, bilineare Interpolation) gilt

$$
c_{k_1,k_2}^{(i_1,i_2)} = \left( \left[ \begin{array}{c} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{array} \right]_{m_{k_2}^{(i_2)}, h_{i_2}} \cdot \left[ \begin{array}{ccc} -\frac{1}{2} & 1 & -\frac{1}{2} \end{array} \right]_{m_{k_1}^{(i_1)}, h_{i_1}} \right) u^{(2)}
$$

$$= \left[ \begin{array}{ccc} \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \end{array} \right]_{m_{k_1}^{(i_1)}, h_{i_1}, m_{k_2}^{(i_2)}, h_{i_2}} \quad u^{(2)}.$$

Im allgemeinen Fall $d > 2$ erhalten wir mit einer vereinfachenden Produkt-Schreibweise

$$(11) \qquad c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)} = \left( \prod_{j=1}^{d} \left[ \begin{array}{ccc} -\frac{1}{2} & 1 & -\frac{1}{2} \end{array} \right]_{m_{k_j}^{(i_j)}, h_{i_j}} \right) u^{(d)} = \left( \prod_{j=1}^{d} I_{m_{k_j}^{(i_j)}, h_{i_j}} \right) u^{(d)}.$$

Aufgrund des beschriebenen Zusammenhangs nennen wir $c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}$ auch *hierarchischen Überschuß* im Gitterpunkt $(m_{k_1}^{(i_1)}, \dots, m_{k_d}^{(i_d)})$.

Lemma 2.4. *Gegeben sei $u^{(d)} \in X_0^{(d)}$ in hierarchischer Darstellung (vgl. (7) und (8)). Dann gelten für die hierarchischen Überschüsse $c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}$ folgende Abschätzungen:*

$$(12) \qquad \begin{aligned} |c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}| &\leq \frac{|u^{(d)}|_\infty}{2^d} \cdot h_{i_1}^2 \cdot \dots \cdot h_{i_d}^2, \\ |c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}| &\leq \frac{|u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|_2}{6^{\frac{d}{2}}} \cdot h_{i_1}^{\frac{3}{2}} \cdot \dots \cdot h_{i_d}^{\frac{3}{2}}, \end{aligned}$$

*wobei $\varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}$ die charakteristische Funktion des Trägers von $g_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}$ ist. Auf dem Träger von $g_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}$ nimmt $\varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}$ also den Wert 1 an, außerhalb den Wert 0.*

*Beweis.* O.B.d.A. sei $m_{k_j}^{(i_j)} = 0 \; \forall j$. Definiere

$$v_{i_j}(x) := \frac{-h_{i_j}}{2} \cdot w_{i_j}(x), \quad -h_{i_j} \leq x \leq h_{i_j},$$

und betrachte folgendes Integral $J$:

$$\begin{aligned} J &:= \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_d}}^{+h_{i_d}} \Big( \prod_{j=1}^{d} v_{i_j}(x_j) \Big) \cdot \frac{\partial^{2d} u^{(d)}(x_1, \dots, x_d)}{\partial x_1^2 \dots \partial x_d^2} \, dx_d \dots dx_1 \\ &= (-1)^d \cdot \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_d}}^{+h_{i_d}} \Big( \prod_{j=1}^{d} v'_{i_j}(x_j) \Big) \cdot \frac{\partial^{d} u^{(d)}(x_1, \dots, x_d)}{\partial x_1 \dots \partial x_d} \, dx_d \dots dx_1 \\ &\qquad \text{(partielle Integration)} \\ &= (-1)^{d-1} \cdot \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_{d-1}}}^{+h_{i_{d-1}}} \Big( \prod_{j=1}^{d-1} v'_{i_j}(x_j) \Big) \cdot \\ &\qquad \cdot \left( \int_{-h_{i_d}}^{+h_{i_d}} -v'_{i_d}(x_d) \cdot \frac{\partial^{d} u^{(d)}(x_1, \dots, x_d)}{\partial x_1 \dots \partial x_d} \, dx_d \right) dx_{d-1} \dots dx_1 \\ &= (-1)^{d-1} \cdot \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_{d-1}}}^{+h_{i_{d-1}}} \Big( \prod_{j=1}^{d-1} v'_{i_j}(x_j) \Big) \cdot \\ &\qquad \cdot I_{0,h_{i_d}} \frac{\partial^{d-1} u^{(d)}(x_1, \dots, x_{d-1}, \cdot)}{\partial x_1 \dots \partial x_{d-1}} \, dx_{d-1} \dots dx_1 \end{aligned}$$

$$
\begin{aligned}
&= (-1)^{d-2} \cdot \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_{d-2}}}^{+h_{i_{d-2}}} \Big(\prod_{j=1}^{d-2} v'_{i_j}(x_j)\Big) \cdot \\
&\qquad \cdot \Big(I_{0,h_{i_{d-1}}} I_{0,h_{i_d}}\Big) \; \frac{\partial^{d-2} u^{(d)}(x_1, \dots, x_{d-2}, \cdot, \cdot)}{\partial x_1 \dots \partial x_{d-2}} \; dx_{d-2} \dots dx_1 \\
&= \dots \\
&= \left(\prod_{j=1}^{d} I_{0,h_{i_j}}\right) \; u^{(d)}.
\end{aligned}
$$

Andererseits gilt

$$
\begin{aligned}
|J| &\leq \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_d}}^{+h_{i_d}} |u^{(d)}|_\infty \cdot \Big(\prod_{j=1}^{d} |v_{i_j}(x_j)|\Big) \; dx_d \dots dx_1 \\
&= |u^{(d)}|_\infty \cdot \left(\prod_{j=1}^{d} \int_{-h_{i_j}}^{+h_{i_j}} |v_{i_j}(x_j)| \; dx_j\right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot h_{i_1}^2 \cdot \dots \cdot h_{i_d}^2,
\end{aligned}
$$

woraus mit (11) die erste Behauptung des Lemmas folgt.

Zum Beweis der $L_2$-Schranke benutzen wir die Cauchy-Schwarz'sche Ungleichung und erhalten

$$
\begin{aligned}
|J| &= \left| \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_d}}^{+h_{i_d}} \Big(\prod_{j=1}^{d} v_{i_j}(x_j)\Big) \cdot \frac{\partial^{2d} u^{(d)}(x_1, \dots, x_d)}{\partial x_1^2 \dots \partial x_d^2} \; dx_d \dots dx_1 \right| \\
&= \left| \int_{-h_{i_1}}^{+h_{i_1}} \dots \int_{-h_{i_d}}^{+h_{i_d}} \Big(\prod_{j=1}^{d} v_{i_j}(x_j)\Big) \cdot \frac{\partial^{2d} (u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)})(x_1, \dots, x_d)}{\partial x_1^2 \dots \partial x_d^2} \; dx_d \dots dx_1 \right| \\
&\leq \left\| \prod_{j=1}^{d} v_{i_j}(x_j) \right\|_2 \cdot \left\| \frac{\partial^{2d} (u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)})(x_1, \dots, x_d)}{\partial x_1^2 \dots \partial x_d^2} \right\|_2 \\
&= |u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|_2 \cdot \left(\prod_{j=1}^{d} \int_{-h_{i_j}}^{+h_{i_j}} v_{i_j}^2(x_j) \; dx_j\right)^{\frac{1}{2}} \\
&= |u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|_2 \cdot \left(\prod_{j=1}^{d} \frac{h_{i_j}^3}{6}\right)^{\frac{1}{2}} \\
&= \frac{|u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|_2}{6^{\frac{d}{2}}} \cdot \left(\prod_{j=1}^{d} h_{i_j}^{\frac{3}{2}}\right),
\end{aligned}
$$

woraus — wiederum mit (11) — die zweite Behauptung des Lemmas folgt. □

(11) zeigt, daß der hierarchische Überschuß bis auf einen Faktor $(-1)^d \cdot \frac{h_{i_1}^2 \cdot \dots \cdot h_{i_d}^2}{2^d}$ gerade der üblichen Differenzenquotienten-Approximation von $\frac{\partial^{2d} u^{(d)}}{\partial x_1^2 \dots \partial x_d^2}$ entspricht. Man beachte, daß dieser Zusammenhang auch in der ersten Abschätzung im soeben bewiesenen Lemma 2.4 sichtbar wird. Wichtig ist außerdem, daß die Schranken für die Faktoren $c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}$ nicht beziehungsweise nur über die charakteristische Funktion des Trägers von

den Indizes $k_j$ abhängen. Ausschlaggebend für die Größenordnung der angegebenen Schranken ist also nur der Teilraum $T_{i_1,\dots,i_d}$, nicht jedoch die konkrete Basisfunktion.

LEMMA 2.5. *Gegeben sei wiederum $u^{(d)} \in X_0^{(d)}$ in hierarchischer Darstellung (7). Dann gelten für jedes $u_{i_1,\dots,i_d} \in T_{i_1,\dots,i_d}$,*

$$u_{i_1,\dots,i_d} = \sum_{k_1=1}^{2^{i_1-1}} \dots \sum_{k_d=1}^{2^{i_d-1}} c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)} \cdot g_{k_1,\dots,k_d}^{(i_1,\dots,i_d)},$$

*folgende Abschätzungen:*

$$\begin{aligned} \|u_{i_1,\dots,i_d}\|_\infty &\le \frac{|u^{(d)}|_\infty}{2^d} \cdot h_{i_1}^2 \cdot \dots \cdot h_{i_d}^2, \\ \|u_{i_1,\dots,i_d}\|_2 &\le \frac{|u^{(d)}|_2}{3^d} \cdot h_{i_1}^2 \cdot \dots \cdot h_{i_d}^2. \end{aligned} \tag{13}$$

*Beweis.* Aufgrund der Disjunktheit der Träger der Basisfunktionen von $T_{i_1,\dots,i_d}$ gilt

$$\begin{aligned} \|u_{i_1,\dots,i_d}\|_\infty &= \Big\| \sum_{k_1=1}^{2^{i_1-1}} \dots \sum_{k_d=1}^{2^{i_d-1}} c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)} \cdot g_{k_1,\dots,k_d}^{(i_1,\dots,i_d)} \Big\|_\infty \\ &= \max_{k_1,\dots,k_d} \Big( |c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}| \cdot \|g_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}\|_\infty \Big), \end{aligned}$$

woraus sich mit Lemma 2.3 und Lemma 2.4 sofort die erste Ungleichung ergibt.

Zum Beweis der zweiten Ungleichung liefern wiederum das Disjunktheitsargument sowie Lemma 2.3

$$\begin{aligned} \|u_{i_1,\dots,i_d}\|_2^2 &= \Big\| \sum_{k_1=1}^{2^{i_1-1}} \dots \sum_{k_d=1}^{2^{i_d-1}} c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)} \cdot g_{k_1,\dots,k_d}^{(i_1,\dots,i_d)} \Big\|_2^2 \\ &= \sum_{k_1=1}^{2^{i_1-1}} \dots \sum_{k_d=1}^{2^{i_d-1}} |c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|^2 \cdot \|g_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}\|_2^2 \\ &= \sum_{k_1=1}^{2^{i_1-1}} \dots \sum_{k_d=1}^{2^{i_d-1}} |c_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|^2 \cdot \left(\frac{2}{3}\right)^d \cdot h_{i_1} \cdot \dots \cdot h_{i_d}. \end{aligned}$$

Aus Lemma 2.4 sowie den disjunkten Trägern folgt schließlich

$$\begin{aligned} \|u_{i_1,\dots,i_d}\|_2^2 &\le \sum_{k_1=1}^{2^{i_1-1}} \dots \sum_{k_d=1}^{2^{i_d-1}} \frac{|u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|_2^2}{6^d} \cdot h_{i_1}^3 \cdot \dots \cdot h_{i_d}^3 \cdot \left(\frac{2}{3}\right)^d \cdot h_{i_1} \cdot \dots \cdot h_{i_d} \\ &= \frac{1}{9^d} \cdot h_{i_1}^4 \cdot \dots \cdot h_{i_d}^4 \cdot \sum_{k_1=1}^{2^{i_1-1}} \dots \sum_{k_d=1}^{2^{i_d-1}} |u^{(d)} \cdot \varphi_{k_1,\dots,k_d}^{(i_1,\dots,i_d)}|_2^2 \\ &= \frac{1}{9^d} \cdot h_{i_1}^4 \cdot \dots \cdot h_{i_d}^4 \cdot |u^{(d)}|_2^2. \end{aligned}$$

□

Damit sind wir jetzt in der Lage, sowohl Schranken für die Maximums- und die $L_2$-Norm des Interpolationsfehlers $u^{(d)} - u_{n,I}^{(d)}$ als auch deren Größenordnung anzugeben.

SATZ 2.6. *Gegeben seien $u^{(d)} \in X_0^{(d)}$ sowie sein stückweise $d$-linearer Interpolant $u_{n,I}^{(d)} \in S_n^{(d)}$ in hierarchischer Darstellung (7). Dann gelten folgende Ungleichungen:*

$$
\begin{aligned}
\|u^{(d)} - u_{n,I}^{(d)}\|_\infty &\leq \frac{d \cdot |u^{(d)}|_\infty}{6^d} \cdot h_n^2 &= O(h_n^2), \\
\|u^{(d)} - u_{n,I}^{(d)}\|_2 &\leq \frac{d \cdot |u^{(d)}|_2}{9^d} \cdot h_n^2 &= O(h_n^2).
\end{aligned} \tag{14}
$$

*Beweis.* Nach (7) gilt

$$
\begin{aligned}
\|u^{(d)} - u_{n,I}^{(d)}\|_\infty &= \Big\| \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} u_{i_1,\dots,i_d} - \sum_{i_1=1}^{n} \dots \sum_{i_d=1}^{n} u_{i_1,\dots,i_d} \Big\|_\infty \\
&\leq \sum_{(i_1,\dots,i_d):\exists i_j > n} \|u_{i_1,\dots,i_d}\|_\infty \\
&\leq \sum_{(i_1,\dots,i_d):\exists i_j > n} \frac{|u^{(d)}|_\infty}{2^d} \cdot h_{i_1}^2 \cdot \dots \cdot h_{i_d}^2 \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} 4^{-i_1} \cdot \dots \cdot 4^{-i_d} - \sum_{i_1=1}^{n} \dots \sum_{i_d=1}^{n} 4^{-i_1} \cdot \dots \cdot 4^{-i_d} \right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \Big( \sum_{i_1=1}^{\infty} 4^{-i_1} \Big)^d - \Big( \sum_{i_1=1}^{n} 4^{-i_1} \Big)^d \right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \Big(\frac{1}{3}\Big)^d - \Big(\frac{1}{3}\Big)^d \cdot \big(1 - 4^{-n}\big)^d \right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \Big(\frac{1}{3}\Big)^d \cdot \big(1 - (1 - 4^{-n})^d\big) \\
&\leq \frac{|u^{(d)}|_\infty}{2^d} \cdot \frac{d \cdot 4^{-n}}{3^d} \\
&= \frac{d \cdot |u^{(d)}|_\infty}{6^d} \cdot h_n^2 \\
&= O(h_n^2).
\end{aligned}
$$

Analog erhält man

$$
\begin{aligned}
\|u^{(d)} - u_{n,I}^{(d)}\|_2 &= \Big\| \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} u_{i_1,\dots,i_d} - \sum_{i_1=1}^{n} \dots \sum_{i_d=1}^{n} u_{i_1,\dots,i_d} \Big\|_2 \\
&\leq \frac{|u^{(d)}|_2}{3^d} \cdot \left( \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} 4^{-i_1} \cdot \dots \cdot 4^{-i_d} - \sum_{i_1=1}^{n} \dots \sum_{i_d=1}^{n} 4^{-i_1} \cdot \dots \cdot 4^{-i_d} \right) \\
&\leq \frac{|u^{(d)}|_2}{3^d} \cdot \frac{d \cdot 4^{-n}}{3^d}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{d \cdot |u^{(d)}|_2}{9^d} \cdot h_n^2 \\
&= O(h_n^2).
\end{aligned}
$$

□

Die Ordnung $O(h_n^2)$ der Maximums- und der $L_2$-Norm des Interpolationsfehlers in Satz 2.6 entspricht dem, was wir bei Verwendung stückweise $d$-linearer Finite-Elemente-Basen und äquidistanter, rechtwinkliger Gitter erwarten durften. Besondere Beachtung verdienen die von $n$ unabhängigen Faktoren $\frac{d \cdot |u^{(d)}|_\infty}{6^d}$ beziehungsweise $\frac{d \cdot |u^{(d)}|_2}{9^d}$, die — anstelle der bei Standard-Fehlerabschätzungen üblichen zweiten Ableitungen — gemischte Ableitungen der Ordnung $2d$ enthalten.

**3. Dünne Gitter auf der Grundlage der $\mathbf{L_2}$- bzw. Maximumsnorm.** In diesem Abschnitt befassen wir uns nun mit $d$-dimensionalen *dünnen Gittern*. Dabei lehnen wir uns an [38] an, wo dünne Gitter vorgestellt und für den Fall von Funktionen zweier Veränderlicher bezüglich der Maximumsnorm untersucht wurden. Die Betrachtungen zu Speicheraufwand und Interpolationsfehler werden auf Funktionen von mehr als zwei Unbekannten erweitert. Die erforderlichen Fehlerabschätzungen werden bezüglich der Maximums- und der $L_2$-Norm durchgeführt.

**3.1. Der Raum $\mathbf{\tilde{S}_n^{(d)}}$ und seine Dimension.** Wir betrachten zunächst nochmals Abbildung 1. Dort entspricht der Voll-Gitter-Raum $S_3^{(2)}$ einem quadratischen Schema von Teilräumen $T_{i_1,i_2}$, $1 \leq i_1, i_2 \leq 3$. Für diese Teilräume gilt nach (5) $dim(T_{i_1,i_2}) = 2^{i_1+i_2-2}$. (13) zufolge sind außerdem $\|u_{i_1,i_2}\|_\infty$ und $\|u_{i_1,i_2}\|_2$ für $u_{i_1,i_2} \in T_{i_1,i_2}$ von der Ordnung $O(4^{-i_1-i_2})$. Betrachten wir nun im quadratischen Schema aus Abbildung 1 Teilräume $T_{i_1,i_2}$ mit $i_1 + i_2 = c$, so sind sie erstens von derselben Dimension und liefern zweitens zum Interpolanten einen Beitrag gleicher Größenordnung. Da die Dimension eines Teilraumes ein natürliches Maß für den zu investierenden Speicheraufwand darstellt und da ferner die Normen von $u_{i_1,i_2}$ ein Maß für den Informationsgehalt des Teilraumes $T_{i_1,i_2}$ bei der Darstellung einer Funktion sind, hat das quadratische Schema klare Nachteile. So wird im Raum $S_3^{(2)}$ der Teilraum $T_{3,3}$ berücksichtigt, obwohl er „teurer“ (mehr Unbekannte) und zugleich „unwichtiger“ (geringerer Beitrag zum Interpolationsfehler) ist als beispielsweise der Teilraum $T_{1,4}$, der nicht in $S_3^{(2)}$ enthalten ist.

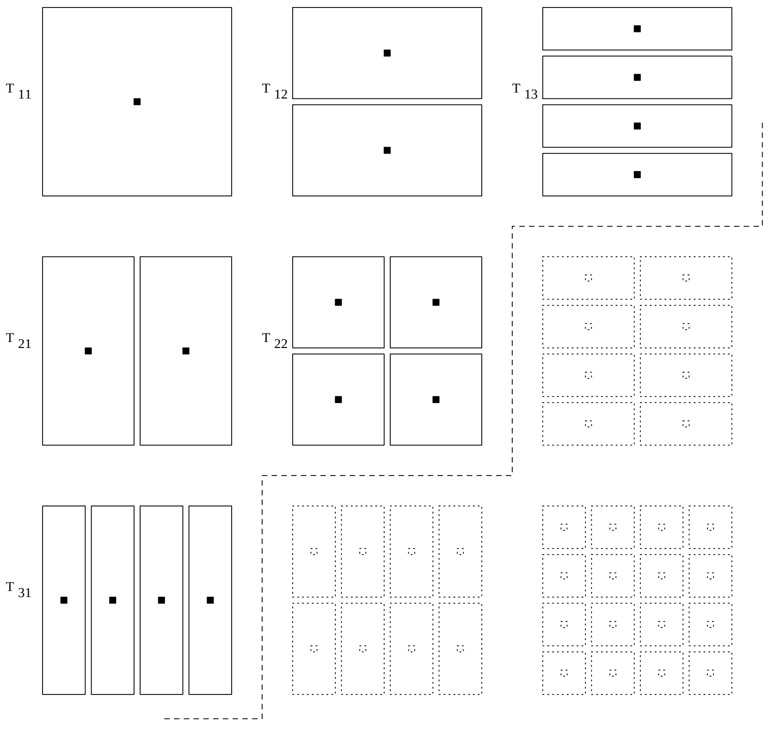



Abb. 4. *$d = 2$: Der Raum $\tilde{S}_3^{(2)}$ und seine Teilräume.*

Diese Erkenntnis führte in [38] zur Einführung zweidimensionaler dünner Gitter, basierend auf einem Dreiecksschema (siehe Abbildung 4). Davon ausgehend definieren wir nun im allgemeinen, $d$-dimensionalen Fall den Dünn-Gitter-Raum $\tilde{S}_n^{(d)}$ wie folgt:

$$\tilde{S}_n^{(d)} := \bigoplus_{i_1+\ldots+i_d \leq n+d-1} T_{i_1,\ldots,i_d}. \tag{15}$$

Dementsprechend ergibt sich für $u^{(d)} \in X_0^{(d)}$ sowie seinen Interpolanten $\tilde{u}_{n,I}^{(d)} \in \tilde{S}_n^{(d)}$ die hierarchische Darstellung ($u_{i_1,\ldots,i_d} \in T_{i_1,\ldots,i_d}$, vgl. (7))

$$\begin{aligned} u^{(d)} &= \sum_{i_1=1}^{\infty} \ldots \sum_{i_d=1}^{\infty} u_{i_1,\ldots,i_d}, \\ \tilde{u}_{n,I}^{(d)} &= \sum_{i_1+\ldots+i_d \leq n+d-1} u_{i_1,\ldots,i_d}. \end{aligned} \tag{16}$$

Übertragen auf die den Räumen zugrundeliegenden Gitter bedeutet diese Definition, daß wir anstelle des *vollen* Gitters $G_n^{(d)}$ ein *dünnes* Gitter $\tilde{G}_n^{(d)}$ erhalten, in dem die den weggelassenen Teilräumen entsprechenden Gitterpunkte fehlen. Abbildung 5 zeigt zwei dünne Gitter im zweidimensionalen Fall $d = 2$ (mit entsprechenden Randpunkten).

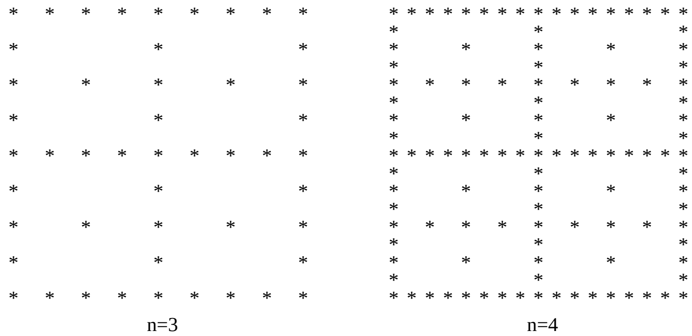



Abb. 5. *$d = 2$: Dünne Gitter $\tilde{G}_3^{(2)}$ und $\tilde{G}_4^{(2)}$.*

Als erstes wollen wir nun die Dimension von $\tilde{S}_n^{(d)}$ berechnen. Diese ist gleich der Anzahl der inneren Gitterpunkte des dünnen Gitters $\tilde{G}_n^{(d)}$. Aus Abbildung 5 können wir entnehmen, daß sich der innere Anteil des dünnen Gitters $\tilde{G}_4^{(2)}$ aus einer Zeile mit $15 = 2^4 - 1$ Gitterpunkten, zwei Zeilen mit 7, vier Zeilen mit 3 sowie acht Zeilen mit einem Gitterpunkt zusammensetzt. Dieser Zusammenhang führt uns zu folgender Rekursionsformel:

Satz 3.1. *Bezeichne*

$$N_n^{(d)} := dim(S_n^{(d)})$$

*die Dimension des Voll-Gitter-Raumes $S_n^{(d)}$, und bezeichne*

$$\tilde{N}_n^{(d)} := dim(\tilde{S}_n^{(d)})$$

*die Dimension des Dünn-Gitter-Raumes* $\tilde{S}_n^{(d)}$. *Dann gilt*

$$\begin{aligned} \tilde{N}_n^{(1)} &= N_n^{(1)} \;=\; 2^n - 1, \\ \tilde{N}_n^{(d)} &= \sum_{i=0}^{n-1} 2^i \cdot \tilde{N}_{n-i}^{(d-1)} \qquad \text{für } \; d > 1. \end{aligned} \tag{17}$$

*Damit gilt insbesondere*

$$\tilde{N}_n^{(d)} \;=\; O(2^n \cdot n^{d-1}) \;=\; O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^{d-1}). \tag{18}$$

*Beweis.* Die Behauptung für $\tilde{N}_n^{(1)}$ folgt unmittelbar aus (3) und (15), da im Falle $d = 1$ für alle $n$ $S_n^{(1)} \;=\; \tilde{S}_n^{(1)}$ gilt. Zum Beweis der Rekursionsformel für $\tilde{N}_n^{(d)}$ bei $d > 1$ schreiben wir zunächst (15) um:

$$\begin{aligned} \tilde{S}_n^{(d)} &= \bigoplus_{i_1+\ldots+i_d \le n+d-1} T_{i_1,\ldots,i_d} \\ &= \underbrace{\bigoplus_{i_1+\ldots+i_{d-1} \le n+d-2} T_{i_1,\ldots,i_{d-1},1}}_{i_d=1} \;\oplus \ldots \oplus\; \underbrace{\bigoplus_{i_1+\ldots+i_{d-1} \le d-1} T_{i_1,\ldots,i_{d-1},n}}_{i_d=n} . \end{aligned}$$

Da es sich hierbei um direkte Summen von Teilräumen handelt, führt (5) sofort zu

$$\begin{aligned} \tilde{N}_n^{(d)} &= \sum_{i_1+\ldots+i_d \le n+d-1} dim(T_{i_1,\ldots,i_d}) \\ &= \sum_{i_d=1}^{n} \left( \sum_{i_1+\ldots+i_{d-1} \le n-i_d+d-1} dim(T_{i_1,\ldots,i_{d-1},i_d}) \right) \\ &= \sum_{i_d=1}^{n} \left( \sum_{i_1+\ldots+i_{d-1} \le n-i_d+d-1} 2^{i_d-1} \cdot dim(T_{i_1,\ldots,i_{d-1}}) \right) \\ &= \sum_{i_d=1}^{n} \left( 2^{i_d-1} \cdot \sum_{i_1+\ldots+i_{d-1} \le n-i_d+d-1} dim(T_{i_1,\ldots,i_{d-1}}) \right) \\ &= \sum_{i=0}^{n-1} \left( 2^{i} \cdot \sum_{i_1+\ldots+i_{d-1} \le n-i+d-2} dim(T_{i_1,\ldots,i_{d-1}}) \right) \\ &= \sum_{i=0}^{n-1} 2^i \cdot \tilde{N}_{n-i}^{(d-1)}. \end{aligned}$$

Den Beweis von (18) führen wir mittels vollständiger Induktion nach $d$. Offensichtlich gilt für $d = 1$

$$\tilde{N}_n^{(1)} \;=\; 2^n - 1 \;=\; O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^0).$$

Aufgrund der Induktionsannahme und (17) gilt dann

$$\tilde{N}_n^{(d+1)} \;=\; \sum_{i=0}^{n-1} 2^i \cdot \tilde{N}_{n-i}^{(d)} \;=\; O\Big(\sum_{i=0}^{n-1} 2^i \cdot 2^{n-i} \cdot (n-i)^{d-1}\Big)$$

$$
\begin{aligned}
&= O\Big(\sum_{i=0}^{n-1} 2^n \cdot (n-i)^{d-1}\Big) \\
&= O\Big(2^n \cdot \sum_{i=0}^{n-1} (n-i)^{d-1}\Big) \\
&= O(2^n \cdot n^d).
\end{aligned}
$$

□

In den Spezialfällen $d = 2$ und $d = 3$ geben wir die resultierenden Formeln explizit an:

KOROLLAR 3.2. *Für die Dimension der Dünn-Gitter-Räume $\tilde{S}_n^{(2)}$ und $\tilde{S}_n^{(3)}$ gilt*

$$
\begin{aligned}
\tilde{N}_n^{(2)} &= 2^n \cdot (n-1) + 1, \\
\tilde{N}_n^{(3)} &= 2^n \cdot \frac{n^2 - n + 2}{2} - 1.
\end{aligned} \tag{19}
$$

*Beweis.* Nach Satz 3.1 gilt

$$
\begin{aligned}
\tilde{N}_n^{(2)} = \sum_{i=0}^{n-1} 2^i \cdot \tilde{N}_{n-i}^{(1)} &= \sum_{i=0}^{n-1} 2^i \cdot (2^{n-i} - 1) \\
&= \sum_{i=0}^{n-1} 2^n - \sum_{i=0}^{n-1} 2^i \\
&= n \cdot 2^n - (2^n - 1) \\
&= 2^n \cdot (n-1) + 1
\end{aligned}
$$

sowie

$$
\begin{aligned}
\tilde{N}_n^{(3)} = \sum_{i=0}^{n-1} 2^i \cdot \tilde{N}_{n-i}^{(2)} &= \sum_{i=0}^{n-1} 2^i \cdot \Big(2^{n-i} \cdot (n-i-1) + 1\Big) \\
&= 2^n \cdot \frac{n \cdot (n-1)}{2} + (2^n - 1) \\
&= 2^n \cdot \frac{n^2 - n + 2}{2} - 1.
\end{aligned}
$$

□

Der folgende Satz zeigt, daß auch bei Berücksichtigung des Randes $\delta\bar{\Omega}^{(d)}$ die Dimension des entsprechenden Dünn-Gitter-Raumes und damit die Anzahl der Gitterpunkte im dünnen Gitter (vgl. Abbildung 5) von der Größenordnung $h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^{d-1}$ ist.

SATZ 3.3. *Bezeichne $M_n^{(d)} = (2^n + 1)^d$ die Anzahl der Gitterpunkte des vollen Gitters $G_n^{(d)}$, und bezeichne $\tilde{M}_n^{(d)}$ die Anzahl der Gitterpunkte des dünnen Gitters $\tilde{G}_n^{(d)}$. Dann gilt*

$$
\begin{aligned}
\tilde{M}_n^{(1)} &= M_n^{(1)} = 2^n + 1, \\
\tilde{M}_n^{(d)} &= 3 \cdot \tilde{M}_n^{(d-1)} + \sum_{i=1}^{n-1} 2^i \cdot \tilde{M}_{n-i}^{(d-1)} \qquad \text{für } d > 1.
\end{aligned} \tag{20}
$$

*Insbesondere gilt*

$$\tilde{M}_n^{(d)} \;=\; O(2^n \cdot n^{d-1}) \;=\; O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^{d-1}). \tag{21}$$

*In den Spezialfällen $d = 2$ und $d = 3$ gilt*

$$\begin{aligned} \tilde{M}_n^{(2)} &= 2^n \cdot (n+3) + 1, \\ \tilde{M}_n^{(3)} &= 2^n \cdot \frac{n^2 + 11 \cdot n + 14}{2} + 1. \end{aligned} \tag{22}$$

*Beweis.* Die Behauptung für $d = 1$ ist wegen $G_n^{(1)} = \tilde{G}_n^{(1)}$ offensichtlich. Zum Beweis der Rekursionsformel für $\tilde{M}_n^{(d)}$, $d > 1$, betrachten wir die entsprechende Formel (17) für $\tilde{N}_n^{(d)}$ aus Satz 3.1. Da wir jetzt Randpunkte in die Betrachtung mit einbeziehen, lauten die einzelnen Summanden für $i > 0$ nun $2^i \cdot \tilde{M}_{n-i}^{(d-1)}$ anstelle von $2^i \cdot \tilde{N}_{n-i}^{(d-1)}$. Außerdem müssen die Ränder $x_d = 0$ und $x_d = 1$ berücksichtigt werden, weshalb der Summand für $i = 0$ statt $2^0 \cdot \tilde{N}_n^{(d-1)}$ jetzt $3 \cdot 2^0 \cdot \tilde{M}_n^{(d-1)}$ lautet.

Den Beweis von (21) führen wir wieder mittels vollständiger Induktion nach $d$. Hierbei ist der Induktionsanfang $d = 1$ offensichtlich. Aufgrund der Induktionsannahme, (18) und (20) gilt

$$\begin{aligned} \tilde{M}_n^{(d+1)} \;=\; 3 \cdot \tilde{M}_n^{(d)} \;+\; \sum_{i=1}^{n-1} 2^i \cdot \tilde{M}_{n-i}^{(d)} &= O\Big(\sum_{i=0}^{n-1} 2^i \cdot \tilde{M}_{n-i}^{(d)}\Big) \\ &= O\Big(\sum_{i=0}^{n-1} 2^i \cdot \tilde{N}_{n-i}^{(d)}\Big) \\ &= O(2^n \cdot n^d). \end{aligned}$$

Als letztes bleibt (22) zu zeigen. Aus (20) folgt

$$\begin{aligned} \tilde{M}_n^{(2)} &= 3 \cdot (2^n + 1) \;+\; \sum_{i=1}^{n-1} 2^i \cdot (2^{n-i} + 1) \\ &= 2 \cdot (2^n + 1) \;+\; \sum_{i=0}^{n-1} 2^n \;+\; \sum_{i=0}^{n-1} 2^i \\ &= 2 \cdot (2^n + 1) \;+\; n \cdot 2^n \;+\; (2^n - 1) \\ &= 2^n \cdot (n+3) + 1 \end{aligned}$$

sowie

$$\begin{aligned} \tilde{M}_n^{(3)} &= 3 \cdot \Big(2^n \cdot (n+3) + 1\Big) \;+\; \sum_{i=1}^{n-1} 2^i \cdot \Big(2^{n-i} \cdot (n-i+3) + 1\Big) \\ &= 2 \cdot \Big(2^n \cdot (n+3) + 1\Big) \;+\; \sum_{i=0}^{n-1} 2^n \cdot (n-i+3) \;+\; \sum_{i=0}^{n-1} 2^i \\ &= 2 \cdot \Big(2^n \cdot (n+3) + 1\Big) \;+\; 2^n \cdot \sum_{i=4}^{n+3} i \;+\; 2^n - 1 \end{aligned}$$

$$
\begin{aligned}
&= 2^n \cdot (2 \cdot n + 7) + 1 \; + \; 2^n \cdot \Big( \frac{(n+3) \cdot (n+4)}{2} - 6 \Big) \\
&= 2^n \cdot \frac{n^2 + 11 \cdot n + 14}{2} + 1.
\end{aligned}
$$

□

Als Ergebnis dieses Abschnitts halten wir fest, daß die in [38] für den Fall $d = 2$ vorgestellten und hier auf beliebiges $d$ verallgemeinerten dünnen Gitter im Vergleich zu herkömmlichen vollen Gittern signifikant weniger Gitterpunkte enthalten. Anstelle der bei vollen Gittern üblichen $O(h_n^{-d})$ Punkte enthalten dünne Gitter lediglich $O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^{d-1})$ Gitterpunkte. In gleichem Maße ist auch die Dimension des zur Interpolation verwendeten Funktionenraumes $\tilde{S}_n^{(d)}$ geringer als die des standardmäßig herangezogenen Raumes $S_n^{(d)}$.

Bevor wir nun den Interpolationsfehler $u^{(d)} - \tilde{u}_{n,I}^{(d)}$ genauer untersuchen, um so den Verlust an Genauigkeit durch den Übergang vom vollen auf das dünne Gitter in den Griff zu bekommen, sollen für $d = 3$ und einige Werte von $n$ die verschiedenen in diesem Abschnitt aufgetretenen Größen gegenübergestellt werden. Die Zahlen in Tabelle 1 zeigen deutlich die Vorteile dünner Gitter bezüglich Raumdimension und somit Speicherplatz.

TABELLE 1
*Vergleich von Raumdimension und Gitterpunktanzahl bei $d = 3$.*

| $n$ | $M_n^{(3)} = \vert G_n^{(3)} \vert$ | $N_n^{(3)} = dim(S_n^{(3)})$ | $\tilde{M}_n^{(3)} = \vert \tilde{G}_n^{(3)} \vert$ | $\tilde{N}_n^{(3)} = dim(\tilde{S}_n^{(3)})$ |
|---|---|---|---|---|
| 4 | 4913 | 3375 | 593 | 111 |
| 5 | 35937 | 29791 | 1505 | 351 |
| 6 | 274625 | 250047 | 3713 | 1023 |
| 7 | $2.1 \cdot 10^{06}$ | $2.0 \cdot 10^{06}$ | 8961 | 2815 |
| 8 | $1.7 \cdot 10^{07}$ | $1.7 \cdot 10^{07}$ | 21249 | 7423 |
| 9 | $1.4 \cdot 10^{08}$ | $1.3 \cdot 10^{08}$ | 49665 | 18943 |
| 10 | $1.1 \cdot 10^{09}$ | $1.1 \cdot 10^{09}$ | 114689 | 47103 |
| 15 | $3.5 \cdot 10^{13}$ | $3.5 \cdot 10^{13}$ | $6.6 \cdot 10^{06}$ | $3.5 \cdot 10^{06}$ |
| 20 | $1.2 \cdot 10^{18}$ | $1.2 \cdot 10^{18}$ | $3.3 \cdot 10^{08}$ | $2.0 \cdot 10^{08}$ |

**3.2. Der Interpolationsfehler im Raum $\mathbf{\tilde{S}_n^{(d)}}$.** Wie wir im vorigen Abschnitt gesehen haben, geht der Übergang von $G_n^{(d)}$ zu $\tilde{G}_n^{(d)}$ beziehungsweise von $S_n^{(d)}$ zu $\tilde{S}_n^{(d)}$ einher mit einer deutlichen Reduzierung der Anzahl der Gitterpunkte beziehungsweise der Raumdimension. Entscheidend für die Effizienz dünner Gitter ist natürlich die Frage, wie stark sich dieser Verzicht auf Gitterpunkte auf die Genauigkeit auswirkt, mit der der Raum $\tilde{S}_n^{(d)}$ eine Funktion $u^{(d)} \in X_0^{(d)}$ interpolieren kann. Zur Beantwortung dieser Frage sollen nun Schranken für die Maximums- und die $L_2$-Norm des Interpolationsfehlers $u^{(d)} - \tilde{u}_{n,I}^{(d)}$ (vgl. (16)) hergeleitet werden. Das nachfolgende Lemma stellt hierzu einige Hilfsmittel bereit.

LEMMA 3.4. *Die Größen* $A_n^{(d)}$ *und* $B_n^{(d)}$,

$$A_n^{(d)} := \sum_{i_1+\ldots+i_d \le n+d-1} 4^{-i_1-\ldots-i_d}, \qquad B_n^{(d)} := \sum_{i_1+\ldots+i_d \le n+d-1} 1,$$

*erfüllen folgende Gleichungen:*

$$\begin{aligned}
d=1: \quad A_n^{(1)} &= \frac{1-4^{-n}}{3}, \\
B_n^{(1)} &= n; \\
d>1: \quad A_n^{(d)} &= \left(\frac{1}{3}\right)^{d-1} \cdot A_n^{(1)} - \sum_{i=1}^{d-1} \left(\left(\frac{1}{3}\right)^{d-i} \cdot 4^{-n-i} \cdot B_n^{(i)}\right), \\
B_n^{(d)} &= \binom{n+d-1}{d}.
\end{aligned} \tag{23}$$

*Beweis.* Die Definitionen von $A_n^{(d)}$ und $B_n^{(d)}$ liefern direkt

$$\begin{aligned}
A_n^{(1)} &= \sum_{i_1=1}^{n} 4^{-i_1} = \frac{1}{4} \cdot \sum_{i_1=0}^{n-1} 4^{-i_1} = \frac{1-4^{-n}}{3} \quad \text{sowie} \\
B_n^{(1)} &= \sum_{i_1=1}^{n} 1 = n.
\end{aligned}$$

Im Falle $d>1$ schreiben wir $A_n^{(d)}$ ausführlich

$$\begin{aligned}
A_n^{(d)} &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n-i_1+1} \cdots \sum_{i_{d-1}=1}^{n-i_1-\ldots-i_{d-2}+d-2} \sum_{i_d=1}^{n-i_1-\ldots-i_{d-1}+d-1} 4^{-i_1-\ldots-i_d} \\
&= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n-i_1+1} \cdots \sum_{i_{d-1}=1}^{n-i_1-\ldots-i_{d-2}+d-2} 4^{-i_1-\ldots-i_{d-1}} \cdot \left( \sum_{i_d=1}^{n-i_1-\ldots-i_{d-1}+d-1} 4^{-i_d} \right) \\
&= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n-i_1+1} \cdots \sum_{i_{d-1}=1}^{n-i_1-\ldots-i_{d-2}+d-2} 4^{-i_1-\ldots-i_{d-1}} \cdot \frac{1}{3} \cdot (1-4^{-n+i_1+\ldots+i_{d-1}-d+1}) \\
&= \frac{1}{3} \cdot \sum_{i_1=1}^{n} \sum_{i_2=1}^{n-i_1+1} \cdots \sum_{i_{d-1}=1}^{n-i_1-\ldots-i_{d-2}+d-2} \left(4^{-i_1-\ldots-i_{d-1}} - 4^{-n-d+1}\right) \\
&= \frac{1}{3} \cdot A_n^{(d-1)} - \frac{1}{3} \cdot 4^{-n-(d-1)} \cdot B_n^{(d-1)} \\
&= \frac{1}{3} \cdot \left(\frac{1}{3} \cdot A_n^{(d-2)} - \frac{1}{3} \cdot 4^{-n-(d-2)} \cdot B_n^{(d-2)}\right) - \frac{1}{3} \cdot 4^{-n-(d-1)} \cdot B_n^{(d-1)} \\
&= \ldots \\
&= \left(\frac{1}{3}\right)^{d-1} \cdot A_n^{(1)} - \sum_{i=1}^{d-1} \left(\left(\frac{1}{3}\right)^{d-i} \cdot 4^{-n-i} \cdot B_n^{(i)}\right).
\end{aligned}$$

Der Definition von $B_n^{(d)}$ schließlich kann man entnehmen, daß diese Größe gerade die Anzahl der $d$-Tupel natürlicher Zahlen angibt, deren Summe kleiner oder gleich $n+d-1$

ist. Mittels vollständiger Induktion nach $k$ zeigen wir allgemein, daß die Anzahl der $k$-Tupel natürlicher Zahlen, deren Summe kleiner oder gleich $m$ ist, durch den Binomialkoeffizienten $\binom{m}{k}$ bestimmt ist. Der Induktionsanfang $k = 1$ ist offensichtlich. Beim Induktionsschritt $k \to k+1$ kann die $(k+1)$-te Zahl die Werte $1, ..., m-k$ annehmen, weshalb sich für die restlichen $k$ Zahlen nach Induktionsannahme $\binom{m-1}{k}, ..., \binom{m-(m-k)}{k}$ Möglichkeiten ergeben. Mit $\binom{m}{k} + \binom{m}{k+1} = \binom{m+1}{k+1}$ gilt schließlich für die Anzahl der $(k+1)$-Tupel natürlicher Zahlen, deren Summe kleiner oder gleich $m$ ist:

$$
\begin{aligned}
& \binom{m-1}{k} + \binom{m-2}{k} + \binom{m-3}{k} + \binom{m-4}{k} + \ldots + \binom{k}{k} \\
= & \binom{m}{k+1} - \binom{m-1}{k+1} + \binom{m-2}{k} + \binom{m-3}{k} + \binom{m-4}{k} + \ldots + \binom{k}{k} \\
= & \binom{m}{k+1} - \binom{m-2}{k+1} + \binom{m-3}{k} + \binom{m-4}{k} + \ldots + \binom{k}{k} \\
= & \binom{m}{k+1} - \binom{m-3}{k+1} + \binom{m-4}{k} + \ldots + \binom{k}{k} \\
= & \ldots \\
= & \binom{m}{k+1} - \binom{k+1}{k+1} + \binom{k}{k} \\
= & \binom{m}{k+1}.
\end{aligned}
$$

Mit $m = n+d-1$ und $k = d$ folgt somit $B_n^{(d)} = \binom{n+d-1}{d}$, womit Lemma 3.4 vollständig bewiesen ist. □

Damit sind wir nun in der Lage, die gewünschten Schranken für den Interpolationsfehler anzugeben:

SATZ 3.5. *Gegeben seien* $u^{(d)} \in X_0^{(d)}$ *und sein stückweise* $d$*-linearer Interpolant* $\tilde{u}_{n,I}^{(d)} \in \tilde{S}_n^{(d)}$ *in hierarchischer Darstellung (16). Dann gelten folgende Abschätzungen:*

$$
\begin{aligned}
\|u^{(d)} - \tilde{u}_{n,I}^{(d)}\|_\infty &\leq \frac{|u^{(d)}|_\infty}{6^d} \cdot \left(1 + \sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot \binom{n+i-1}{i}\right) \cdot h_n^2, \\
\|u^{(d)} - \tilde{u}_{n,I}^{(d)}\|_2 &\leq \frac{|u^{(d)}|_2}{9^d} \cdot \left(1 + \sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot \binom{n+i-1}{i}\right) \cdot h_n^2.
\end{aligned}
\tag{24}
$$

*Insbesondere gilt für die Ordnung des Interpolationsfehlers*

$$
\begin{aligned}
\|u^{(d)} - \tilde{u}_{n,I}^{(d)}\|_\infty &= O\left(h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^{d-1}\right), \\
\|u^{(d)} - \tilde{u}_{n,I}^{(d)}\|_2 &= O\left(h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^{d-1}\right).
\end{aligned}
\tag{25}
$$

*Beweis.* Nach (13) sowie (16) gilt

$$\begin{aligned}
\|u^{(d)} - \tilde{u}^{(d)}_{n,I}\|_\infty &= \Big\| \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} u_{i_1,\dots,i_d} - \sum_{i_1+\dots+i_d \le n+d-1} u_{i_1,\dots,i_d} \Big\|_\infty \\
&\le \sum_{i_1+\dots+i_d > n+d-1} \|u_{i_1,\dots,i_d}\|_\infty \\
&\le \sum_{i_1+\dots+i_d > n+d-1} \frac{|u^{(d)}|_\infty}{2^d} \cdot h_{i_1}^2 \cdot \dots \cdot h_{i_d}^2 \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} 4^{-i_1-\dots-i_d} - \sum_{i_1+\dots+i_d \le n+d-1} 4^{-i_1-\dots-i_d} \right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \Big( \sum_{i_1=1}^{\infty} 4^{-i_1} \Big)^d - \sum_{i_1+\dots+i_d \le n+d-1} 4^{-i_1-\dots-i_d} \right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \left(\frac{1}{3}\right)^d - A_n^{(d)} \right).
\end{aligned}$$

Setzt man jetzt für $A_n^{(d)}$ das Ergebnis aus Lemma 3.4 ein, so erhält man

$$\begin{aligned}
\|u^{(d)} - \tilde{u}^{(d)}_{n,I}\|_\infty &\le \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \left(\frac{1}{3}\right)^d - \left(\frac{1}{3}\right)^{d-1} \cdot A_n^{(1)} + \sum_{i=1}^{d-1} \left(\frac{1}{3}\right)^{d-i} \cdot 4^{-n-i} \cdot B_n^{(i)} \right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \left(\frac{1}{3}\right)^d - \left(\frac{1}{3}\right)^{d-1} \cdot A_n^{(1)} + \left(\frac{1}{3}\right)^d \cdot 4^{-n} \cdot \sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot B_n^{(i)} \right) \\
&= \frac{|u^{(d)}|_\infty}{2^d} \cdot \left( \left(\frac{1}{3}\right)^d \cdot 4^{-n} + \left(\frac{1}{3}\right)^d \cdot 4^{-n} \cdot \sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot \binom{n+i-1}{i} \right) \\
&= \frac{|u^{(d)}|_\infty}{6^d} \cdot \left( 1 + \sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot \binom{n+i-1}{i} \right) \cdot h_n^2.
\end{aligned}$$

Analog ergibt sich

$$\begin{aligned}
\|u^{(d)} - \tilde{u}^{(d)}_{n,I}\|_2 &= \Big\| \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} u_{i_1,\dots,i_d} - \sum_{i_1+\dots+i_d \le n+d-1} u_{i_1,\dots,i_d} \Big\|_2 \\
&\le \frac{|u^{(d)}|_2}{3^d} \cdot \left( \left(\frac{1}{3}\right)^d - A_n^{(d)} \right) \\
&= \frac{|u^{(d)}|_2}{9^d} \cdot \left( 1 + \sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot \binom{n+i-1}{i} \right) \cdot h_n^2,
\end{aligned}$$

womit (24) bewiesen ist.

Zum Beweis der Ordnung der Normen des Interpolationsfehlers (25) ist nun noch

$$\sum_{i=1}^{d-1} \left(\frac{3}{4}\right)^i \cdot \binom{n+i-1}{i} = O\left( (\mathrm{ld}(h_n^{-1}))^{d-1} \right)$$

zu zeigen, was jedoch unmittelbar aus $n = \mathrm{ld}(h_n^{-1})$ und $\binom{n+i-1}{i} = O(n^i)$ folgt. □

Für die beiden Spezialfälle $d = 2$ und $d = 3$ geben wir die aus Satz 3.5 resultierenden Schranken explizit an:

KOROLLAR 3.6. *Gegeben seien $u^{(d)} \in X_0^{(d)}$ und sein stückweise d-linearer Interpolant $\tilde{u}_{n,I}^{(d)} \in \tilde{S}_n^{(d)}$ in hierarchischer Darstellung (16). Dann gelten für $d = 2$ beziehungsweise $d = 3$ die Abschätzungen*

$$
\begin{aligned}
\|u^{(2)} - \tilde{u}_{n,I}^{(2)}\|_\infty &\leq \frac{|u^{(2)}|_\infty}{36} \cdot \left(1 + \frac{3}{4} \cdot n\right) \cdot h_n^2, \\
\|u^{(2)} - \tilde{u}_{n,I}^{(2)}\|_2 &\leq \frac{|u^{(2)}|_2}{81} \cdot \left(1 + \frac{3}{4} \cdot n\right) \cdot h_n^2, \\
\|u^{(3)} - \tilde{u}_{n,I}^{(3)}\|_\infty &\leq \frac{|u^{(3)}|_\infty}{216} \cdot \left(1 + \frac{33}{32} \cdot n + \frac{9}{32} \cdot n^2\right) \cdot h_n^2, \\
\|u^{(3)} - \tilde{u}_{n,I}^{(3)}\|_2 &\leq \frac{|u^{(3)}|_2}{729} \cdot \left(1 + \frac{33}{32} \cdot n + \frac{9}{32} \cdot n^2\right) \cdot h_n^2.
\end{aligned}
\tag{26}
$$

*Beweis.* Aus (24) folgt mit $d = 2$ beziehungsweise $d = 3$ sofort

$$
\begin{aligned}
\|u^{(2)} - \tilde{u}_{n,I}^{(2)}\|_\infty &\leq \frac{|u^{(2)}|_\infty}{36} \cdot \left(1 + \frac{3}{4} \cdot \binom{n}{1}\right) \cdot h_n^2 \\
&= \frac{|u^{(2)}|_\infty}{36} \cdot \left(1 + \frac{3}{4} \cdot n\right) \cdot h_n^2
\end{aligned}
$$

sowie

$$
\begin{aligned}
\|u^{(3)} - \tilde{u}_{n,I}^{(3)}\|_\infty &\leq \frac{|u^{(3)}|_\infty}{216} \cdot \left(1 + \frac{3}{4} \cdot n + \frac{9}{16} \cdot \frac{(n+1) \cdot n}{2}\right) \cdot h_n^2 \\
&= \frac{|u^{(3)}|_\infty}{216} \cdot \left(1 + \frac{33}{32} \cdot n + \frac{9}{32} \cdot n^2\right) \cdot h_n^2.
\end{aligned}
$$

Die Aussagen bezüglich der $L_2$-Norm erhält man analog. □

Neben Satz 3.1 stellt Satz 3.5 das zweite wichtige Ergebnis im Zusammenhang mit dünnen Gittern dar. In Satz 3.1 wurde gezeigt, daß die Dimension der Dünn-Gitter-Räume $\tilde{S}_n^{(d)}$ im Vergleich zu den Voll-Gitter-Räumen $S_n^{(d)}$ von $O(h_n^{-d})$ auf $O\left(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^{d-1}\right)$ abnimmt. Satz 3.5 nun stellt sicher, daß dieser signifikanten Reduktion der Anzahl der Unbekannten ein nur geringer Verlust an Approximationsgenauigkeit gegenübersteht. Anstelle der auf vollen Gittern erzielten Genauigkeit der Ordnung $O(h_n^2)$ erreicht man auf dünnen Gittern $O\left(h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^{d-1}\right)$, was jedoch keine wesentliche Verschlechterung bedeutet. Somit bieten dünne Gitter eine sehr effiziente Möglichkeit zur Speicherung und Darstellung glatter Funktionen.

Abschließend sei noch darauf hingewiesen, daß die Verwendung dünner Gitter zur Repräsentation glatter Funktionen auch als Interpolationsverfahren höherer Ordnung angesehen werden kann. Sieht man einmal von den logarithmischen Faktoren ab, so ergeben sich hier Interpolationsfehler der Ordnung $O(h_n^2)$ mit $O(h_n^{-1})$ Gitterpunkten, was auf vollen Gittern bereits einem Verfahren der Ordnung $O(h_n^{2d})$ entspricht. Auch hieran sieht man wiederum die mit wachsender Dimensionalität $d$ zunehmende Attraktivität dünner Gitter.

**4. Dünne Gitter auf der Grundlage der Energienorm.** Unsere bisherige Vorgehensweise orientierte sich stark an der Maximumsnorm und an der $L_2$-Norm. In diesem Abschnitt stellen wir nun ein alternatives Dünn-Gitter-Konzept vor, welches sich in verschiedener Hinsicht als vorteilhafter erweisen wird.

**4.1. Der Raum $\hat{\mathbf{S}}_{\mathbf{n}}^{(\mathbf{d})}$ und seine Dimension.** Wir führen zunächst einen weiteren Dünn-Gitter-Raum $\hat{S}_n^{(d)}$ mit zugehörigem dünnen Gitter $\hat{G}_n^{(d)}$ ein. Dieser entsteht aus dem Voll-Gitter-Raum $S_n^{(d)}$, indem man im Vergleich zum bereits eingeführten Dünn-Gitter-Raum $\tilde{S}_n^{(d)}$ auf noch mehr Teilräume $T_{i_1,\ldots,i_d}$ verzichtet.

Bei der Definition von $\hat{S}_n^{(d)}$ lassen wir uns jetzt im Gegensatz zur bisherigen Vorgehensweise nicht mehr von der $L_2$- oder der Maximumsnorm leiten, sondern greifen auf die Energienorm zurück. Dies erscheint gerade deshalb als sinnvoll, weil der im zweiten Teil dieser Arbeit beschriebene Poisson-Löser auf dünnen Gittern als Finite-Elemente-Algorithmus den Fehler bezüglich der Energienorm minimiert (siehe [35]). Es wird gezeigt, daß $\hat{G}_n^{(d)}$ nur noch $O(h_n^{-1})$ Gitterpunkte enthält. Dennoch ist der Interpolationsfehler bezüglich der Energienorm von derselben Ordnung $O(h_n)$ wie bei Verwendung eines vollen Gitters.

Dem Finite-Elemente-Kontext entlehnen wir die folgende Definition der Energienorm einer Funktion $u^{(d)} \in X_0^{(d)}$ (beziehungsweise ihres Quadrats, der Energie):

$$\|u^{(d)}\|_E^2 \;:=\; \int\ldots\int_\Omega \left( \left(\frac{\partial u^{(d)}}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial u^{(d)}}{\partial x_d}\right)^2 \right) dx_d \,\ldots\, dx_1 .$$

Davon ausgehend soll nun zunächst der Beitrag $\|u_{i_1,\ldots,i_d}\|_E^2$ eines Teilraumes $T_{i_1,\ldots,i_d}$ zur Gesamtenergie $\|u^{(d)}\|_E^2$ untersucht werden.

Lemma 4.1. *Die Energie einer Basisfunktion* $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \in T_{i_1,\ldots,i_d}$ , $1 \le k_j \le 2^{i_j-1}$, $i_j \ge 1$, $1 \le j \le d$ , *beträgt*

$$\|g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}\|_E^2 \;=\; 2 \cdot \left(\frac{2}{3}\right)^{d-1} \cdot \sum_{j=1}^{d} \frac{h_{i_1} \cdot \ldots \cdot h_{i_d}}{h_{i_j}^2}. \tag{27}$$

*Für die Energie einer beliebigen Funktion* $u_{i_1,\ldots,i_d} \in T_{i_1,\ldots,i_d}$ *gilt*

$$\|u_{i_1,\ldots,i_d}\|_E^2 \;\le\; \frac{d \cdot |u^{(d)}|_\infty^2}{4 \cdot 12^{d-1}} \cdot h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2 \cdot \max_{1\le j\le d} \left( \frac{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2}{h_{i_j}^2} \right). \tag{28}$$

*Beweis.* O.B.d.A. sei der Ursprung Mittelpunkt des Trägers der betrachteten Basisfunktion $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$. Mit Definition (6) erhält man

$$\begin{aligned}
\|g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}\|_E^2 \;&=\; \int_{-h_{i_1}}^{+h_{i_1}} \ldots \int_{-h_{i_d}}^{+h_{i_d}} \sum_{j=1}^{d} \left( \frac{\partial g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}}{\partial x_j} \right)^2 dx_d \,\ldots\, dx_1 \\
&=\; \frac{2^d}{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2} \cdot \int_0^{h_{i_1}} \ldots \int_0^{h_{i_d}} \sum_{j=1}^{d} \frac{x_1^2 \cdot \ldots \cdot x_d^2}{x_j^2} \; dx_d \,\ldots\, dx_1
\end{aligned}$$

$$
\begin{aligned}
&= \frac{2^d}{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2} \cdot \sum_{j=1}^{d} \frac{h_{i_1}^3 \cdot \ldots \cdot h_{i_d}^3}{3^{d-1} \cdot h_{i_j}^2} \\
&= 2 \cdot \left(\frac{2}{3}\right)^{d-1} \cdot \sum_{j=1}^{d} \frac{h_{i_1} \cdot \ldots \cdot h_{i_d}}{h_{i_j}^2}.
\end{aligned}
$$

Zum Beweis von (28) schreiben wir zunächst $u_{i_1,\ldots,i_d} \in T_{i_1,\ldots,i_d}$ als Linearkombination von Basisfunktionen $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \in T_{i_1,\ldots,i_d}$ (beachte $dim(T_{i_1,\ldots,i_d}) = 2^{i_1-1} \cdot \ldots \cdot 2^{i_d-1}$ sowie die Disjunktheit der Träger der $g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \in T_{i_1,\ldots,i_d}$):

$$
\begin{aligned}
\|u_{i_1,\ldots,i_d}\|_E^2 &= \left\| \sum_{k_1=1}^{2^{i_1-1}} \ldots \sum_{k_d=1}^{2^{i_d-1}} c_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \cdot g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)} \right\|_E^2 \\
&= \sum_{k_1=1}^{2^{i_1-1}} \ldots \sum_{k_d=1}^{2^{i_d-1}} |c_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}|^2 \cdot \|g_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}\|_E^2 \ .
\end{aligned}
$$

Für die Koeffizienten $c_{k_1,\ldots,k_d}^{(i_1,\ldots,i_d)}$ wurde in Abschnitt 2 bereits eine obere Schranke angegeben (siehe (12)). Damit und mit (27) folgt

$$
\begin{aligned}
\|u_{i_1,\ldots,i_d}\|_E^2 &\leq \sum_{k_1=1}^{2^{i_1-1}} \ldots \sum_{k_d=1}^{2^{i_d-1}} \left( \left( |u^{(d)}|_\infty \cdot \frac{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2}{2^d} \right)^2 \cdot 2 \cdot \left(\frac{2}{3}\right)^{d-1} \cdot \sum_{j=1}^{d} \frac{h_{i_1} \cdot \ldots \cdot h_{i_d}}{h_{i_j}^2} \right) \\
&= \sum_{k_1=1}^{2^{i_1-1}} \ldots \sum_{k_d=1}^{2^{i_d-1}} \left( \frac{|u^{(d)}|_\infty^2}{2 \cdot 6^{d-1}} \cdot h_{i_1}^3 \cdot \ldots \cdot h_{i_d}^3 \cdot \sum_{j=1}^{d} \frac{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2}{h_{i_j}^2} \right) \\
&= \frac{1}{2^d \cdot h_{i_1} \cdot \ldots \cdot h_{i_d}} \cdot \left( \frac{|u^{(d)}|_\infty^2}{2 \cdot 6^{d-1}} \cdot h_{i_1}^3 \cdot \ldots \cdot h_{i_d}^3 \cdot \sum_{j=1}^{d} \frac{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2}{h_{i_j}^2} \right) \\
&= \frac{|u^{(d)}|_\infty^2}{4 \cdot 12^{d-1}} \cdot h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2 \cdot \sum_{j=1}^{d} \frac{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2}{h_{i_j}^2} \\
&\leq \frac{d \cdot |u^{(d)}|_\infty^2}{4 \cdot 12^{d-1}} \cdot h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2 \cdot \max_{1 \leq j \leq d} \left( \frac{h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2}{h_{i_j}^2} \right) .
\end{aligned}
$$

□

Dieses Resultat führt nun zur Definition des neuen Dünn-Gitter-Raumes $\hat{S}_n^{(d)}$. (13) zeigt, daß der in der $L_2$-Norm gemessene Beitrag eines Teilraumes $T_{i_1,\ldots,i_d}$ zum Interpolationsfehler für alle $T_{i_1,\ldots,i_d}$ mit $i_1 + \ldots + i_d = c$ von derselben Ordnung $O(h_{i_1}^2 \cdot \ldots \cdot h_{i_d}^2)$ ist. Betrachtet man dagegen den Interpolationsfehler in der Energienorm, so liefern nach (28) Teilräume $T_{i_1,\ldots,i_d}$ mit $i_1 + \ldots + i_d = c$ und $i_1 \approx \ldots \approx i_d$ erheblich kleinere Beiträge als Teilräume mit $i_1 + \ldots + i_d = c$ und sehr unterschiedlichen Werten $i_j$ ($\min(i_j) \ll \max(i_j)$). Es liegt deshalb nahe, diejenigen Teilräume $T_{i_1,\ldots,i_d}$ in $S_n^{(d)}$ beziehungsweise $\tilde{S}_n^{(d)}$ zu vernachlässigen, die sich bezüglich der Energienorm als weniger informationstragend herausstellen und dennoch eine hohe Dimension aufweisen. Wir betrachten also das Verhältnis von Nutzen (Beitrag zum Interpolationsfehler) zu Kosten

(Raumdimension) und erhalten mit (5) sowie (28) für einen Teilraum $T_{i_1,...,i_d}$

$$\begin{aligned}\frac{\text{Nutzen}}{\text{Kosten}} &= \frac{O\left(h_{i_1}\cdot ...\cdot h_{i_d}\cdot \max_{1\le j\le d}\left(\frac{h_{i_1}\cdot ...\cdot h_{i_d}}{h_{i_j}}\right)\right)}{2^{-d}\cdot h_{i_1}^{-1}\cdot ...\cdot h_{i_d}^{-1}} \\ &= O\left(h_{i_1}^2\cdot ...\cdot h_{i_d}^2\cdot \max_{1\le j\le d}\left(\frac{h_{i_1}\cdot ...\cdot h_{i_d}}{h_{i_j}}\right)\right).\end{aligned}$$

Dieses Verhältnis wird nun für die Berücksichtigung des Teilraumes $T_{i_1,...,i_d}$ in $\hat{S}_n^{(d)}$ maßgeblich sein. Mit der Vereinbarung

$$n-4 \;=\; k\cdot 5+r, \quad 0\le r\le 4,$$

führen wir zum Zwecke der formalen Definition des Raumes $\hat{S}_n^{(d)}$ zunächst drei Indexmengen $B_n^{(d)}$, $\tilde{B}_n^{(d)}$ und $\hat{B}_n^{(d)}$ ein, die für alle natürlichen Zahlen $n$ und $d$ die Beziehung $\hat{B}_n^{(d)}\subseteq\tilde{B}_n^{(d)}\subseteq B_n^{(d)}$ erfüllen. Im folgenden gelte $(i_1,...,i_d)\in\mathbb{N}^d$:

$$\begin{aligned}B_n^{(d)} &:= \Big\{(i_1,...,i_d):\ i_j\le n\ \forall j\Big\} \;=\; \Big\{(i_1,...,i_d):\ h_{i_j}\ge h_n\ \forall j\Big\}, \\ \tilde{B}_n^{(d)} &:= \Big\{(i_1,...,i_d):\ \sum_{j=1}^d i_j \;\le\; n+d-1\Big\} \\ &= \Big\{(i_1,...,i_d):\ \prod_{j=1}^d h_{i_j} \;\ge\; h_{n+d-1}\Big\}, \\ \hat{B}_n^{(d)} &:= \Big\{(i_1,...,i_d):\ \sum_{j=1}^d i_j + \frac{1}{2}\min_{1\le j\le d}\sum_{l\ne j} i_l \;\le\; n+\frac{3}{2}\cdot d-1\Big\} \cup \tilde{B}_{n-k-1}^{(d)} \\ &= \Big\{(i_1,...,i_d):\ \prod_{j=1}^d h_{i_j}\cdot\Big(\max_{1\le j\le d}\prod_{l\ne j} h_{i_l}\Big)^{\frac{1}{2}} \;\ge\; h_{n+\frac{3}{2}\cdot d-1}\Big\} \cup \tilde{B}_{n-k-1}^{(d)}.\end{aligned} \tag{29}$$

Die Vereinigung mit $\tilde{B}_{n-k-1}^{(d)}$ hat lediglich beweistechnische Gründe. Dadurch wird sichergestellt, daß von den $d$-Tupeln aus $\tilde{B}_n^{(d)}$ nur solche mit $i_1+...+i_d\ge n+d-1-k$ entfernt werden können, um zu $\hat{B}_n^{(d)}$ zu gelangen.

Mit Hilfe dieser Indexmengen können wir nun unsere drei Funktionenräume wie folgt beschreiben:

$$\begin{aligned}S_n^{(d)} &= \bigoplus_{(i_1,...,i_d)\in B_n^{(d)}} T_{i_1,...,i_d}\ , \\ \tilde{S}_n^{(d)} &= \bigoplus_{(i_1,...,i_d)\in \tilde{B}_n^{(d)}} T_{i_1,...,i_d}\ , \\ \hat{S}_n^{(d)} &= \bigoplus_{(i_1,...,i_d)\in \hat{B}_n^{(d)}} T_{i_1,...,i_d}\ .\end{aligned} \tag{30}$$

Dementsprechend ergibt sich für $u^{(d)}\in X_0^{(d)}$ sowie seinen Interpolanten $\hat{u}_{n,I}^{(d)}\in\hat{S}_n^{(d)}$ die hierarchische Darstellung ($u_{i_1,...,i_d}\in T_{i_1,...,i_d}$, vgl. (7) und (16))

$$
(31) \qquad
\begin{aligned}
u^{(d)} &= \lim_{n \to \infty} \sum_{(i_1,\ldots,i_d) \in B_n^{(d)}} u_{i_1,\ldots,i_d} \;, \\
\hat{u}_{n,I}^{(d)} &= \sum_{(i_1,\ldots,i_d) \in \hat{B}_n^{(d)}} u_{i_1,\ldots,i_d} \;.
\end{aligned}
$$

Abbildung 6 veranschaulicht den zweidimensionalen Fall. $\tilde{S}_n^{(2)}$ entstand aus $S_n^{(2)}$, indem aus dem quadratischen Schema der Teilräume $T_{i_1,i_2}$ diejenigen mit $i_1 + i_2 > n + 1$ weggelassen wurden. Zu $\hat{S}_n^{(2)}$ gelangen wir nun, indem wir aus den Diagonalen $i_1 + i_2 = n+1-s$, $s = 0, ..., k$, weitere Teilräume entfernen. Im dreidimensionalen Fall verzichten wir auf weitere Teilräume $T_{i_1,i_2,i_3}$ aus den Ebenen $i_1 + i_2 + i_3 = n + 2 - s$, $s = 0, ..., k$.

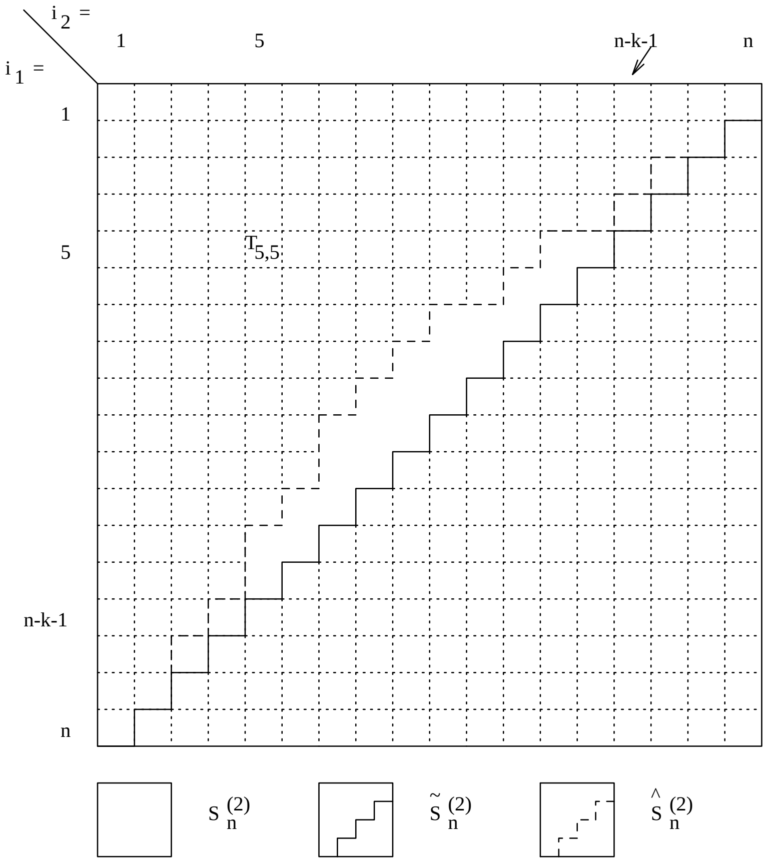



ABB. 6. $S_n^{(2)}$, $\tilde{S}_n^{(2)}$ *und* $\hat{S}_n^{(2)}$.

Der folgende Satz stellt nun eine Formel bereit, die für beliebiges $d$ die Berechnung der Dimension von $\hat{S}_n^{(d)}$ und damit der Anzahl der Gitterpunkte des zugehörigen dünnen Gitters $\hat{G}_n^{(d)}$ gestattet.

SATZ 4.2. *Für die Dimension des Raumes $\hat{S}_n^{(d)}$ gilt ($n - 4 = 5k + r$, $0 \leq r \leq 4$):*

$$
(32) \qquad dim(\hat{S}_n^{(d)}) \;=\; dim(\tilde{S}_{n-k-1}^{(d)}) \;+\; \sum_{s=0}^{k} d \cdot \binom{d+2s}{d-1} \cdot 2^{n-1-s} \;=\; O(h_n^{-1}).
$$

*Beweis.* Mit der Vereinbarung

$$\hat{B}^{(d)}_{n;s} \;:=\; \Big\{(i_1,...,i_d):\; \sum_{j=1}^{d} i_j = n+d-1-s \;\wedge\; \sum_{j=1}^{d} i_j \;+\; \frac{1}{2}\min_{1\le j\le d}\sum_{l\ne j} i_l \;\le\; n+\frac{3}{2}\cdot d-1\Big\}$$

für $0 \le s \le k$ können wir $\hat{B}^{(d)}_n$ wie folgt beschreiben:

$$\hat{B}^{(d)}_n \;=\; \tilde{B}^{(d)}_{n-k-1} \;\cup\; \bigcup_{s=0}^{k} \hat{B}^{(d)}_{n;s}\;.$$

Aufgrund der Disjunktheit der auftretenden Indexmengen gilt

$$\hat{S}^{(d)}_n \;=\; \tilde{S}^{(d)}_{n-k-1} \;\oplus\; \bigoplus_{s=0}^{k}\Big(\bigoplus_{(i_1,...,i_d)\in\hat{B}^{(d)}_{n;s}} T_{i_1,...,i_d}\Big)$$

und somit wegen $dim(T_{i_1,...,i_d}) = 2^{n-1-s} \;\; \forall (i_1,...,i_d)$ mit $i_1+...+i_d = n+d-1-s$

$$dim(\hat{S}^{(d)}_n) \;=\; dim(\tilde{S}^{(d)}_{n-k-1}) \;+\; \sum_{s=0}^{k} |\hat{B}^{(d)}_{n;s}|\cdot 2^{n-1-s}\;.$$

Zur Berechnung von $|\hat{B}^{(d)}_{n;s}|$ schreiben wir zunächst die Mengen $\hat{B}^{(d)}_{n;s}$ in etwas anderer Weise:

$$\begin{aligned}\hat{B}^{(d)}_{n;s} &= \Big\{(i_1,...,i_d):\; \sum_{j=1}^{d} i_j = n+d-1-s \;\wedge\; \sum_{j=1}^{d} i_j + \frac{1}{2}\min_{1\le j\le d}\sum_{l\ne j} i_l \;\le\; n+\frac{3}{2}\cdot d-1\Big\}\\ &= \Big\{(i_1,...,i_d):\; \sum_{j=1}^{d} i_j = n+d-1-s \;\wedge\; \min_{1\le j\le d}\sum_{l\ne j} i_l \;\le\; d+2s\Big\}.\end{aligned}$$

<u>Behauptung 1:</u> $\forall\, (i_1,...,i_d)\in\hat{B}^{(d)}_{n;s}$ gilt: $\exists\, j,\; 1\le j\le d,$ mit $i_j > i_l \;\; \forall\, l\ne j.$
*Beweis:* Wir führen die Annahme zum Widerspruch, daß $\hat{B}^{(d)}_{n;s}$ ein $d$-Tupel mit mehrfach auftretendem maximalen Eintrag enthält. Sei also $i_j \le c\; \forall j$ sowie o.B.d.A. $i_1 = i_2 = c$. Dann muß $c \le \frac{n+d-1-s-(d-2)}{2} = \frac{n-s+1}{2}$ gelten, und wir erhalten

$$\begin{aligned} & (i_1,...,i_d) \;\in\; \hat{B}^{(d)}_{n;s}\\ \Rightarrow\quad & \min_{1\le j\le d}\sum_{l\ne j} i_l \;=\; n+d-1-s-c \;\le\; d+2s\\ \Rightarrow\quad & n \;\le\; 3s+c+1 \;\le\; \frac{n}{2}+\frac{5s}{2}+\frac{3}{2}\\ \Rightarrow\quad & n \;\le\; 5s+3 \;\le\; 5k+3 \;=\; n-1-r \qquad (\text{da }\; 0\le s\le k)\\ \Rightarrow\quad & n \;\le\; n-1,\end{aligned}$$

womit Behauptung 1 bewiesen ist. Somit gilt

$$\begin{aligned}|\hat{B}^{(d)}_{n;s}| \;&=\; \sum_{j=1}^{d}\Big|\Big\{(i_1,...,i_d)\in\hat{B}^{(d)}_{n;s}:\;\; i_j \;=\; \max_{1\le l\le d} i_l\Big\}\Big|\\ &=\; d\cdot\Big|\Big\{(i_1,...,i_d)\in\hat{B}^{(d)}_{n;s}:\;\; i_1 \;=\; \max_{1\le l\le d} i_l\Big\}\Big|\\ &=\; d\cdot\Big|\Big\{(i_1,...,i_d):\; \sum_{j=1}^{d} i_j = n+d-1-s \;\wedge\; \sum_{j=2}^{d} i_j \;\le\; d+2s\Big\}\Big|,\end{aligned}$$

da aus $\sum_{j=1}^{d} i_j = n + d - 1 - s$ und $\sum_{j=2}^{d} i_j \leq d + 2s$ nach kurzer Rechnung $i_1 = \max_{1 \leq l \leq d} i_l$ folgt.
Behauptung 2: Es gibt $\binom{n}{k}$ $k$-Tupel natürlicher Zahlen, deren Summe kleiner oder gleich $n$ ist.
*Beweis:* Siehe den Beweis von Lemma 3.4.
Mit diesem Resultat erhalten wir sofort

$$|\hat{B}_{n;s}^{(d)}| \;=\; d \cdot \left|\left\{(i_1, ..., i_d) : \sum_{j=1}^{d} i_j = n + d - 1 - s \;\wedge\; \sum_{j=2}^{d} i_j \;\leq\; d + 2s\right\}\right| \;=\; d \cdot \binom{d+2s}{d-1},$$

woraus schließlich

$$dim(\hat{S}_n^{(d)}) \;=\; dim(\tilde{S}_{n-k-1}^{(d)}) \;+\; \sum_{s=0}^{k} d \cdot \binom{d+2s}{d-1} \cdot 2^{n-1-s}$$

folgt. Es bleibt noch zu zeigen, daß dieser Ausdruck von der Ordnung $O(h_n^{-1})$ ist. Nach (18) gilt

$$\begin{aligned} dim(\tilde{S}_{n-k-1}^{(d)}) &= O(2^{n-k-1} \cdot (n-k-1)^{d-1}) \\ &= O(2^n \cdot \frac{(n-k-1)^{d-1}}{2^{k+1}}) \\ &= O(h_n^{-1}). \end{aligned}$$

Für die Ordnung der Summe ergibt sich

$$\begin{aligned} \sum_{s=0}^{k} d \cdot \binom{d+2s}{d-1} \cdot 2^{n-1-s} &= d \cdot 2^{n-1} \cdot \sum_{s=0}^{k} \binom{d+2s}{d-1} \cdot 2^{-s} \\ &= d \cdot 2^{n-1} \cdot \sum_{s=0}^{k} O(s^{d-1}) \cdot 2^{-s} \\ &= O(h_n^{-1}). \end{aligned}$$

□

(32) gestattet die Berechnung der Dimension unseres neuen Dünn-Gitter-Raumes $\hat{S}_n^{(d)}$ bei beliebigem $d$. Für die Spezialfälle $d = 2$ und $d = 3$ geben wir die resultierende Formel explizit an.

KOROLLAR 4.3. *Sei wiederum $n - 4 = 5k + r,\; 0 \leq r \leq 4$. Dann gilt*

$$\begin{aligned} dim(\hat{S}_n^{(2)}) &= 8 \cdot 2^n \;+\; 1 \;-\; (10 - r) \cdot 2^{n-1-k}, \\ dim(\hat{S}_n^{(3)}) &= 42 \cdot 2^n \;-\; 1 \;+\; \left(k^2 + \frac{4r - 29}{2}k + \frac{r^2 + 5r}{4} - \frac{71}{2}\right) \cdot 2^{n-k}. \end{aligned} \tag{33}$$

*Beweis.* Aus (32) folgt im Falle $d = 2$

$$dim(\hat{S}_n^{(2)}) \;=\; dim(\tilde{S}_{n-k-1}^{(2)}) \;+\; \sum_{s=0}^{k} 2 \cdot \binom{2+2s}{1} \cdot 2^{n-1-s}$$

$$
\begin{aligned}
&= 2^{n-k-1} \cdot (n-k-2) + 1 \ + \ 2^{n+1} \cdot \sum_{s=0}^{k} 2^{-s} \ + \ 2^{n+1} \cdot \sum_{s=0}^{k} s \cdot 2^{-s} \\
&= 2^{n-k-1} \cdot (n-k-2) + 1 \ + \ 2^{n+1} \cdot \Big(2 - 2^{-k} \ + \ 2 - (k+2) \cdot 2^{-k}\Big) \\
&= 8 \cdot 2^n \ + \ 1 \ + (n-k-2-4-4k-8) \cdot 2^{n-1-k} \\
&= 8 \cdot 2^n \ + \ 1 \ - (10-r) \cdot 2^{n-1-k}.
\end{aligned}
$$

Für $d = 3$ liefert (32)

$$
\begin{aligned}
dim(\hat{S}_n^{(3)}) &= dim(\tilde{S}_{n-k-1}^{(3)}) \ + \ \sum_{s=0}^{k} 3 \cdot \binom{3+2s}{2} \cdot 2^{n-1-s} \\
&= 2^{n-k-1} \cdot \left( \frac{(n-k-1)^2}{2} - \frac{n-k-1}{2} + 1 \right) - 1 \\
&\qquad + 3 \cdot 2^{n-2} \cdot \left( 4 \cdot \sum_{s=0}^{k} s^2 \cdot 2^{-s} \ + \ 10 \cdot \sum_{s=0}^{k} s \cdot 2^{-s} \ + \ 6 \cdot \sum_{s=0}^{k} 2^{-s} \right) \\
&= 2^{n-k-2} \cdot \Big(n^2 - (2k+3) \cdot n + k^2 + 3k + 4\Big) - 1 \\
&\qquad + 3 \cdot 2^{n-2} \cdot \Big(4 \cdot \Big(6 - 2^{-k} \cdot (k^2 + 4k + 6)\Big) \\
&\qquad\qquad + 10 \cdot \Big(2 - 2^{-k} \cdot (k+2)\Big) \ + \ 6 \cdot (2 - 2^{-k})\Big) \\
&= 2^{n-k-2} \cdot \Big(n^2 - (2k+3) \cdot n + k^2 + 3k + 4\Big) - 1 \\
&\qquad + 3 \cdot 2^{n-2} \cdot \Big(2^{-k} \cdot (-4k^2 - 26k - 50) \ + \ 56\Big) \\
&= 42 \cdot 2^n \ - \ 1 \\
&\qquad + \ \left( \frac{n^2 - (2k+3) \cdot n + k^2 + 3k + 4}{4} - 3k^2 - 19.5k - 37.5 \right) \cdot 2^{n-k} \\
&= 42 \cdot 2^n \ - \ 1 \ + \ \left( k^2 + \frac{4r-29}{2} k + \frac{r^2 + 5r}{4} - \frac{71}{2} \right) \cdot 2^{n-k},
\end{aligned}
$$

da $n = 4 + 5k + r$. $\square$

Abbildung 7 zeigt die Situation für $d = 3$ und $n = 13$. Wegen $n - 4 = 9 = 1 \cdot 5 + 4$ gilt $k = 1$ und $r = 4$. Aufgrund von Definition (29) können aus $\tilde{B}_{13}^{(3)}$ nur Tripel $(i_1, i_2, i_3)$ mit $i_1 + i_2 + i_3 \in \{14, 15\}$ entfernt werden, um zu $\hat{B}_{13}^{(3)}$ zu gelangen. Das Diagramm der Tripel $(i_1, i_2, i_3)$ beziehungsweise der zugeordneten Teilräume $T_{i_1,i_2,i_3}$, das bei $\tilde{B}_{13}^{(3)}$ beziehungsweise $\tilde{S}_{13}^{(3)}$ ein Simplex darstellt, wird also — anschaulich gesprochen — von der durch die drei Punkte $(13, 1, 1)$, $(1, 13, 1)$ und $(1, 1, 13)$ charakterisierten Fläche aus in Richtung auf den Ursprung eingedrückt.

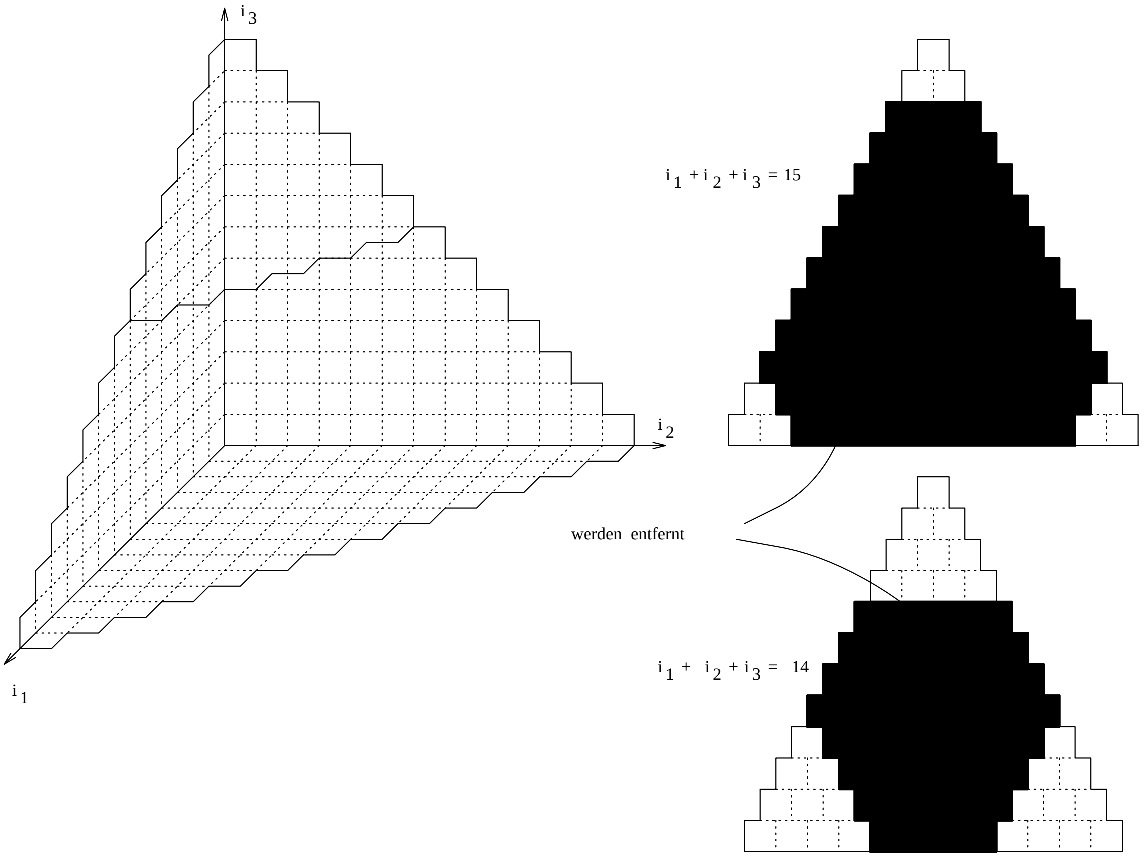



ABB. 7. *Der Übergang von $\tilde{B}_{13}^{(3)}$ bzw. $\tilde{S}_{13}^{(3)}$ (links) zu $\hat{B}_{13}^{(3)}$ bzw. $\hat{S}_{13}^{(3)}$.*

Abschließend sollen noch einige Zahlenbeispiele in Tabelle 2 die Speicherplatzvorteile des neuen Dünn-Gitter-Raumes $\hat{S}_n^{(3)}$ gegenüber $\tilde{S}_n^{(3)}$ hervorheben.

TABELLE 2
*Raumdimensionen im Vergleich, $d = 3$.*

| $n$ | $dim(S_n^{(3)})$ | $dim(\tilde{S}_n^{(3)})$ | $dim(\hat{S}_n^{(3)})$ |
|---|---|---|---|
| 5 | 29791 | 351 | 255 |
| 6 | 250047 | 1023 | 639 |
| 7 | $2.0 \cdot 10^{06}$ | 2815 | 1599 |
| 8 | $1.7 \cdot 10^{07}$ | 7423 | 3967 |
| 9 | $1.3 \cdot 10^{08}$ | 18943 | 8959 |
| 10 | $1.1 \cdot 10^{09}$ | 47103 | 19711 |
| 15 | $3.5 \cdot 10^{13}$ | $3.5 \cdot 10^{6}$ | 925695 |
| 20 | $1.2 \cdot 10^{18}$ | $2.0 \cdot 10^{8}$ | $3.6 \cdot 10^{7}$ |

Es ist klar, daß sich bei $d = 3$ der Wegfall des logarithmischen Faktors $(\mathrm{ld}(h_n^{-1}))^{d-1}$ erst bei größeren $n$ signifikant auswirkt. Mit zunehmendem Dimensionsparameter $d$ wird die Ersparnis an Gitterpunkten jedoch bereits bei kleinen Werten für $n$ spürbar.

**4.2. Der Interpolationsfehler im Raum $\hat{\mathbf{S}}_{\mathbf{n}}^{(\mathbf{d})}$.** Für die in der Energienorm gemessene Abweichung $u^{(d)} - \hat{u}_{n,I}^{(d)}$ des Interpolanten $\hat{u}_{n,I}^{(d)}$ von $u^{(d)} \in X_0^{(d)}$ in den Raum $\hat{S}_n^{(d)}$ von $u^{(d)}$ selbst gilt mit $u^{(d)} = \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} u_{i_1,\dots,i_d},\ u_{i_1,\dots,i_d} \in T_{i_1,\dots,i_d}$,

$$
\begin{aligned}
\|u^{(d)} - \hat{u}_{n,I}^{(d)}\|_E &= \Big\| \sum_{i_1=1}^{\infty} \dots \sum_{i_d=1}^{\infty} u_{i_1,\dots,i_d} - \sum_{(i_1,\dots,i_d)\in \hat{B}_n^{(d)}} u_{i_1,\dots,i_d} \Big\|_E \\
&= \Big\| \sum_{(i_1,\dots,i_d)\notin \hat{B}_n^{(d)}} u_{i_1,\dots,i_d} \Big\|_E \\
&\leq \sum_{(i_1,\dots,i_d)\notin \hat{B}_n^{(d)}} \|u_{i_1,\dots,i_d}\|_E \\
&= \sum_{s=-k}^{\infty} \sum_{\substack{i_1+\dots+i_d=n+d-1+s \\ (i_1,\dots,i_d)\notin \hat{B}_n^{(d)}}} \|u_{i_1,\dots,i_d}\|_E \ ,
\end{aligned}
$$

woraus sich mit (28) ergibt:

$$
\begin{aligned}
&\|u^{(d)} - \hat{u}_{n,I}^{(d)}\|_E \\
\leq\ & \frac{\sqrt{d} \cdot |u^{(d)}|_\infty}{2 \cdot 12^{\frac{d-1}{2}}} \cdot \sum_{s=-k}^{\infty} \sum_{\substack{i_1+\dots+i_d=n+d-1+s \\ (i_1,\dots,i_d)\notin \hat{B}_n^{(d)}}} h_{i_1} \cdot \dots \cdot h_{i_d} \cdot \max_{1\leq j\leq d} \left( \frac{h_{i_1} \cdot \dots \cdot h_{i_d}}{h_{i_j}} \right) \\
=\ & \frac{\sqrt{d} \cdot |u^{(d)}|_\infty}{2 \cdot 12^{\frac{d-1}{2}}} \cdot \sum_{s=-k}^{\infty} \sum_{\substack{i_1+\dots+i_d=n+d-1+s \\ (i_1,\dots,i_d)\notin \hat{B}_n^{(d)}}} 2^{-(n+d-1+s)} \cdot \max_{1\leq j\leq d} \left( \frac{h_{i_1} \cdot \dots \cdot h_{i_d}}{h_{i_j}} \right).
\end{aligned} \tag{34}
$$

Ungleichung (34) gestattet für beliebiges, fest gewähltes $d$ die Abschätzung des Interpolationsfehlers bezüglich der Energienorm im Raum $\hat{S}_n^{(d)}$. Wir geben hier der Kürze halber nur die Resultate für die praktisch bedeutsamen Spezialfälle $d = 2$ und $d = 3$ an, bemühen uns aber dafür um eine möglichst genaue Abschätzung der konstanten Faktoren bei $h_n$ in den Fehlerschranken.

Satz 4.4. *Gegeben seien $u^{(d)} \in X_0^{(d)}$ und sein stückweise d-linearer Interpolant $\hat{u}_{n,I}^{(d)} \in \hat{S}_n^{(d)}$ in hierarchischer Darstellung (31). Dann gelten für die Energienorm des Interpolationsfehlers in den Fällen $d = 2$ und $d = 3$ folgende Abschätzungen:*

$$
\begin{aligned}
\|u^{(2)} - \hat{u}_{n,I}^{(2)}\|_E &\leq 2^{-n} \cdot \frac{\sqrt{6} \cdot |u^{(2)}|_\infty}{8} = O(h_n), \\
\|u^{(3)} - \hat{u}_{n,I}^{(3)}\|_E &\leq 2^{-n} \cdot \frac{5 \cdot \sqrt{3} \cdot |u^{(3)}|_\infty}{64} = O(h_n).
\end{aligned} \tag{35}
$$

*Beweis.* Aus (34) folgt im Falle $d = 2$

$$
\|u^{(2)} - \hat{u}_{n,I}^{(2)}\|_E \leq \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \sum_{s=-k}^{\infty} \sum_{\substack{i_1+i_2=n+1+s \\ (i_1,i_2)\notin \hat{B}_n^{(2)}}} 2^{-(n+1+s)} \cdot \max_{1\leq j\leq 2} 2^{-i_j}
$$

$$= \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \sum_{s=-k}^{0} \sum_{\substack{i_1+i_2=n+1+s \\ (i_1,i_2)\notin \hat{B}_n^{(2)}}} 2^{-(n+1+s)} \cdot \max_{1\le j\le 2} 2^{-i_j} \quad +$$
$$\frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \sum_{s=1}^{\infty} \sum_{i_1+i_2=n+1+s} 2^{-(n+1+s)} \cdot \max_{1\le j\le 2} 2^{-i_j},$$

da $(i_1, i_2) \notin \hat{B}_n^{(2)}$, falls $i_1 + i_2 > n+1$ (siehe (29)). Wir summieren also längs der Diagonalen $i_1 + i_2 = n+1+s,\ s = -k, ..., 0, 1, ...$ . Dabei unterscheiden wir zwischen $s > 0$, wo alle Indexpaare $(i_1, i_2)$ mit $i_1 + i_2 = n+1+s$ zu berücksichtigen sind, und $s \le 0$, wo die Indexpaare $(i_1, i_2)$ mit $i_1 + i_2 = n+1+s$ und $\min(i_1, i_2) < 3 - 2s$ in $\hat{B}_n^{(2)}$ liegen und deshalb bei der Summation nicht berücksichtigt werden dürfen (siehe Abbildung 8).

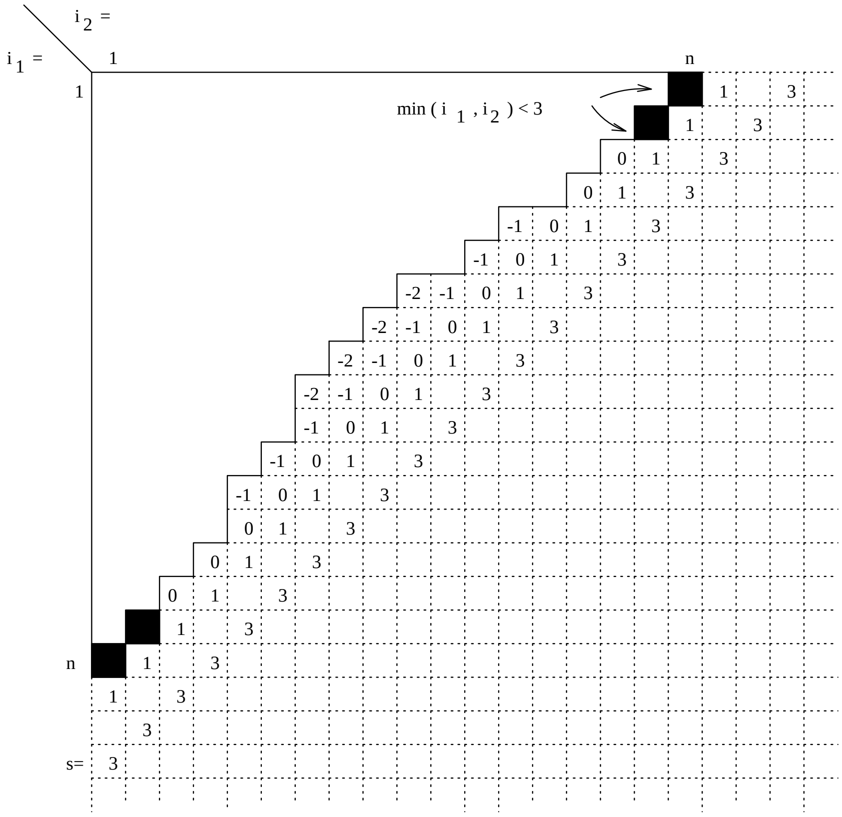



Abb. 8. *Summationsgebiete bei der Berechnung des Interpolationsfehlers für $d = 2$.*

Somit erhalten wir

$$\|u^{(2)} - \hat{u}_{n,I}^{(2)}\|_E \le \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \sum_{s=-k}^{0} 2^{-(n+1+s)} \cdot 2 \cdot \left(\frac{1}{2^{3-2s}} + \frac{1}{2^{3-2s+1}} + \frac{1}{2^{3-2s+2}} + \dots\right) \quad +$$
$$\frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \sum_{s=1}^{\infty} 2^{-(n+1+s)} \cdot 2 \cdot \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots\right)$$
$$= \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \left(\sum_{s=-k}^{0} 2^{-(n+3-s)} \cdot \left(1 + \frac{1}{2} + \frac{1}{4} + \dots\right) \quad + \quad \sum_{s=1}^{\infty} 2^{-(n+s)}\right)$$

$$
\begin{aligned}
&= 2^{-n} \cdot \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \left( \frac{1}{4} \cdot \sum_{s=-k}^{0} 2^s \quad + \quad \sum_{s=1}^{\infty} 2^{-s} \right) \\
&\leq 2^{-n} \cdot \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \left( \frac{1}{4} \cdot 2 \quad + \quad 1 \right) \\
&= 2^{-n} \cdot \frac{\sqrt{6} \cdot |u^{(2)}|_\infty}{8}.
\end{aligned}
$$

Für $d = 3$ liefert (34):

$$
\begin{aligned}
\|u^{(3)} - \hat{u}_{n,I}^{(3)}\|_E \quad &\leq \quad \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{24} \cdot \sum_{s=-k}^{\infty} \sum_{\substack{i_1+i_2+i_3=n+2+s \\ (i_1,i_2,i_3) \notin \hat{B}_n^{(3)}}} 2^{-(n+2+s)} \cdot \max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l} \\
&= \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{24} \cdot \sum_{s=-k}^{0} \sum_{\substack{i_1+i_2+i_3=n+2+s \\ (i_1,i_2,i_3) \notin \hat{B}_n^{(3)}}} 2^{-(n+2+s)} \cdot \max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l} \quad + \\
&\quad \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{24} \cdot \sum_{s=1}^{\infty} \sum_{i_1+i_2+i_3=n+2+s} 2^{-(n+2+s)} \cdot \max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l} \\
&= \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{24} \cdot \sum_{s=-k}^{0} \left( 2^{-(n+2+s)} \cdot \sum_{\substack{i_1+i_2+i_3=n+2+s \\ (i_1,i_2,i_3) \notin \hat{B}_n^{(3)}}} \max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l} \right) \quad + \\
&\quad \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{24} \cdot \sum_{s=1}^{\infty} \left( 2^{-(n+2+s)} \cdot \sum_{i_1+i_2+i_3=n+2+s} \max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l} \right),
\end{aligned}
$$

da $(i_1, i_2, i_3) \notin \hat{B}_n^{(3)}$, falls $i_1 + i_2 + i_3 > n + 2$ (siehe (29)). Analog zum zweidimensionalen Fall summieren wir über die Ebenen $i_1 + i_2 + i_3 = n + 2 + s,\ \ s = -k, ..., 0, 1, ...$ (vgl. Abbildung 7). Dabei unterscheiden wir zwischen $s > 0$, wo alle Indextripel $(i_1, i_2, i_3)$ mit $i_1 + i_2 + i_3 = n + 2 + s$ zu berücksichtigen sind, und $s \leq 0$, wo die Indextripel $(i_1, i_2, i_3)$ mit $i_1 + i_2 + i_3 = n + 2 + s$ und $\min(i_1 + i_2, i_1 + i_3, i_2 + i_3) < 4 - 2s$ in $\hat{B}_n^{(3)}$ liegen und deshalb bei der Summation nicht berücksichtigt werden dürfen (siehe Abbildung 9).

Als nächstes wenden wir uns der Ermittlung der Werte obiger max-Terme

$$
\max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l}, \qquad i_1 + i_2 + i_3 = n + 2 + s, \quad (i_1, i_2, i_3) \notin \hat{B}_n^{(3)},
$$

bei festem $s$ zu. Offensichtlich gibt es drei Tripel $(i_1, i_2, i_3)$ mit $i_1 + i_2 + i_3 = n + 2 + s$, für die der max-Term den Wert $\frac{1}{4}$ annimmt (nämlich $(1, 1, n + s)$, $(1, n + s, 1)$ und $(n + s, 1, 1)$), sechs Tripel mit $\frac{1}{8}$, neun mit $\frac{1}{16}$ usw.. Daraus folgt für $s \geq 1$

$$
\begin{aligned}
\sum_{i_1+i_2+i_3=n+2+s} \max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l} \quad &\leq \quad \sum_{i=2}^{\infty} 3 \cdot (i-1) \cdot 2^{-i} \\
&= \frac{3}{2} \cdot \sum_{i=1}^{\infty} i \cdot 2^{-i} \\
&= 3
\end{aligned}
$$

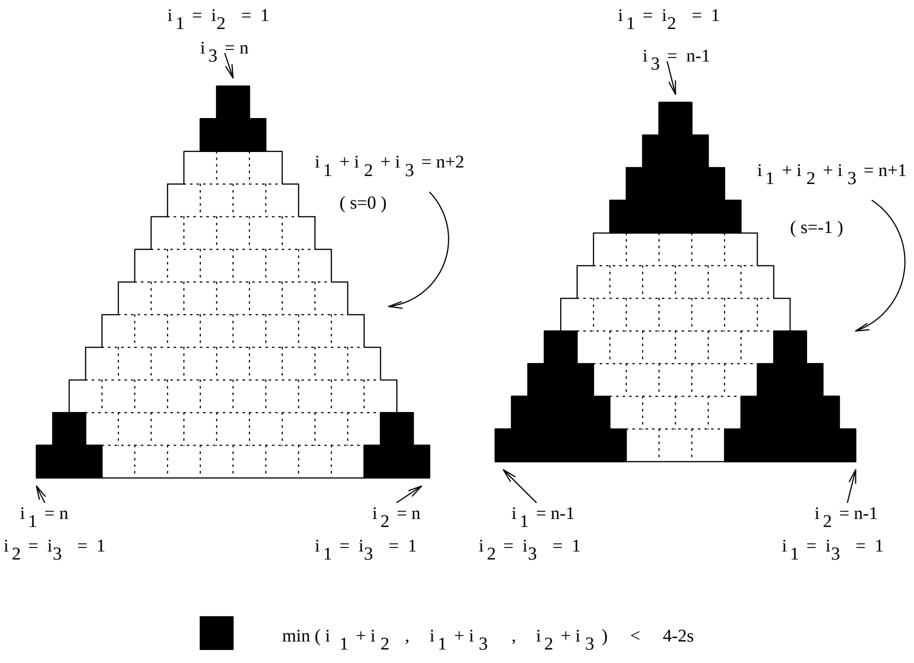



ABB. 9. *Summationsgebiete bei der Berechnung des Interpolationsfehlers für* $d = 3$.

sowie für $s \leq 0$

$$
\begin{aligned}
\sum_{\substack{i_1+i_2+i_3=n+2+s \\ (i_1,i_2,i_3)\notin \hat{B}_n^{(3)}}} \max_{\substack{1 \leq k,l \leq 3 \\ k \neq l}} 2^{-i_k-i_l} &\leq \sum_{i=4-2s}^{\infty} 3 \cdot (i-1) \cdot 2^{-i} \\
&= \frac{3}{2} \cdot \sum_{i=3-2s}^{\infty} i \cdot 2^{-i} \\
&= \frac{3}{2} \cdot \left( \sum_{i=1}^{\infty} i \cdot 2^{-i} - \sum_{i=1}^{2-2s} i \cdot 2^{-i} \right) \\
&= \frac{3}{2} \cdot 2^{2s-2} \cdot (4-2s) \\
&= 3 \cdot 2^{2s-2} \cdot (2-s).
\end{aligned}
$$

Somit ergibt sich

$$
\begin{aligned}
\|u^{(3)} - \hat{u}_{n,I}^{(3)}\|_E &\leq \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{24} \cdot \sum_{s=-k}^{0} 3 \cdot 2^{2s-2} \cdot (2-s) \cdot 2^{-(n+2+s)} \quad + \\
&\quad \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{24} \cdot \sum_{s=1}^{\infty} 3 \cdot 2^{-(n+2+s)} \\
&= 2^{-n} \cdot \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{32} \cdot \left( \sum_{s=-k}^{0} (2-s) \cdot 2^{s-2} \quad + \quad \sum_{s=1}^{\infty} 2^{-s} \right) \\
&= 2^{-n} \cdot \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{32} \cdot \left( \sum_{s=2}^{k+2} s \cdot 2^{-s} \quad + \quad 1 \right)
\end{aligned}
$$

$$\leq \quad 2^{-n} \cdot \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{32} \cdot \left( \sum_{s=2}^{\infty} s \cdot 2^{-s} \quad + \quad 1 \right)$$
$$= \quad 2^{-n} \cdot \frac{5 \cdot \sqrt{3} \cdot |u^{(3)}|_\infty}{64}.$$

□

Damit ist gezeigt, daß unsere neuen Dünn-Gitter-Räume $\hat{S}_n^{(d)}$ für $d = 2$ und $d = 3$ die Repräsentation einer Funktion $u^{(d)} \in X_0^{(d)}$ mit einem Interpolationsfehler $\|u^{(d)} - \hat{u}_{n,I}^{(d)}\|_E$ der Ordnung $O(h_n)$ gestatten. Dies entspricht einem Fehler derselben Größenordnung wie bei Verwendung eines vollen Gitters, obwohl die Anzahl der benötigten Gitterpunkte signifikant kleiner ist, nämlich von der Ordnung $O(h_n^{-1})$ statt $O(h_n^{-d})$. Auch wenn jedoch beim Übergang vom vollen Gitter $G_n^{(d)}$ auf das dünne Gitter $\hat{G}_n^{(d)}$ die Ordnung des Interpolationsfehlers erhalten bleibt, so wird dessen Wert aufgrund der Verringerung der Anzahl der Unbekannten dennoch steigen. Das heißt, der konstante Faktor vor dem führenden $h_n$-Term der Abschätzung des Interpolationsfehlers wird zunehmen. Um diesen Zuwachs beschreiben zu können, geben wir zum Schluß noch Schranken für den in der Energienorm gemessenen Interpolationsfehler auf $G_n^{(d)}$ sowie dem dünnen Gitter $\tilde{G}_n^{(d)}$ gemäß Abschnitt 3 an und vergleichen die Faktoren vor den $h_n$-Termen der Abschätzungen.

Satz 4.5. *Gegeben seien $u^{(d)} \in X_0^{(d)}$ und seine stückweise d-linearen Interpolanten $u_{n,I}^{(d)} \in S_n^{(d)}$ beziehungsweise $\tilde{u}_{n,I}^{(d)} \in \tilde{S}_n^{(d)}$ in hierarchischer Darstellung gemäß (7) beziehungsweise (16). Dann gelten für die Energienorm des Interpolationsfehlers in den Fällen $d = 2$ und $d = 3$ folgende Abschätzungen:*

$$\begin{aligned}
\|u^{(2)} - u_{n,I}^{(2)}\|_E &\leq h_n \cdot \frac{|u^{(2)}|_\infty}{3 \cdot \sqrt{6}} \; + \; O(h_n^2), \\
\|u^{(3)} - u_{n,I}^{(3)}\|_E &\leq h_n \cdot \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{72} \; + \; O(h_n^3), \\
\|u^{(2)} - \tilde{u}_{n,I}^{(2)}\|_E &\leq h_n \cdot \frac{|u^{(2)}|_\infty}{\sqrt{24}}, \\
\|u^{(3)} - \tilde{u}_{n,I}^{(3)}\|_E &\leq h_n \cdot \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{32}.
\end{aligned} \tag{36}$$

*Beweis.* Wir zeigen zunächst die Aussagen für $\tilde{u}_{n,I}^{(d)}$. Analog zur Herleitung von (34) gilt mit (28)

$$\begin{aligned}
& \|u^{(d)} - \tilde{u}_{n,I}^{(d)}\|_E \\
= \; & \Big\| \sum_{i_1=1}^{\infty} \cdots \sum_{i_d=1}^{\infty} u_{i_1,\ldots,i_d} \; - \sum_{(i_1,\ldots,i_d) \in \tilde{B}_n^{(d)}} u_{i_1,\ldots,i_d} \Big\|_E \\
\leq \; & \sum_{(i_1,\ldots,i_d) \notin \tilde{B}_n^{(d)}} \|u_{i_1,\ldots,i_d}\|_E \\
= \; & \sum_{s=1}^{\infty} \; \sum_{i_1+\ldots+i_d=n+d-1+s} \|u_{i_1,\ldots,i_d}\|_E
\end{aligned}$$

$$\leq \frac{\sqrt{d}\cdot|u^{(d)}|_\infty}{2\cdot 12^{\frac{d-1}{2}}}\cdot\sum_{s=1}^{\infty}\sum_{i_1+\ldots+i_d=n+d-1+s} 2^{-(n+d-1+s)}\cdot\max_{1\leq j\leq d}\left(\frac{h_{i_1}\cdot\ldots\cdot h_{i_d}}{h_{i_j}}\right),$$

woraus für die Spezialfälle $d=2$ und $d=3$ folgt:

$$\|u^{(2)}-\tilde{u}^{(2)}_{n,I}\|_E \leq \frac{|u^{(2)}|_\infty}{\sqrt{24}}\cdot\sum_{s=1}^{\infty}\sum_{i_1+i_2=n+1+s} 2^{-(n+1+s)}\cdot\max_{1\leq j\leq 2} 2^{-i_j},$$

$$\|u^{(3)}-\tilde{u}^{(3)}_{n,I}\|_E \leq \frac{\sqrt{3}\cdot|u^{(3)}|_\infty}{24}\cdot\sum_{s=1}^{\infty}\sum_{i_1+i_2+i_3=n+2+s} 2^{-(n+2+s)}\cdot\max_{\substack{1\leq k,l\leq 3\\ k\neq l}} 2^{-i_k-i_l}.$$

Die auftretenden max-Terme wurden bereits im Beweis von Satz 4.4 abgeschätzt. Mit den dortigen Resultaten erhält man

$$\|u^{(2)}-\tilde{u}^{(2)}_{n,I}\|_E \leq \frac{|u^{(2)}|_\infty}{\sqrt{24}}\cdot\sum_{s=1}^{\infty}2^{-(n+s)} = \frac{|u^{(2)}|_\infty}{\sqrt{24}}\cdot 2^{-n}$$

sowie

$$\|u^{(3)}-\tilde{u}^{(3)}_{n,I}\|_E \leq \frac{\sqrt{3}\cdot|u^{(3)}|_\infty}{32}\cdot\sum_{s=1}^{\infty}2^{-(n+s)} = \frac{\sqrt{3}\cdot|u^{(3)}|_\infty}{32}\cdot 2^{-n}.$$

Für $u^{(d)}_{n,I}$ nun gilt mit (28)

$$\begin{aligned}
\|u^{(d)}-u^{(d)}_{n,I}\|_E &= \Big\|\sum_{i_1=1}^{\infty}\ldots\sum_{i_d=1}^{\infty}u_{i_1,\ldots,i_d} - \sum_{(i_1,\ldots,i_d)\in B_n^{(d)}}u_{i_1,\ldots,i_d}\Big\|_E\\
&\leq \sum_{(i_1,\ldots,i_d)\notin B_n^{(d)}}\|u_{i_1,\ldots,i_d}\|_E\\
&\leq \frac{\sqrt{d}\cdot|u^{(d)}|_\infty}{2\cdot 12^{\frac{d-1}{2}}}\cdot\sum_{\max(i_j)>n}2^{-i_1-\ldots-i_d}\cdot\max_{1\leq j\leq d}\left(\frac{h_{i_1}\cdot\ldots\cdot h_{i_d}}{h_{i_j}}\right).
\end{aligned}\tag{37}$$

Im Falle $d=2$ zerlegen wir den Summationsbereich $\mathbb{N}^2\setminus B_n^{(2)}$ in $\mathbb{N}^2\setminus\tilde{B}_{2n}^{(2)}$ und $\tilde{B}_{2n}^{(2)}\setminus B_n^{(2)}$ und erhalten somit

$$\begin{aligned}
\|u^{(2)}-u^{(2)}_{n,I}\|_E &\leq \frac{|u^{(2)}|_\infty}{\sqrt{24}}\cdot\Bigg(\sum_{s=1}^{\infty}\sum_{i_1+i_2=2n+1+s}2^{-(2n+1+s)}\cdot\max_{1\leq j\leq 2}2^{-i_j} +\\
&\qquad 2\cdot\sum_{s=1}^{n}\sum_{i_1+i_2=s+1}2^{-(n+1+s)}\cdot 2^{-i_2}\Bigg)\\
&\leq \frac{|u^{(2)}|_\infty}{\sqrt{24}}\cdot\Bigg(\sum_{s=1}^{\infty}2^{-(2n+1+s)}\cdot 2\cdot\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots\right)+\\
&\qquad \sum_{s=1}^{n}2^{-(n+s)}\cdot\sum_{i_1+i_2=s+1}2^{-i_2}\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \left( \sum_{s=1}^{\infty} 2^{-(2n+s)} \;+\; \sum_{s=1}^{n} 2^{-(n+s)} \cdot \left(1 - 2^{-s}\right) \right) \\
&= \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \left(2^{-2n} \;+\; 2^{-n} \cdot \left( \sum_{s=1}^{n} 2^{-s} \;-\; \sum_{s=1}^{n} 2^{-2s} \right)\right) \\
&= \frac{|u^{(2)}|_\infty}{\sqrt{24}} \cdot \left(2^{-2n} \;+\; 2^{-n} \cdot \left(\frac{2}{3} - 2^{-n} + \frac{1}{3} \cdot 4^{-n}\right)\right) \\
&= \frac{|u^{(2)}|_\infty}{3 \cdot \sqrt{6}} \cdot h_n \;+\; O(h_n^2).
\end{aligned}
$$

Im Falle $d = 3$ schließlich zerlegen wir den Summationsbereich $\mathbb{N}^3 \setminus B_n^{(3)} = \{(i_1, i_2, i_3) : \max(i_j) > n\}$ in die Teilbereiche I (alle drei Indizes größer als $n$), II (genau zwei Indizes größer als $n$) sowie III (genau ein Index größer als $n$). Die Bereiche II und III zerfallen ihrerseits aus Symmetriegründen in je drei gleich große Teilbereiche (welcher Index ist kleiner beziehungsweise größer als $n$?). Damit gilt für Teilbereich I

$$
\begin{aligned}
\sum_{I} 2^{-i_1-i_2-i_3} \cdot \max_{\substack{1 \le k,l \le 3 \\ k \ne l}} 2^{-i_k-i_l} &= \sum_{i_1=n+1}^{\infty} \sum_{i_2=n+1}^{\infty} \sum_{i_3=n+1}^{\infty} 2^{-i_1-i_2-i_3} \cdot \max_{\substack{1 \le k,l \le 3 \\ k \ne l}} 2^{-i_k-i_l} \\
&\le \sum_{i_1=n+1}^{\infty} \sum_{i_2=n+1}^{\infty} \sum_{i_3=n+1}^{\infty} 2^{-i_1-i_2-i_3} \cdot 2^{-2n} \\
&= 2^{-5n}.
\end{aligned}
$$

Für Teilbereich II erhalten wir

$$
\begin{aligned}
\sum_{II} 2^{-i_1-i_2-i_3} \cdot \max_{\substack{1 \le k,l \le 3 \\ k \ne l}} 2^{-i_k-i_l} &= 3 \cdot \sum_{i_1=n+1}^{\infty} \sum_{i_2=n+1}^{\infty} \sum_{i_3=1}^{n} 2^{-i_1-i_2-i_3} \cdot \max_{\substack{1 \le k,l \le 3 \\ k \ne l}} 2^{-i_k-i_l} \\
&\le 3 \cdot \sum_{i_1=n+1}^{\infty} \sum_{i_2=n+1}^{\infty} \sum_{i_3=1}^{n} 2^{-i_1-i_2-i_3} \cdot 2^{-i_3-n} \\
&= 3 \cdot 2^{-3n} \cdot \sum_{i_3=1}^{n} 4^{-i_3} \\
&\le 2^{-3n},
\end{aligned}
$$

für Teilbereich III schließlich ergibt sich

$$
\begin{aligned}
\sum_{III} 2^{-i_1-i_2-i_3} \cdot \max_{\substack{1 \le k,l \le 3 \\ k \ne l}} 2^{-i_k-i_l} &= 3 \cdot \sum_{i_1=n+1}^{\infty} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} 2^{-i_1-i_2-i_3} \cdot \max_{\substack{1 \le k,l \le 3 \\ k \ne l}} 2^{-i_k-i_l} \\
&= 3 \cdot \sum_{i_1=n+1}^{\infty} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} 2^{-i_1-i_2-i_3} \cdot 2^{-i_2-i_3} \\
&= 3 \cdot 2^{-n} \cdot \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} 4^{-i_2-i_3} \\
&\le \frac{1}{3} \cdot 2^{-n}.
\end{aligned}
$$

Zusammen mit (37) folgt nun

$$\begin{aligned}
\|u^{(3)} - u^{(3)}_{n,I}\|_E &\leq \frac{\sqrt{3}\cdot|u^{(3)}|_\infty}{24} \cdot \sum_{I,II,III} 2^{-i_1-i_2-i_3} \cdot \max_{\substack{1\,\leq\, k,l\,\leq\, 3\\ k\,\neq\, l}} 2^{-i_k-i_l} \\
&\leq \frac{\sqrt{3}\cdot|u^{(3)}|_\infty}{24} \cdot \left(2^{-5n} + 2^{-3n} + \frac{1}{3}\cdot 2^{-n}\right) \\
&= \frac{\sqrt{3}\cdot|u^{(3)}|_\infty}{72} \cdot 2^{-n} \;+\; O(h_n^3).
\end{aligned}$$

□

Tabelle 3 stellt die Faktoren gegenüber, mit denen $h_n$ in den Schranken aus Satz 4.4 und Satz 4.5 multipliziert wird. Wie erwartet sind diese Faktoren bei Verwendung des vollen Gitters $G_n^{(d)}$ am kleinsten und bei Verwendung des dünnen Gitters $\hat{G}_n^{(d)}$ am größten. Da wir hier aber nur Schranken für die Energienorm des Interpolationsfehlers und nicht dessen Energienorm selbst betrachtet haben, liefert Tabelle 3 lediglich einen Hinweis auf die tatsächliche Zunahme des Fehlers bei Verwendung eines dünnen Gitters. Im zweidimensionalen Fall ist nach Tabelle 3 der Faktor vor $h_n$ beim dünnen Gitter $\hat{G}_n^{(d)}$ gut zweimal so groß wie beim vollen Gitter. Dies kann als Indiz dafür gelten, daß auch der tatsächliche Interpolationsfehler beim Übergang von $G_n^{(d)}$ zu $\hat{G}_n^{(d)}$ nur unwesentlich zunimmt. Im dreidimensionalen Fall nimmt der Faktor vor $h_n$ bei $\hat{G}_n^{(d)}$ einen etwa sechsmal größeren Wert an als bei $G_n^{(d)}$, was auf eine stärkere Zunahme des Interpolationsfehlers hinweist, als dies bei $d = 2$ der Fall ist. An dieser Stelle sei jedoch daran erinnert, daß im Falle $d = 3$ die Anzahl der benötigten Gitterpunkte beim Übergang von $G_n^{(d)}$ zu $\hat{G}_n^{(d)}$ erheblich stärker abnimmt als im zweidimensionalen Fall, da der Unterschied zwischen $O(h_n^{-d})$ und $O(h_n^{-1}\cdot(\mathrm{ld}(h_n^{-1}))^{d-1})$ mit wachsendem $d$ immer größer wird.

TABELLE 3

*Faktoren vor $h_n$ in den Schranken für die Energienorm des Interpolationsfehlers.*

| $d$ | $G_n^{(d)}$ | $\tilde{G}_n^{(d)}$ | $\hat{G}_n^{(d)}$ |
|---|---|---|---|
| 2 | $0.136\cdot\|u^{(2)}\|_\infty$ | $0.204\cdot\|u^{(2)}\|_\infty$ | $0.306\cdot\|u^{(2)}\|_\infty$ |
| 3 | $0.024\cdot\|u^{(3)}\|_\infty$ | $0.054\cdot\|u^{(3)}\|_\infty$ | $0.135\cdot\|u^{(3)}\|_\infty$ |

Schließlich wollen wir noch kurz darauf eingehen, wie sich die Verwendung des dünnen Gitters $\hat{G}_n^{(d)}$ auf die $L_2$- beziehungsweise Maximumsnorm des Interpolationsfehlers auswirkt. Dazu erinnern wir uns, daß $\hat{G}_n^{(d)}$ gemäß Definition das dünne Gitter $\tilde{G}_{n-k-1}^{(d)}$, $n-4=5k+r,\ 0\leq r\leq 4$, enthält (siehe (29)). Wegen $n-k-1\approx\frac{4}{5}\cdot n$ folgt daraus mit Satz 3.5, daß der in der $L_2$- beziehungsweise Maximumsnorm gemessene Interpolationsfehler mindestens von der Ordnung $O(h_n^{1.6}\cdot(\mathrm{ld}(h_n^{-0.8}))^{d-1})$ ist.

Die in Satz 4.2 und Satz 4.4 gezeigten Eigenschaften von $\hat{S}_n^{(d)}$ sind aber nur ein Grund dafür, daß wir uns in diesem Abschnitt eingehender mit der Energienorm und einem darauf basierenden dünnen Gitter auseinandergesetzt haben. Ein weiterer Vorteil eines in der Energienorm gemessenen Interpolationsfehlers wird im zweiten Teil dieser

Arbeit zum Tragen kommen, wenn wir einen Finite-Elemente-Algorithmus zur Lösung von Poisson-Problemen auf dünnen Gittern vorstellen werden.

**5. Adaptivität und Datenstrukturen.** Als Abschluß des ersten Teils dieser Arbeit wollen wir noch auf die Möglichkeit adaptiver Verfeinerung im Zusammenhang mit dünnen Gittern hinweisen und einige Anmerkungen zu den verwendeten Datenstrukturen machen.

**5.1. Reguläre und adaptiv verfeinerte dünne Gitter.** Obwohl die bisher definierten dünnen Gitter $\tilde{G}_n^{(d)}$ und $\hat{G}_n^{(d)}$ Gitter von lokal sehr unterschiedlicher Feinheit darstellen, liegt ihnen beziehungsweise den entsprechenden Funktionenräumen $\tilde{S}_n^{(d)}$ und $\hat{S}_n^{(d)}$ eine regelmäßige Struktur zugrunde. Für den zweidimensionalen Fall zeigt dies Abbildung 5 auf sehr anschauliche Weise. Ob ein Punkt des vollen Gitters $G_n^{(d)}$ auch im jeweiligen dünnen Gitter enthalten ist, hängt dabei nicht von der konkreten Gestalt der darzustellenden Funktion ab. Deshalb bezeichnen wir dünne Gitter gemäß (15) oder (30) auch als *reguläre* dünne Gitter.

Demzufolge sollten (reguläre) dünne Gitter nicht mit adaptiv verfeinerten vollen Gittern verwechselt werden (siehe z.B. MLAT [2], [7]; FAC [30], [31]; PLTMG [3], [4]; KASKADE [9]). Die in den genannten Ansätzen verwendeten Adaptionstechniken führen zu lokal verfeinerten Gittern, welche aber lokal wiederum die Charakteristik eines *vollen* Gitters aufweisen. Verläßt man dagegen den engen Rahmen der bisherigen Definition dünner Gitter und läßt auch für diese die Möglichkeit adaptiver Verfeinerung zu, so erhält man lokal verfeinerte *dünne* Gitter, die auch lokal die typischen Eigenschaften dünner Gitter besitzen. Damit bleibt insbesondere die Eigenschaft, daß dünne Gitter signifikant weniger Punkte enthalten als volle Gitter, auch im wichtigen Fall adaptiver Verfeinerung erhalten.

Wir halten fest, daß dünne Gitter nicht einfach als eine weitere Technik zur adaptiven Verfeinerung voller Gitter angesehen werden dürfen. Vielmehr ist die Unterscheidung zwischen vollen Gittern einerseits und dünnen Gittern andererseits von grundsätzlich anderer Qualität als die Unterscheidung zwischen regulären und adaptiv verfeinerten Gittern.

Nachdem in den vorangegangenen drei Abschnitten reguläre dünne Gitter gemäß (15) und (30) ausführlich besprochen worden sind, wollen wir uns im folgenden noch mit adaptiv verfeinerten dünnen Gittern beschäftigen.

Bei der Konstruktion adaptiv verfeinerter dünner Gitter weichen wir in zweierlei Hinsicht von $\tilde{G}_n^{(d)}$ beziehungsweise $\hat{G}_n^{(d)}$ ab. Zum einen entfallen der Parameter $n$ und damit die Kriterien $(i_1, ..., i_d) \in \tilde{B}_n^{(d)}$ (d.h. $i_1 + ... + i_d \leq n + d - 1$) beziehungsweise $(i_1, ..., i_d) \in \hat{B}_n^{(d)}$, wodurch zum Beispiel im Falle $d = 2$ andere Schemata als die in Abbildung 6 gezeigten auftreten können. Zum anderen wird jetzt nicht mehr für einen gesamten Teilraum $T_{i_1,...,i_d}$ entschieden, ob er bei der Darstellung berücksichtigt werden soll oder nicht. Vielmehr wird für jedes Basiselement des Teilraumes $T_{i_1,...,i_d}$ und seine zugehörige Basisfunktion $g_{k_1,...,k_d}^{(i_1,...,i_d)}$ die Frage der Berücksichtigung gesondert gestellt und beantwortet. Benötigt wird folglich ein Kriterium, anhand dessen wir entscheiden können, ob die Basisfunktion $g_{k_1,...,k_d}^{(i_1,...,i_d)}$ des Teilraumes $T_{i_1,...,i_d}$ zur Repräsentation herangezogen werden soll oder nicht, ob also der zugehörige Gitterpunkt $(m_{k_1}^{(i_1)}, ..., m_{k_d}^{(i_d)})$ — der Mittelpunkt des Trägers von $g_{k_1,...,k_d}^{(i_1,...,i_d)}$ — zum dünnen Gitter gehören soll oder nicht.

Wir wollen hier zwei Adaptionsstrategien angeben, um so für jede unserer beiden Klassen $\tilde{G}_n^{(d)}$ und $\hat{G}_n^{(d)}$ regulärer dünner Gitter ein adaptives Analogon bereitzustellen. Soll sich unsere Adaptionsstrategie und somit das resultierende adaptive dünne Gitter wie $\tilde{G}_n^{(d)}$ an der Maximumsnorm orientieren, so gibt $\|c_{k_1,...,k_d}^{(i_1,...,i_d)} \cdot g_{k_1,...,k_d}^{(i_1,...,i_d)}\|_\infty$ den Beitrag eines Basiselements des Teilraumes $T_{i_1,...,i_d}$ zum Interpolanten wieder. Für diese Größe gilt nach (10)

$$\begin{aligned} \left\| c_{k_1,...,k_d}^{(i_1,...,i_d)} \cdot g_{k_1,...,k_d}^{(i_1,...,i_d)} \right\|_\infty &= |c_{k_1,...,k_d}^{(i_1,...,i_d)}| \cdot \left\| g_{k_1,...,k_d}^{(i_1,...,i_d)} \right\|_\infty \\ &= |c_{k_1,...,k_d}^{(i_1,...,i_d)}|. \end{aligned}$$

Bei gegebenem Adaptionsparameter $\epsilon$ beantwortet demnach der Test

$$|c_{k_1,...,k_d}^{(i_1,...,i_d)}| \ \geq \ \epsilon \ \ ? \tag{38}$$

die Frage, ob der betreffende Punkt $(m_{k_1}^{(i_1)}, ..., m_{k_d}^{(i_d)})$ zum dünnen Gitter gehören soll ($\geq \epsilon$) oder nicht ($< \epsilon$).

Wollen wir dagegen eine energieorientierte Adaptionsstrategie verfolgen, so ist die entscheidende Größe $\|c_{k_1,...,k_d}^{(i_1,...,i_d)} \cdot g_{k_1,...,k_d}^{(i_1,...,i_d)}\|_E$. Mit der Definition

$$a(i_1, ..., i_d) \ := \ \sqrt{2 \cdot \left(\frac{2}{3}\right)^{d-1} \cdot \sum_{j=1}^{d} \frac{h_{i_1} \cdot ... \cdot h_{i_d}}{h_{i_j}^2}}$$

gilt nun nach (27)

$$\left\| c_{k_1,...,k_d}^{(i_1,...,i_d)} \cdot g_{k_1,...,k_d}^{(i_1,...,i_d)} \right\|_E \ = \ |c_{k_1,...,k_d}^{(i_1,...,i_d)}| \cdot \left\| g_{k_1,...,k_d}^{(i_1,...,i_d)} \right\|_E \ = \ |c_{k_1,...,k_d}^{(i_1,...,i_d)}| \cdot a(i_1, ..., i_d).$$

Es liegt daher nahe, bei gegebenem $\epsilon$ das Ergebnis des Tests

$$|c_{k_1,...,k_d}^{(i_1,...,i_d)}| \cdot a(i_1, ..., i_d) \ \geq \ \epsilon \ \ ?$$

als Entscheidungskriterium für die Aufnahme eines Gitterpunktes in das adaptive dünne Gitter zu verwenden. Die hieraus resultierenden dünnen Gitter sind zwar den regelmäßigen energieorientierten Gittern $\hat{G}_n^{(d)}$ verwandt, sie bestehen jedoch aufgrund der Faktoren $a(i_1, ..., i_d)$, die in Extremfällen sehr groß werden können (betrachte beispielsweise $h_1 = ... = h_{d-1} = \frac{1}{2}$, $h_d = 2^{-n}$), aus mehr Punkten als die mittels (38) erzeugten Gitter. In Anlehnung an die Definition von $\hat{G}_n^{(d)}$ (siehe (29) sowie (30)) benutzen wir deshalb zur Beantwortung der Frage, ob weiter verfeinert werden soll oder nicht, die Bedingung

$$|c_{k_1,...,k_d}^{(i_1,...,i_d)}| \cdot \left( \max_{1 \leq j \leq d} \prod_{l \neq j} h_{i_l} \right) \ \geq \ \epsilon \ \ ? \tag{39}$$

als Entscheidungskriterium. Mit dieser Adaptionsstrategie gelangt man nun zu adaptiv verfeinerten dünnen Gittern, die nach denselben Prinzipien wie $\hat{G}_n^{(d)}$ aufgebaut sind und im Vergleich zu den mittels (38) erzeugten Gittern weniger Gitterpunkte enthalten.

Sowohl (38) als auch (39) verwenden den hierarchischen Überschuß $c_{k_1,...,k_d}^{(i_1,...,i_d)}$ als (Interpolations-) Fehlerindikator. Aufgrund von Beziehung (11) kann dieser leicht aus dem Funktionswert am Punkt $(m_{k_1}^{(i_1)}, ..., m_{k_d}^{(i_d)})$ selbst und aus Funktionswerten an Punkten auf dem Rand des Trägers von $g_{k_1,...,k_d}^{(i_1,...,i_d)}$ errechnet werden. Ferner haben wir in Abschnitt 2 darauf hingewiesen, daß der hierarchische Überschuß bis auf einen Faktor $(-1)^d \cdot \frac{h_{i_1}^2 \cdot ... \cdot h_{i_d}^2}{2^d}$ gerade der üblichen Differenzenquotienten-Approximation von $\frac{\partial^{2d} u^{(d)}}{\partial x_1^2 ... \partial x_d^2}$ entspricht. Den Wert dieser gemischten Ableitung wiederum verwenden wir aber gerade als Maß für die Glattheit einer Funktion. Somit stellen (38) und (39) vernünftige Kriterien zur Beantwortung der Frage dar, ob $(m_{k_1}^{(i_1)}, ..., m_{k_d}^{(i_d)})$ zum dünnen Gitter gehören soll oder nicht.

(12) lieferte für die hierarchischen Überschüsse $c_{k_1,...,k_d}^{(i_1,...,i_d)}$ eine obere Schranke, die nur indirekt über $\sum_{j=1}^{d} i_j$ von den Werten $i_j$ abhängt. Ferner treten die $k_j$ in der Schranke nicht auf. Daran kann man erkennen, daß bei glatten Funktionen $u^{(d)}$ Strategie (38) zu adaptiven Dünn-Gitter-Räumen führt, die den regulären Dünn-Gitter-Räumen $\tilde{S}_n^{(d)}$ aus Abschnitt 3 verwandt sind. Strategie (39) hingegen versieht den hierarchischen Überschuß mit einem zusätzlichen Faktor, der für $i_1 \approx ... \approx i_d$ erheblich kleiner ist als für sehr unterschiedliche Werte $i_j$ (bei gleicher Summe $i_1 + ... + i_d$). Deshalb resultieren aus dieser Adaptionsstrategie tatsächlich Dünn-Gitter-Räume, die den energieorientierten (regulären) Räumen $\hat{S}_n^{(d)}$ aus Abschnitt 4 verwandt sind.

An dieser Stelle sei erwähnt, daß sich unsere Implementierung dünner Gitter auf Binärbäume stützt. Durch die Verwendung einer dynamischen Datenstruktur läßt sich die Adaptivität auf sehr einfache und natürliche Weise realisieren: Der Baum wächst dort in die Tiefe, wo die Adaptionsstrategie dies erfordert. Insbesondere lassen sich adaptive und reguläre dünne Gitter mit denselben Programmen erzeugen. Dadurch spiegelt sich auch in der Implementierung wider, daß reguläre dünne Gitter einen Spezialfall adaptiv verfeinerter dünner Gitter darstellen. Wir beschreiben im folgenden die verwendete Datenstruktur.

**5.2. Eine Datenstruktur für dünne Gitter.** Wie bereits erwähnt benutzen wir zur Darstellung einer Funktion auf einem (regulären oder adaptiv verfeinerten) dünnen Gitter Binärbäume als Datenstruktur. Eine Funktion in einer Unbekannten wird dabei als Baum von Gitterpunkten dargestellt. Jeder Knoten im Baum entspricht dann einem Punkt im dünnen Gitter. Eine Funktion zweier Veränderlicher wird durch einen Baum von Gitterlinien beschrieben, in dem jeder Knoten einer Linie im dünnen Gitter und deshalb selbst einem Baum von Gitterpunkten (den Punkten auf der betreffenden Gitterlinie) entspricht. Abbildung 10 zeigt die Baumstruktur bei Funktionen in einer Unbekannten.

Die beiden Randpunkte $x = 0$ und $x = 1$ erfahren dabei eine gewisse Sonderbehandlung: Sie werden oberhalb der eigentlichen Wurzel $x = 0.5$ gespeichert und nehmen so im Baum die hierarchisch höchste Position ein. Dies ist deshalb sinnvoll, da unsere Interpolationstechnik die Werte einer Funktion auf dem Rand zur Berechnung von Innenwerten heranzieht.

Abbildung 11 zeigt nun die Baumstruktur bei Funktionen zweier Veränderlicher. Wiederum wird der Rand — die beiden Randzeilen $y = 0$ und $y = 1$ — oberhalb der eigentlichen Wurzel gespeichert. Jeder Knoten im Baum entspricht einer Gitterlinie

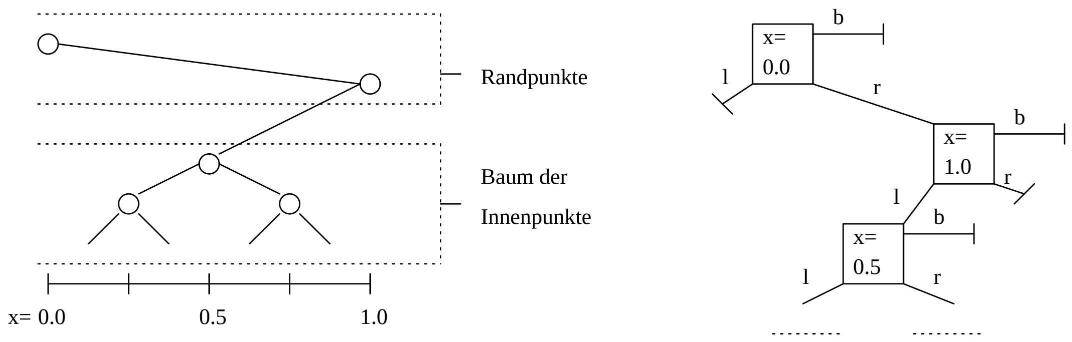



ABB. 10. *Baum zur Darstellung einer Funktion* $u(x)$ *auf dünnen Gittern.*

$y = \bar{y}$, die ihrerseits durch einen Baum zur Darstellung von $u(x, \bar{y})$ als Funktion in einer Unbekannten $x$ gemäß Abbildung 10 repräsentiert wird.

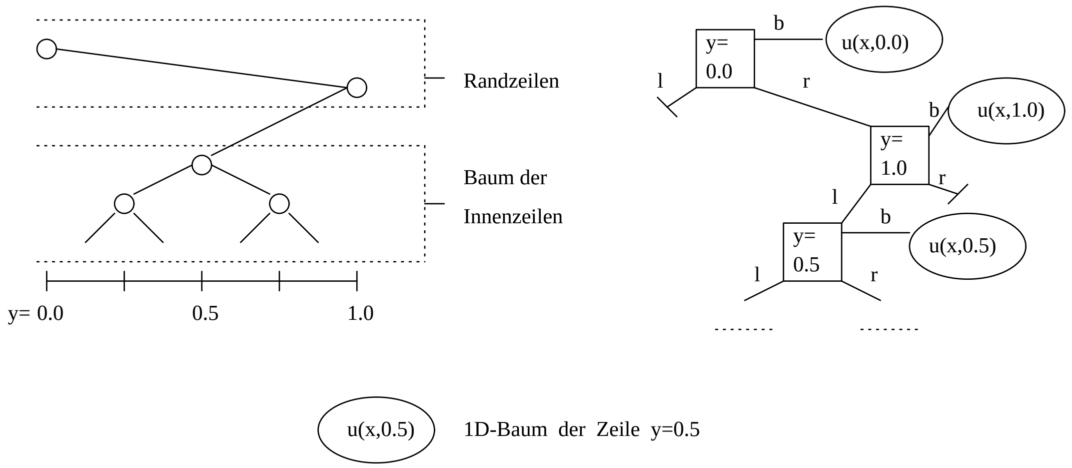



ABB. 11. *Baum zur Darstellung einer Funktion* $u(x, y)$ *auf dünnen Gittern.*

Wie man sofort sieht, läßt sich das geschilderte Darstellungsprinzip unmittelbar auf Funktionen von drei oder mehr Unbekannten verallgemeinern, wobei dieselbe Binärbaum-Datenstruktur rekursiv zur Anwendung gelangt.

Für den Spezialfall $d = 2$ beschreiben wir abschließend noch den Prozeß des Baumaufbaus für den Dünn-Gitter-Interpolanten einer Funktion $u^{(2)} \in X^{(2)}$, wobei wir jetzt auf dem Rand auch Werte ungleich Null zulassen. Man beginnt mit einem Baum, der nur Knoten für die vier Eckpunkte des Einheitsquadrats enthält. Ein solcher Baum repräsentiert gerade den bilinearen Interpolanten dieser vier Randwerte (ohne Beiträge aus dem Inneren). Entsprechend der hierarchischen Anordnung der Teilräume $T_{i_1,i_2}$ (vgl. (2)) wird dann zunächst der Knoten für den Mittelpunkt $x = y = 0.5$ erzeugt und in den Baum eingehängt ($T_{1,1}$). Zusammen mit den ebenfalls jetzt zu erzeugenden Knoten für die vier Randpunkte $(0, 0.5)$, $(1, 0.5)$, $(0.5, 0)$ sowie $(0.5, 1)$ ist damit das Gitter $\tilde{G}_1^{(2)}$ realisiert. Anschließend werden Knoten für diejenigen Gitterpunkte eingefügt, die den Teilräumen $T_{1,2}$ und $T_{2,1}$ entsprechen oder auf dem Rand liegen und zu $\tilde{G}_2^{(2)}$ gehören. Mit unserem Baum können wir in seinem jetzigen Zustand Funktionen darstellen, die sich aus einer Randkomponente sowie einer Komponente aus $\tilde{S}_2^{(2)}$ zusammensetzen. In einem weiteren Schritt erzeugt man nun Knoten für die zu $T_{1,3}$, $T_{2,2}$ und $T_{3,1}$ gehörigen Gitterpunkte sowie die noch fehlenden Randpunkte von $\tilde{G}_3^{(2)}$ und ermöglicht somit die Darstellung von Funktionen auf $\tilde{G}_3^{(2)}$. Diese Vorgehensweise führt offensichtlich zu regulären dünnen Gittern, wie sie in Abschnitt 3 dieser Arbeit beschrieben wurden. Macht man die Knotenerzeugung dagegen von einem der in Abschnitt 5.1 angegebenen Kriterien (38) oder (39) abhängig, so erhält man mit demselben Verfahren Bäume für Funktionen auf adaptiven dünnen Gittern.

Von zentraler Bedeutung beim Baumaufbau ist die Einhaltung der hierarchischen Ordnung der Teilräume. Zur Berechnung des hierarchischen Überschusses in einem Gitterpunkt werden nach (11) Funktionswerte an gewissen benachbarten Gitterpunkten benötigt, die hierarchisch höher angesiedelten Teilräumen zugeordnet sind. Existiert nun zum Zeitpunkt der Erzeugung eines Gitterpunktes ein solcher Nachbarpunkt nicht, so muß für ihn ebenfalls ein Knoten in den Baum eingehängt werden. Im Falle regulärer dünner Gitter verhindert der beschriebene Konstruktionsalgorithmus natürlich — abgesehen vom Rand — das Auftreten einer derartigen Situation, im Falle adaptiver Verfeinerung ist die Existenz hierarchisch höherer Nachbarn jedoch auch im Inneren des Einheitsquadrats nicht gesichert, so daß hier der Algorithmus gegebenenfalls solche fehlenden Knoten automatisch nachträglich erzeugen muß.

Nach diesen Bemerkungen zur Implementierung dünner Gitter verlassen wir nun den Bereich der Darstellung beziehungsweise Interpolation von Funktionen und wenden uns im zweiten Teil dieser Arbeit mit der numerischen Lösung partieller Differentialgleichungen einem neuen Anwendungsgebiet dünner Gitter zu. Vom Blickwinkel der Differentialgleichungen aus betrachtet zeigt sich hierbei die Schwierigkeit, daß nicht nur die Lösung, sondern auch das sie darstellende dünne Gitter im Laufe der Iteration erst ermittelt werden muß. Im Dünn-Gitter-Kontext stellt uns dieser Sachverhalt vor das Problem, ein auf eine gewisse Funktion – nämlich die Lösung der Differentialgleichung – zugeschnittenes dünnes Gitter ohne Kenntnis der betreffenden Funktion sukzessive zu konstruieren. Mit dieser Thematik befaßt sich der folgende Abschnitt.

**6. Zweidimensionale Poisson-Probleme auf dünnen Gittern.** Gegeben sei auf dem Einheitsquadrat folgendes Poisson-Modellproblem mit Dirichlet'schen Nebenbedingungen:

$$
\begin{aligned}
\Delta u(x_1,x_2) &= f(x_1,x_2) \text{ für } (x_1,x_2) \in \Omega^{(2)} = (0,1) \times (0,1), \\
u(x_1,x_2) &= g(x_1,x_2) \text{ für } (x_1,x_2) \in \delta\bar{\Omega}^{(2)}.
\end{aligned} \tag{40}
$$

Wir stellen ein iteratives Verfahren vor, welches eine Näherungslösung für (40) auf einem dünnen Gitter bereitstellt.

Dabei müssen wir unser Augenmerk zugleich auf zwei unterschiedliche Problemstellungen richten. Zum einen muß ein dünnes Gitter konstruiert werden, welches sich für die Darstellung der zu produzierenden Näherungslösung von (40) eignet. Hierzu sind die Reihenfolge der Knotenerzeugung sowie eine reguläre oder adaptive Verfeinerungsstrategie festzulegen. Zum anderen benötigen wir ein Verfahren, welches auf einem vorliegenden dünnen Gitter eine Näherungslösung für unser Modellproblem (40) liefert. Dabei muß zunächst ein geeigneter Ansatz zur Diskretisierung des Differentialgleichungsoperators gewählt werden. Ebenso muß die rechte Seite $f$ der Differentialgleichung passend diskretisiert werden. Schließlich ist das aus der Diskretisierung des Laplace-Operators und der rechten Seite resultierende lineare Gleichungssystem iterativ zu lösen.

Es ist klar, daß die beiden genannten Teilaufgaben (Gittergenerierung und Lösung des Modellproblems) im algorithmischen Ablauf nicht getrennt werden können; denn ein sukzessive aufgebautes dünnes Gitter wird nur dann sinnvoll zur Darstellung einer Näherungslösung für (40) verwendet werden können, wenn schon frühzeitig Information über die Gestalt dieser Funktion herangezogen wird. Somit muß bereits iteriert werden, bevor der Aufbau des Gitters abgeschlossen ist. Umgekehrt wiederum bedarf es einer ausreichenden Anzahl von Gitterpunkten, wenn die Iteration eine brauchbare Näherungslösung produzieren soll.

Zunächst befassen wir uns kurz mit dem ersten Aspekt unserer Methode, der Gittergenerierung. Anschließend wenden wir uns dann dem eigentlichen Lösungsverfahren zu. Als Diskretisierungsansatz wählen wir dabei die Methode der Finiten Elemente, wobei als Testfunktionen die in (6) eingeführten bilinearen hierarchischen Basisfunktionen verwendet werden. Die Steifigkeitsmatrix wird für reguläre wie adaptiv erzeugte dünne Gitter explizit angegeben. Die Diskretisierung der rechten Seite $f$ der Differentialgleichung erfordert beim Finite-Elemente-Ansatz die Berechnung gewichteter Integrale von $f$. Deshalb gehen wir auch auf die Darstellung der Funktion $f$ ein und geben eine Technik zur numerischen Quadratur auf dünnen Gittern an. Bei der iterativen Lösung des entstehenden linearen Gleichungssystems gelangt dann das Gauß-Seidel-Verfahren zur Anwendung. Schließlich widmen wir uns noch der Frage nach der Genauigkeit der erzielten Näherungslösung sowie der Konvergenzgeschwindigkeit.

**6.1. Die sukzessive Erzeugung des dünnen Gitters.** Als erste Näherung für die Lösung $u$ von (40) verwenden wir den bilinearen Interpolanten aus den vier Eckwerten der Dirichlet'schen Randfunktion $g$. Folglich enthält unser Startbaum lediglich

Knoten für die vier Eckpunkte von $\bar{\Omega}^{(2)}$. Ein Iterationsschritt besteht nun aus einem Durchlauf durch den Teil des Baumes, der das Innere des Einheitsquadrats repräsentiert. Dieser Baumdurchlauf erfolgt in hierarchischer Ordnung, wie dies in Abschnitt 5.2 beim Baumaufbau beschrieben worden ist — beginnend beim Punkt $x_1 = x_2 = 0.5$. Dabei wird in jedem Knoten der Funktionswert der Näherung gemäß einer noch zu besprechenden Iterationsvorschrift aktualisiert.

Erreicht man ein Blatt des Baumes, so muß eine Entscheidung gefällt werden, ob weiter verfeinert werden soll oder nicht. Dabei unterscheiden wir zwei Klassen von Blättern beziehungsweise Knoten (siehe Abbildung 10 und Abbildung 11): Knoten im Baum von Gitterlinien, die eine ganze Gitterlinie $x_2 = \bar{x}_2$ repräsentieren, und Knoten innerhalb eines Teilbaumes für eine Gitterlinie, die einen einzelnen Gitterpunkt $(\bar{x}_1, \bar{x}_2)$ repräsentieren. In jedem Fall erhält das Blatt bei positiver Entscheidung zwei Söhne. Diese stehen dann im Baum für die Gitterlinien $x_2 = \bar{x}_2 - h_{i_2}/2$ sowie $x_2 = \bar{x}_2 + h_{i_2}/2$ beziehungsweise für die Gitterpunkte $(\bar{x}_1 - h_{i_1}/2, \bar{x}_2)$ sowie $(\bar{x}_1 + h_{i_1}/2, \bar{x}_2)$. Als Entscheidungskriterium verwendet man dabei im Falle regulärer dünner Gitter die Zugehörigkeit des entsprechenden Teilraumes $T_{i_1+1,i_2}$ beziehungsweise $T_{i_1,i_2+1}$ zu $\tilde{S}_n^{(2)}$ oder $\hat{S}_n^{(2)}$. Falls wir hingegen adaptiv verfeinern wollen, so kommt eines der Kriterien (38) oder (39) aus Abschnitt 5.1 zum Tragen.

Werden während des Iterationsprozesses für die Ermittlung der hierarchischen Überschüsse zusätzliche Gitterpunkte auf dem Rand benötigt, so erzeugt der Algorithmus die entsprechenden Knoten und hängt sie in den Baum ein.

Nach diesen einleitenden Ausführungen zur Gittererzeugung wenden wir uns nun dem eigentlichen Thema dieses Abschnitts zu, dem Differentialgleichungslöser auf dünnen Gittern.

**6.2. Das Iterationsverfahren zur Lösung des Randwertproblems.** Zur Lösung von (40) auf beliebigen regulären oder adaptiven dünnen Gittern wählen wir einen Ansatz gemäß der Methode der Finiten Elemente. Das sich hieraus ergebende lineare Gleichungssystem lösen wir iterativ mit Hilfe des Gauß-Seidel-Verfahrens.

**6.2.1. Die Methode der Finiten Elemente auf dünnen Gittern.** Gegeben sei ein reguläres oder adaptiv erzeugtes dünnes Gitter mit $N_I$ Innen- und $N_R$ Randpunkten. Dazu betrachten wir die zugehörigen hierarchischen Basen $\{\varphi_j, 1 \leq j \leq N_I\}$ und $\{\zeta_j, 1 \leq j \leq N_R\}$, bestehend aus den bilinearen hierarchischen Basisfunktionen aus Abschnitt 2 im Innern $(\varphi_j)$ und auf dem Rand $(\zeta_j)$. Im Falle eines regulären dünnen Gitters $\tilde{G}_n^{(2)}$ gemäß (15) gilt nach (19) und (22)

$$\begin{aligned} N_I &= \tilde{N}_n^{(2)} &&= 2^n \cdot (n-1) + 1, \\ N_R &= \tilde{M}_n^{(2)} - \tilde{N}_n^{(2)} &&= 2^n \cdot 4. \end{aligned}$$

Als Basisfunktionen $\varphi_j$ erhalten wir bei Betrachtung von $\tilde{G}_n^{(2)}$ mit (6)

$$\left\{\varphi_j, 1 \leq j \leq N_I\right\} = \left\{g_{k_1,k_2}^{(i_1,i_2)}, (i_1,i_2) \in \tilde{B}_n^{(2)}, 1 \leq k_1 \leq 2^{i_1-1}, 1 \leq k_2 \leq 2^{i_2-1}\right\},$$

wobei der Index $j$ die Hierarchie der Teilräume widerspiegeln soll, d.h., für $\varphi_i \in T_{i_1,i_2}$ und $\varphi_j \in T_{j_1,j_2}$ gelte (vgl. (2))

$$T_{i_1,i_2} \; < \; T_{j_1,j_2} \quad \Rightarrow \quad i < j. \tag{41}$$

Die $\zeta_j$ werden analog als hierarchische Basisfunktionen auf dem Rand definiert.

Die tatsächliche Lösung von (40) soll nun durch eine geeignete Funktion $u$ auf dem vorliegenden dünnen Gitter angenähert werden. Weil $\{\varphi_j, 1 \leq j \leq N_I\}$ und $\{\zeta_j, 1 \leq j \leq N_R\}$ zusammen eine Basis des zugehörigen Dünn-Gitter-Raumes bilden, gibt es reelle Werte $\lambda_j, 1 \leq j \leq N_I$, sowie $\mu_j, 1 \leq j \leq N_R$, so daß gilt

$$u \;=\; \sum_{j=1}^{N_R} \mu_j \cdot \zeta_j \;\;+\;\; \sum_{j=1}^{N_I} \lambda_j \cdot \varphi_j. \tag{42}$$

Dabei sind die $\mu_j$ wie die $\lambda_j$ als Koeffizienten hierarchischer Basisfunktionen hierarchische Werte (d.h. Funktionswerte in den vier Eckpunkten, eindimensionale hierarchische Überschüsse in den restlichen Randpunkten und zweidimensionale hierarchische Überschüsse im Innern). Da die Dirichlet-Randbedingung aus (40) die Werte $\mu_j$ eindeutig festlegt, reduziert sich unser Problem auf die Bestimmung der Koeffizienten $\lambda_j$. Mit der abkürzenden Schreibweise

$$\varphi_0 \;:=\; \sum_{j=1}^{N_R} \mu_j \cdot \zeta_j \tag{43}$$

erhalten wir durch Einsetzen von (42) in (40)

$$\Delta u \;=\; \Delta\Big(\varphi_0 \;+\; \sum_{j=1}^{N_I} \lambda_j \cdot \varphi_j\Big) \;=\; \Delta\varphi_0 \;+\; \sum_{j=1}^{N_I} \lambda_j \cdot \Delta\varphi_j \;=\; f \tag{44}$$

und somit als Residuum

$$R \;:=\; \Delta u - f \;=\; \Delta\varphi_0 \;+\; \sum_{j=1}^{N_I} \lambda_j \cdot \Delta\varphi_j - f. \tag{45}$$

Nach der Idee von Galerkin wird nun das Verschwinden der Integrale

$$\int_0^1 \int_0^1 R \cdot \varphi_i \; dx_2 \; dx_1, \quad 1 \leq i \leq N_I, \tag{46}$$

gefordert. Dies entspricht dem Übergang zur sogenannten schwachen Form der Differentialgleichung, bei der alle Testfunktionen $v$ aus dem Dünn-Gitter-Raum die Gleichung

$$\int_0^1 \int_0^1 \Delta u \cdot v \; dx_2 \; dx_1 \;=\; \int_0^1 \int_0^1 f \cdot v \; dx_2 \; dx_1 \tag{47}$$

erfüllen müssen.

Ersetzt man jetzt in (46) das Residuum $R(x_1, x_2)$ gemäß (45), so erhält man für die $N_I$ Koeffizienten $\lambda_j$ $N_I$ Bestimmungsgleichungen ($1 \leq i \leq N_I$):

$$\int_0^1 \int_0^1 \Big( \Delta\varphi_0 \;+\; \sum_{j=1}^{N_I} \lambda_j \cdot \Delta\varphi_j \Big) \cdot \varphi_i \; dx_2 \; dx_1 \;=\; \int_0^1 \int_0^1 f \cdot \varphi_i \; dx_2 \; dx_1. \tag{48}$$

Nach Vertauschung von Integration und Summation läßt sich dieses System von Bestimmungsgleichungen für die $\lambda_j$ mit Hilfe der ersten Green'schen Formel in das gleichwertige System ($1 \leq i \leq N_I$)

$$\sum_{j=1}^{N_I} \lambda_j \cdot \int_0^1 \int_0^1 \langle \nabla\varphi_j, \nabla\varphi_i \rangle \; dx_2 \; dx_1 \quad = \quad -\int_0^1 \int_0^1 \Big( \langle \nabla\varphi_0, \nabla\varphi_i \rangle + f \cdot \varphi_i \Big) \; dx_2 \; dx_1 \tag{49}$$

transformieren. (49) läßt sich nun als lineares Gleichungssystem $Av = b$ schreiben, wobei

$$\begin{aligned}
A &= (a_{ij})_{1 \leq i,j \leq N_I} \quad \in \mathbb{R}^{N_I \times N_I}, \\
a_{ij} &= \int_0^1 \int_0^1 \langle \nabla\varphi_i, \nabla\varphi_j \rangle \; dx_2 \; dx_1, \\
v &= (v_1, ..., v_{N_I})^T \quad \in \mathbb{R}^{N_I}, \\
b &= (b_1, ..., b_{N_I})^T \quad \in \mathbb{R}^{N_I}, \\
b_i &= -\int_0^1 \int_0^1 \langle \nabla\varphi_0, \nabla\varphi_i \rangle \; dx_2 \; dx_1 \quad - \int_0^1 \int_0^1 f \cdot \varphi_i \; dx_2 \; dx_1.
\end{aligned} \tag{50}$$

Die Komponenten $b_i$ der rechten Seite $b$ unseres Gleichungssystems stellen dabei Summen zweier Integrale dar. Die Integranden $f \cdot \varphi_i$ geben den Einfluß der rechten Seite $f$ der Poisson-Gleichung wieder, in den Integranden $\langle \nabla\varphi_0, \nabla\varphi_i \rangle$ schlägt sich die Dirichlet-Randfunktion $g$ aus (40) nieder (vgl. die Definition von $\varphi_0$ in (43)). Auf die Berechnung von $b$ werden wir an späterer Stelle eingehen. Zunächst befassen wir uns mit der Gestalt der Steifigkeitsmatrix $A$. Die Matrixeinträge $a_{ij}$ von $A$ und damit die Integrale

$$\int_0^1 \int_0^1 \langle \nabla\varphi_i, \nabla\varphi_j \rangle \; dx_2 \; dx_1, \quad 1 \leq i, j \leq N_I,$$

wurden in [29] explizit berechnet. Wir fassen die Ergebnisse im folgenden Satz 6.1 zusammen, dessen Aussagen für Steifigkeitsmatrizen bezüglich regulärer wie adaptiv erzeugter dünner Gitter gelten.

SATZ 6.1. *Seien $\varphi_i \in T_{i_1,i_2}$ und $\varphi_j \in T_{j_1,j_2}$, $1 \leq i,j \leq N_I$, zwei verschiedene Basisfunktionen unserer hierarchischen Dünn-Gitter-Basis.*

*(i) Für die Diagonalelemente der Steifigkeitsmatrix $A$ gilt*

$$a_{ii} \;=\; \frac{4}{3} \cdot \Big( \frac{h_{i_1}}{h_{i_2}} + \frac{h_{i_2}}{h_{i_1}} \Big). \tag{51}$$

*(ii) Gibt es keine achsenparallele Gerade, auf der die zu i und j gehörigen Gitterpunkte (die Mittelpunkte der Träger von $\varphi_i$ und $\varphi_j$) beide liegen, oder sind die Träger von $\varphi_i$ und $\varphi_j$ disjunkt, so gilt*

$$a_{ij} = 0. \tag{52}$$

*(iii) Andernfalls liegen die zu i und j gehörenden Gitterpunkte auf einer achsenparallelen Geraden, einer der beiden Träger von $\varphi_i$ oder $\varphi_j$ ist Teilmenge des anderen, und es gilt $a_{ij} \neq 0$. Sei o.B.d.A. die verbindende Gerade parallel zur $x_1$-Achse, und sei ebenfalls o.B.d.A. der Träger von $\varphi_i$ Teilmenge des Trägers von $\varphi_j$. Dann gilt*

$$a_{ij} \;=\; 2 \cdot \frac{h_{i_1}}{h_{j_1} \cdot h_{j_2}} \cdot (h_{i_1} + w), \tag{53}$$

*wobei w die Einrücktiefe des Trägers von $\varphi_i$ im Träger von $\varphi_j$ bezeichnet (siehe Abbildung 12).*

*Beweis.* Siehe [29]. □

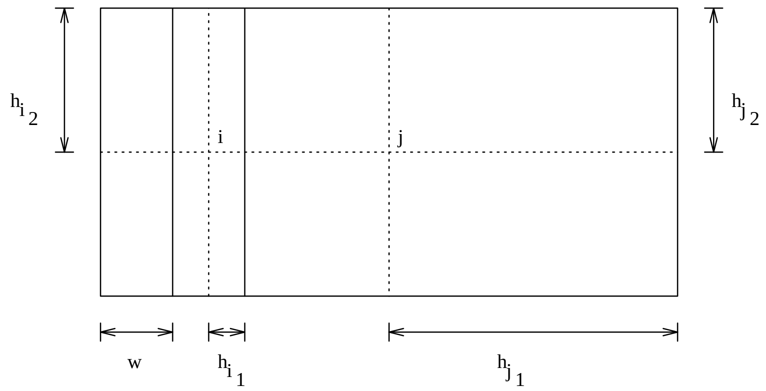



ABB. 12. *Träger von $\varphi_i$ und $\varphi_j$ im Fall (iii).*

Wir fügen die Bemerkung an, daß für Integrale des Typs

$$\int_0^1 \int_0^1 \langle \nabla \zeta_j, \nabla \varphi_i \rangle \; dx_2 \; dx_1,$$

die nach Einsetzen von $\varphi_0$ gemäß (43) in (50) bei der Berechnung der rechten Seite $b$ des Gleichungssystems auftreten, analoge Beziehungen gelten.

Erwartungsgemäß ist die Steifigkeitsmatrix bezüglich eines dünnen Gitters aufgrund der Verwendung hierarchischer Basen nicht so dünn besiedelt wie die klassischen, auf Stützpunktbasen beruhenden Steifigkeitsmatrizen. Die Anzahl von Null verschiedener Matrixeinträge in einer Zeile $k$ variiert dabei stark in Abhängigkeit von $k$. Im Falle des regulären dünnen Gitters $\tilde{G}_n^{(2)}$ ist die Anzahl der Nicht-Nullen in Zeile 1 ($\varphi_1 \in T_{1,1}$, vgl. (41)) von der Ordnung $O(h_n^{-1}) = O(2^n)$, in Zeile $k$ mit $\varphi_k \in T_{1,n}$ beispielsweise dagegen nur von der Ordnung $O(\mathrm{ld}(h_n^{-1})) = O(n)$. Abbildung 13 zeigt am Beispiel des dünnen Gitters $\tilde{G}_4^{(2)}$ die Besetzungsstruktur der resultierenden Steifigkeitsmatrix $A$. Im Mittel ergeben sich für $\tilde{G}_n^{(2)}$ pro Zeile $O(n)$ von Null verschiedene Einträge, wohingegen die Steifigkeitsmatrix beim vollen Gitter $G_n^{(2)}$ und bei Verwendung der Standard-Stützpunktbasis nur $O(1)$ Nicht-Nullen je Zeile hat. Allerdings hat letztere Matrix, bedingt durch die größere Zahl an Unbekannten im vollen Gitter, ja auch wesentlich mehr Zeilen. Eventuelle Auswirkungen der Matrixgestalt auf den Rechenaufwand oder das Konvergenzverhalten unseres Verfahrens werden wir anschließend bei der Besprechung des iterativen Prozesses zur Lösung des linearen Gleichungssystems $Av = b$ sowie bei den Betrachtungen zur Konvergenz behandeln.

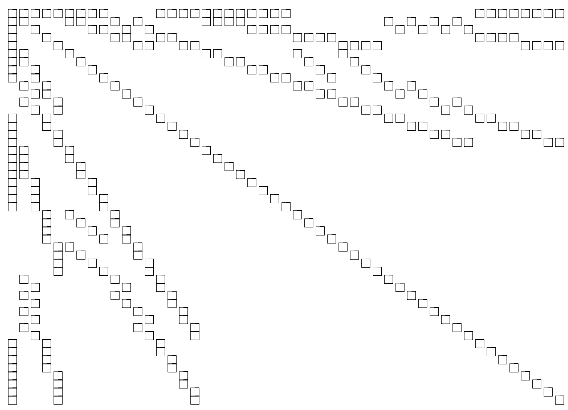


ABB. 13. *Von Null verschiedene Einträge in der Steifigkeitsmatrix zu* $\tilde{G}_4^{(2)}$.

**6.2.2. Die iterative Lösung des Gleichungssystems.** Das aus dem Finite-Elemente-Ansatz resultierende lineare Gleichungssystem $Av = b$ wird nun iterativ mit Hilfe des Gauß-Seidel-Verfahrens gelöst. Dabei werden in einem Iterationsschritt neue Näherungswerte $v_i$ berechnet gemäß

$$(54) \qquad i = 1, ..., N_I : \qquad v_i \; := \; \Big(b_i - \sum_{\substack{j\,=\,1 \\ j\,\neq\,i}}^{N_I} a_{ij} \cdot v_j\Big) / a_{ii}.$$

Die Reihenfolge, in der die einzelnen Werte $v_i$ in einem Iterationsschritt korrigiert werden, wird durch die Hierarchie der Teilräume $T_{i_1,i_2}$ (2) bestimmt. So wird zunächst der den Teilraum $T_{1,1}$ repräsentierende Mittelpunkt $(0.5, 0.5)$ behandelt. Anschließend werden die Gitterpunkte zu $T_{i_1,i_2}$ mit $i_1 + i_2 = 3$ mit neuen Näherungswerten versehen, gefolgt von den Gitterpunkten zu Teilräumen $T_{i_1,i_2}$ mit $i_1 + i_2 = 4$, usw.. Da sich die Baumstruktur nach Abschnitt 6.1 ebenfalls an der Hierarchie der Teilräume orientiert, kann ein vollständiger Gauß-Seidel-Schritt auf einem dünnen Gitter während eines Durchlaufs durch den gesamten Baum erledigt werden.

Im vorigen Abschnitt wurde für die hierarchische Dünn-Gitter-Steifigkeitsmatrix $A$ bei Verwendung eines regulären dünnen Gitters $\tilde{G}_n^{(2)}$ gezeigt, daß sie im Mittel $O(n)$ von Null verschiedene Einträge pro Zeile enthält. Um zu verhindern, daß daraus pro Iterationsschritt ein Rechenaufwand von $O(N_I \cdot n) = O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^2)$ entsteht, berechnen wir in der vorliegenden Implementierung die Matrixeinträge $a_{ij}$ nicht explizit.

Vielmehr wird in jedem Gitterpunkt die zugehörige Gauß-Seidel-Gleichung (54) unter Verwendung während des Baumdurchlaufs akkumulierter Größen aufgestellt und gelöst.

Sei nun $\varphi_i$ eine Basisfunktion des Teilraumes $T_{i_1,i_2}$. Mit $i$ bezeichnen wir kurz den zugehörigen Gitterpunkt, den Mittelpunkt des Trägers von $\varphi_i$. Nach Satz 6.1 ist $a_{ij} \neq 0$ für $i \neq j$ nur möglich, wenn $i$ und $j$ dieselbe $x_1$- oder $x_2$-Koordinate haben und wenn einer der beiden Träger von $\varphi_i$ und $\varphi_j$ den anderen enthält. Nur diejenigen Gitterpunkte $j$ mit $a_{ij} \neq 0$ können aber nach (54) den neuen Wert $v_i$ beeinflussen. Je nachdem, in welcher Koordinate $i$ und $j$ übereinstimmen und welcher Träger den anderen enthält, teilen wir die Gitterpunkte $j$ mit $a_{ij} \neq 0$ in vier Gruppen ein und erhalten so folgende Gestalt der Iterationsgleichung (54):

$$(55) \qquad v_i := \frac{-3 \cdot \left(h_{i_1}^2 \cdot R_i^{(1)} + h_{i_2}^2 \cdot R_i^{(2)} \; + \; h_{i_1} \cdot r_i^{(1)} + h_{i_2} \cdot r_i^{(2)} - h_{i_1} \cdot h_{i_2} \cdot b_i\right)}{4 \cdot (h_{i_1}^2 + h_{i_2}^2)}.$$

(55) spiegelt dabei sowohl die Iterationsvorschrift als auch implizit die Diskretisierung, d.h. die Matrix $A$, wider. $R_i^{(1)}$ enthält nun die Einflüsse von den Gitterpunkten $j$, die dieselbe $x_2$-Koordinate wie $i$ haben und bei denen der Träger von $\varphi_j$ den Träger von $\varphi_i$ enthält. Auf (54) bezogen bedeutet dies

$$R_i^{(1)} \;=\; \frac{h_{i_2}}{h_{i_1}} \cdot \sum_{\substack{x_2(j) = x_2(i) \\ \text{Träger}(\varphi_j) \supset \text{Träger}(\varphi_i)}} a_{ij} \cdot v_j \;.$$

$R_i^{(2)}$ enthält die Einflüsse von den Gitterpunkten $j$, die dieselbe $x_1$-Koordinate wie $i$ haben und bei denen der Träger von $\varphi_j$ den Träger von $\varphi_i$ enthält.

Dementsprechend werden in den Variablen $r_i^{(1)}$ sowie $r_i^{(2)}$ die Einflüsse akkumuliert, die von Gitterpunkten $j$ im Inneren des Trägers von $\varphi_i$ mit gleicher $x_2$- beziehungsweise $x_1$-Komponente wie $i$ herrühren:

$$r_i^{(1)} \;=\; h_{i_2} \cdot \sum_{\substack{x_2(j) = x_2(i) \\ \text{Träger}(\varphi_j) \subset \text{Träger}\ (\varphi_i)}} a_{ij} \cdot v_j \;.$$

Abbildung 14 zeigt für einen Gitterpunkt $i$, welche Punkte $j$ wegen $a_{ij} \neq 0$ den neuen Wert $v_i$ beeinflussen, und veranschaulicht die Bedeutung der Größen $R_i^{(1)}$, $R_i^{(2)}$, $r_i^{(1)}$ sowie $r_i^{(2)}$.

Entscheidend ist nun, daß wir die in (55) auftretenden Summen $R_i^{(1)}$, $R_i^{(2)}$, $r_i^{(1)}$ und $r_i^{(2)}$ während des Baumdurchlaufs mit konstantem Aufwand je Gitterpunkt bilden können. Hierbei stützt sich die Berechnung obiger vier Größen für einen Gitterpunkt $i$ auf die entsprechenden Größen hierarchisch tieferer Gitterpunkte, so daß wir auch bei diesem Akkumulationsprozeß die Baumstruktur unmittelbar ausnutzen. Damit folgt für unsere Gauß-Seidel-Iteration, daß trotz der durchschnittlich $O(n)$ von Null verschiedenen Matrixeinträge pro Zeile ein Iterationsschritt (einmalige Korrektur jeder Variablen $v_i$) mit $O(N_I)$ Additionen und Multiplikationen durchgeführt werden kann. Die im Vergleich zu vollen Gittern geringere Anzahl an Unbekannten kommt also voll zum Tragen.

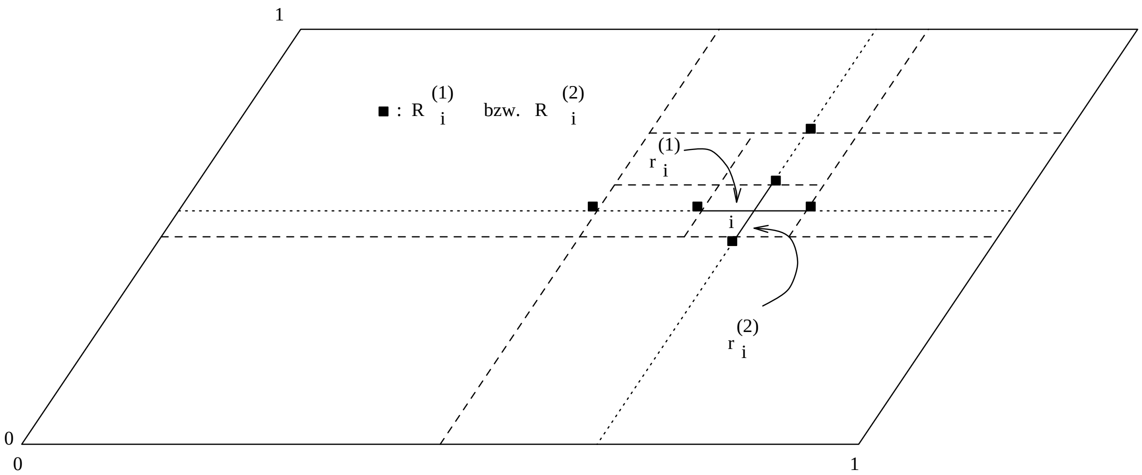



ABB. 14. *Punkte $j$ mit $a_{ij} \neq 0$ im Inneren, auf dem Rand und außerhalb des Trägers von $\varphi_i$.*

**6.2.3. Die Behandlung der rechten Seite der Poisson-Gleichung.** Nach (50) gilt für die Komponenten $b_i,\ 1 \leq i \leq N_I$, der rechten Seite des zu lösenden linearen Gleichungssystems $Av = b$

$$b_i \ = \ -\int_0^1\int_0^1 \langle \nabla\varphi_0, \nabla\varphi_i\rangle \ dx_2 \ dx_1 \ - \ \int_0^1\int_0^1 f\cdot\varphi_i \ dx_2 \ dx_1. \tag{56}$$

Der Integrand $\langle \nabla\varphi_0, \nabla\varphi_i\rangle$ hat dabei nach Einsetzen von (43) strukturell dieselbe Gestalt wie die bei der Berechnung der Matrixkoeffizienten $a_{ij}$ auftretenden Integranden. Deshalb gelten auch hierfür die Aussagen aus Satz 6.1 (vgl. die Bemerkung nach Satz 6.1).

Demgegenüber ist der Integrand $f \cdot \varphi_i$ nicht mehr Skalarprodukt der Gradienten zweier stückweise bilinearer Funktionen, sondern Produkt einer stückweise bilinearen Funktion $\varphi_i$ mit einer beliebigen Funktion $f$. Damit haben wir für jedes $i,\ 1 \leq i \leq N_I$, das Integral

$$\mathcal{I}_i \ := \ \int_0^1\int_0^1 f\cdot\varphi_i \ dx_2 \ dx_1$$

zu berechnen. Hierzu muß im allgemeinen ein Verfahren zur numerischen Quadratur herangezogen werden. Wir stellen eine Methode vor, die sich auf dünne Gitter stützt und eine effiziente näherungsweise Berechnung sämtlicher obiger Integrale $\mathcal{I}_i$ gestattet.

Als Näherung für $\mathcal{I}_i$ betrachten wir

$$\tilde{\mathcal{I}}_i \ := \ \int_0^1\int_0^1 \tilde{f}_I\cdot\varphi_i \ dx_2 \ dx_1, \tag{57}$$

wobei $\tilde{f}_I$ den Interpolanten von $f$ in einen geeignet gewählten Dünn-Gitter-Raum bezeichne. Nach Erzeugung eines Binärbaumes für $f$ beziehungsweise für $\tilde{f}_I$ liegt $\tilde{f}_I$ nun gemäß (42) in hierarchischer Darstellung vor ($\alpha_j, \beta_j \in \mathbb{R}$):

$$\tilde{f}_I \ = \ \sum_{j=1}^{N_R} \alpha_j\cdot\zeta_j \ + \ \sum_{j=1}^{N_I}\beta_j\cdot\varphi_j. \tag{58}$$

Aus (57) und (58) folgt somit

$$(59)\quad \begin{aligned} \tilde{\mathcal{I}}_i &= \int_0^1\int_0^1 \Big( \sum_{j=1}^{N_R} \alpha_j \cdot \zeta_j \;+\; \sum_{j=1}^{N_I} \beta_j \cdot \varphi_j \Big) \cdot \varphi_i \; dx_2 \; dx_1 \\ &= \sum_{j=1}^{N_R} \; \alpha_j \cdot \int_0^1\int_0^1 \zeta_j \cdot \varphi_i \; dx_2 \; dx_1 \;+\; \sum_{j=1}^{N_I} \; \beta_j \cdot \int_0^1\int_0^1 \varphi_j \cdot \varphi_i \; dx_2 \; dx_1 . \end{aligned}$$

Im Gegensatz zu den bei der Berechnung der Koeffizienten der Steifigkeitsmatrix auftretenden Integralen führen hier nur solche Paare von Basisfunktionen zu einem Integralwert Null, deren Träger disjunkt sind. Dennoch lassen sich alle $\tilde{\mathcal{I}}_i$ mit $O(N_I + N_R)$ Additionen und Multiplikationen berechnen. Denn wie wir zuvor die Bestimmung der Matrixeinträge $a_{ij}$ vermieden haben, so müssen auch hier die Integrale

$$\int_0^1\int_0^1 \zeta_j \cdot \varphi_i \; dx_2 \; dx_1 \quad \text{und} \quad \int_0^1\int_0^1 \varphi_j \cdot \varphi_i \; dx_2 \; dx_1$$

nicht explizit angegeben werden. Vielmehr gewährleistet ein geeigneter Summationsprozeß die Bereitstellung aller Werte $\tilde{\mathcal{I}}_i$ während eines einzigen Baumdurchlaufs.

Vom algorithmischen Standpunkt aus betrachtet kann die Berechnung der $\tilde{\mathcal{I}}_i$ an zwei verschiedenen Stellen geschehen. Es liegt zunächst nahe, die Integrale während der Iteration zu berechnen. Dabei wird in jedem Iterationsschritt für alle zu diesem Zeitpunkt existierenden Knoten $i$ der zugehörige Wert $\tilde{\mathcal{I}}_i$ ermittelt, wobei als dünnes Gitter für die Integralberechnung das Gitter zur Darstellung der aktuellen Näherungslösung verwendet wird. Diese Vorgehensweise hat den Vorteil, daß für $f$ kein Baum explizit konstruiert werden muß. Vielmehr wird alles über den Baum für $u$ gesteuert. Allerdings müssen dann sämtliche $\tilde{\mathcal{I}}_i$ in jedem Iterationsschritt neu berechnet werden. Zudem ist die Genauigkeit der Integrale zu Beginn der Iteration aufgrund der zunächst geringen Baumtiefe schlecht.

Um Mehrfachberechnungen zu vermeiden und um die Integralwerte von Anfang an mit der gewünschten Genauigkeit zur Verfügung stellen zu können, bestimmen wir dagegen die $\tilde{\mathcal{I}}_i$ vor Beginn der eigentlichen Iteration. Dazu bauen wir in einer Art Vorlaufrechnung explizit einen Baum für $f$ auf. Die anschließend startende Iteration zur Bestimmung der Näherungslösung $u$ unseres Poisson-Problems benutzt dann einen geeigneten Ausschnitt dieses Baumes zur Darstellung von $u$. Diese Vorgehensweise gestattet eine stärkere Berücksichtigung der Gestalt von $f$, da die Adaptionsprozesse bezüglich $f$ und $u$ entkoppelt sind. Insbesondere im Falle einer nicht glatten Funktion $f$ sollte dies von Vorteil sein. Wir wählen deshalb bei der Konstruktion des Baumes für $f$ und bei der nachfolgenden Iteration verschiedene Gitterparameter $n_f$ und $n_u$ (regelmäßige dünne Gitter) beziehungsweise $\epsilon_f$ und $\epsilon_u$ (adaptive dünne Gitter). Zur Vermeidung von Situationen, in denen die Gestalt von $u$ das Einfügen neuer Knoten in den $f$-Baum erzwingt, sollte der $f$-Baum dabei mindestens eine Ebene tiefer gehen, als wir dies später für den $u$-Baum beabsichtigen. Denn während der Iteration neu eingefügte Knoten würden einen Neustart der Vorlaufrechnung erfordern. Als groben Anhaltspunkt für die Wahl der Gitterparameter halten wir also fest:

$$n_f > n_u \quad \text{beziehungsweise} \quad \epsilon_f < \epsilon_u. \tag{60}$$

Zur Gewinnung präziserer Kriterien für die Wahl der Gitterparameter bei $f$ und $u$ — insbesondere im adaptiven Fall — werden die Auswirkungen des Zusammenspiels der Gitterparameter detailliert analysiert werden müssen. Hier wollen wir uns jedoch mit dem Kriterium (60) begnügen.

Abschließend sei noch darauf hingewiesen, daß wir — gewissermaßen als Nebenprodukt — eine adaptive Technik zur numerischen Berechnung spezieller Integrale mittels dünner Gitter gewonnen haben. Es liegt deshalb nahe, ausgehend von dünnen Gittern und hierarchischen Produktbasen ein allgemeines Verfahren zur numerischen Quadratur im $d$-dimensionalen Raum zu entwickeln. Neben der Lösung partieller Differentialgleichungen stehen derartige Bemühungen zur Zeit im Brennpunkt des Dünn-Gitter-Interesses.

**6.2.4. Betrachtungen zur Konvergenz.** Für die Analyse des in diesem Abschnitt vorgestellten Verfahrens zur Lösung zweidimensionaler Poisson-Probleme auf dünnen Gittern wollen wir noch einige Bemerkungen zu seinem Konvergenzverhalten machen. Dabei gehen wir auf Konvergenz sowohl im Sinne der Genauigkeit der erzielten Lösung als auch im Sinne der Anzahl der erforderlichen Iterationen ein.

Das Hauptproblem bei der Beurteilung der Qualität der berechneten Näherungslösung besteht in der Übertragung der in den Abschnitten 3 und 4 für den Interpolationsfehler gewonnenen Resultate auf die Abweichung der Näherungslösung von der exakten Lösung. Für die Maximums- und die $L_2$-Norm zeigen zwar numerische Experimente, daß diese Übertragung erlaubt ist, daß also auch der in einer dieser beiden Normen gemessene Fehler unserer Näherungslösung von der Größenordnung $O(h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^{d-1})$ ist, ein Beweis hierfür konnte bislang jedoch nicht angegeben werden. Hier zeigt sich ein weiterer Vorteil des energieorientierten Ansatzes. Denn die bei der Finite-Elemente-Methode übliche Abschätzung der Energienorm des Fehlers der gewonnenen Näherungslösung über die Energienorm des Interpolationsfehlers liefert auch in unserer Situation eine Schranke für die Energienorm des Fehlers der Näherungslösung.

Satz 6.2. *Sei $u^{(2)} \in X_0^{(2)}$ die exakte Lösung des Problems (40) mit $g \equiv 0$. Seien ferner $\tilde{u}_n^{(2)} \in \tilde{S}_n^{(2)}$ und $\hat{u}_n^{(2)} \in \hat{S}_n^{(2)}$ die von unserem Verfahren auf den dünnen Gittern $\tilde{G}_n^{(2)}$ beziehungsweise $\hat{G}_n^{(2)}$ produzierten Näherungslösungen. Dann gilt:*

$$\begin{aligned} \|u^{(2)} - \tilde{u}_n^{(2)}\|_E &\leq h_n \cdot \frac{|u^{(2)}|_\infty}{\sqrt{24}}, \\ \|u^{(2)} - \hat{u}_n^{(2)}\|_E &\leq h_n \cdot \frac{|u^{(2)}|_\infty \cdot \sqrt{6}}{8}. \end{aligned} \tag{61}$$

*Beweis.* Eine der grundlegenden Eigenschaften der Methode der Finiten Elemente im Zusammenhang mit Poisson-Problemen ist die Tatsache, daß die produzierte Näherungslösung von allen möglichen Kandidaten (Testfunktionen) den kleinsten Energienorm-Abstand zur exakten Lösung aufweist (siehe z.B. [35] und [1]). Deshalb folgt für die mittels unseres hier vorgestellten Verfahrens auf den verschiedenen dünnen Gittern errechneten Näherungslösungen sofort

$$\|u^{(2)} - \tilde{u}_n^{(2)}\|_E \;\leq\; \|u^{(2)} - \tilde{v}_n^{(2)}\|_E \qquad \forall \tilde{v}_n^{(2)} \in \tilde{S}_n^{(2)}$$

und

$$\|u^{(2)} - \hat{u}_n^{(2)}\|_E \;\leq\; \|u^{(2)} - \hat{v}_n^{(2)}\|_E \qquad \forall \hat{v}_n^{(2)} \in \hat{S}_n^{(2)}.$$

Diese Abschätzungen gelten insbesondere für die Interpolanten von $u^{(2)}$ in die verschiedenen Dünn-Gitter-Räume, also für $\tilde{v}_n^{(2)} = \tilde{u}_{n,I}^{(2)}$ und $\hat{v}_n^{(2)} = \hat{u}_{n,I}^{(2)}$. Somit erhalten wir

$$\|u^{(2)} - \tilde{u}_n^{(2)}\|_E \;\leq\; \|u^{(2)} - \tilde{u}_{n,I}^{(2)}\|_E$$

und

$$\|u^{(2)} - \hat{u}_n^{(2)}\|_E \;\leq\; \|u^{(2)} - \hat{u}_{n,I}^{(2)}\|_E,$$

woraus mit (35) sowie (36) unmittelbar die Behauptung folgt. □

Die Einschränkung auf den Spezialfall $g \equiv 0$ und damit auf homogene Dirichlet-Randbedingungen stellt keine wesentliche Beschränkung der Allgemeinheit dar. Denn im Falle $g \not\equiv 0$ läßt sich zu (40) sofort ein äquivalentes Problem mit homogenen Randbedingungen angeben. Dazu führen wir Interpolationsoperatoren $P_1$ und $P_2$ ein und definieren für eine auf $\delta\bar{\Omega}^{(2)}$ gegebene Funktion $w$ sowie $(x_1, x_2) \in \Omega^{(2)}$

$$\begin{aligned}
P_1 w(x_1, x_2) &:= (1 - x_1) \cdot w(0, x_2) \;+\; x_1 \cdot w(1, x_2), \\
P_2 w(x_1, x_2) &:= (1 - x_2) \cdot w(x_1, 0) \;+\; x_2 \cdot w(x_1, 1), \\
P_1 P_2 w(x_1, x_2) &= (1 - x_1) \cdot (1 - x_2) \cdot w(0,0) + x_1 \cdot (1 - x_2) \cdot w(1,0) + \\
&\quad\; (1 - x_1) \cdot x_2 \cdot w(0,1) + x_1 \cdot x_2 \cdot w(1,1).
\end{aligned}$$

Der Operator $G := P_1 + P_2 - P_1 P_2$ wird in der Literatur auch als *Gordon-Projektor* bezeichnet, die Interpolation mit $G$ heißt *transfinite Interpolation* (siehe z.B. [26], [13] sowie [14]). Wir wenden nun $G$ auf die Dirichlet-Randfunktion $g$ an und definieren

$$g_I(x_1, x_2) \;:=\; Gg(x_1, x_2), \qquad (x_1, x_2) \in \bar{\Omega}^{(2)}.$$

Aufgrund der Eigenschaften der transfiniten Interpolation gilt $g_I(x_1, x_2) = g(x_1, x_2)$ auf dem Rand des Einheitsquadrats. Deshalb ist $u$ genau dann Lösung unseres Modellproblems (40), wenn $v := u - g_I$ das folgende Problem löst:

$$\begin{aligned}
\Delta v(x_1, x_2) &= f(x_1, x_2) - \Delta g_I(x_1, x_2) \quad \text{für } (x_1, x_2) \in \Omega^{(2)} = (0,1) \times (0,1), \\
v(x_1, x_2) &= 0 \quad \text{für } (x_1, x_2) \in \delta\bar{\Omega}^{(2)}.
\end{aligned}$$

Damit ist gezeigt, daß das hier besprochene Verfahren zur Lösung von Poisson-Problemen auf dünnen Gittern eine Näherungslösung liefert, die bezüglich der Energienorm mit einem Fehler der Ordnung $O(h_n)$ behaftet ist. Da nun der Finite-Elemente-Ansatz bei Verwendung stückweise bilinearer Basisfunktionen auch auf vollen Gittern bezüglich der Energienorm zu einem Fehler der Ordnung $O(h_n)$ führt, leisten dünne Gitter in diesem Sinne größenordnungsmäßig gleich viel wie volle Gitter, obwohl dünne Gitter mit signifikant weniger Unbekannten auskommen. Mit anderen Worten: Um bezüglich der Energienorm eine vorgegebene Genauigkeit $\epsilon$ zu erzielen, benötigen dünne Gitter im zweidimensionalen Fall nur $O(\epsilon^{-1})$ Gitterpunkte und damit Unbekannte, wohingegen volle Gitter $O(\epsilon^{-2})$ Gitterpunkte erfordern.

Die Untersuchung der Genauigkeit der erzielten Näherungslösung liefert uns Aussagen über die Qualität der gewählten Diskretisierung. Zur Beurteilung des Iterationsverfahrens, mit der wir das aus der Diskretisierung resultierende Gleichungssystem lösen, ist die (algebraische) Konvergenzgeschwindigkeit, das heißt die Anzahl der Iterationen, die zur Verminderung des Fehlers um einen vorgegebenen Faktor $\epsilon$ erforderlich sind, von zentraler Bedeutung. In dieser Hinsicht optimal sind bekanntlich Mehrgittermethoden, bei denen die Anzahl der benötigten Iterationen von der Ordnung $O(1)$ ist und folglich nicht von der Auflösung beziehungsweise von der Anzahl der Verfeinerungsstufen abhängt.

Verwenden wir zur Lösung des linearen Gleichungssystems $Av = b$ ein lineares Iterationsverfahren

$$v^{(neu)} \;:=\; Mv^{(alt)} \;+\; Nb \tag{62}$$

wie das Gauß-Seidel-Verfahren, so wird sowohl zur Klärung der Frage, ob überhaupt Konvergenz vorliegt, als auch als Maß für die Konvergenzgeschwindigkeit in obigem Sinne üblicherweise der Spektralradius $\rho(M)$ der Iterationsmatrix $M$ betrachtet. Denn zum einen erweist sich die Bedingung

$$\rho(M) \;<\; 1 \tag{63}$$

als notwendig und hinreichend für die Konvergenz des Verfahrens (62), zum anderen erlaubt $\rho(M)$ auch Aussagen über das Ausmaß der Fehlerreduktion pro Iterationsschritt. Die folgenden Ausführungen haben das Ziel, $\rho(M)$ geeignet abzuschätzen.

Wir lösen das lineare Gleichungssystem $Av = b$ mit symmetrischer, positiv definiter Steifigkeitsmatrix $A$ mit Hilfe des Gauß-Seidel-Verfahrens. Bezeichne $\bar{v}$ die exakte Lösung dieses Systems, $v^{(i)}$ die $i$-te Näherungslösung sowie $e^{(i)} := v^{(i)} - \bar{v}$ den zugehörigen Fehler. Dann gilt (siehe z.B. [21])

$$M \;=\; I - (D+L)^{-1}A, \qquad N \;=\; (D+L)^{-1}, \tag{64}$$

wobei die Diagonalmatrix $D$ und die strikte untere Dreiecksmatrix $L$ eine additive Zerlegung der Matrix $A$ gemäß

$$A \;=\; D + L + L^T \tag{65}$$

bilden. Wie man leicht sieht, gilt für den Fehler $e^{(i+1)}$ die Beziehung

$$e^{(i+1)} \;=\; Me^{(i)}. \tag{66}$$

Wir skalieren nun zunächst die Matrix $A$ so, daß alle Diagonalelemente den Wert Eins annehmen:

$$\tilde{A} \;:=\; D^{-\frac{1}{2}} A D^{-\frac{1}{2}}. \tag{67}$$

Mit $\tilde{b} := D^{-\frac{1}{2}} b$ und $y := D^{\frac{1}{2}} v$ lautet das zu lösende Gleichungssystem nun

$$\tilde{A}y \;=\; D^{-\frac{1}{2}} A D^{-\frac{1}{2}} D^{\frac{1}{2}} v \;=\; D^{-\frac{1}{2}} b \;=\; \tilde{b}. \tag{68}$$

Für die Iterationsmatrix $\tilde{M} = I - (\tilde{D} + \tilde{L})^{-1}\tilde{A}$ des skalierten Systems $\tilde{A}y = \tilde{b}$ gilt somit wegen $\tilde{A} = \tilde{D} + \tilde{L} + \tilde{L}^T$, $\tilde{D} = D^{-\frac{1}{2}} D D^{-\frac{1}{2}} = I$ sowie $\tilde{L} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$

$$\tilde{M} \;=\; I - (\tilde{D} + \tilde{L})^{-1}\tilde{A} \;=\; D^{\frac{1}{2}} M D^{-\frac{1}{2}}. \tag{69}$$

Die Iterationsmatrizen $M$ und $\tilde{M}$ sind also ähnlich. Insbesondere erhalten wir für den Spektralradius $\rho(\tilde{M})$

$$\rho(\tilde{M}) \;=\; \rho(M). \tag{70}$$

Der nachfolgende Satz zeigt nun einen Weg zur Konstruktion einer Schranke für den Spektralradius $\rho(M)$ der Iterationsmatrix $M$.

Satz 6.3. *Sei $A$ die Steifigkeitsmatrix zum dünnen Gitter $\tilde{G}_n^{(2)}$, und seien $\tilde{A}$, $\tilde{D}$, $\tilde{L}$, $\tilde{M}$ und $M$ wie oben definiert. Unter der Zusatzvoraussetzung, daß $\frac{1}{c}\cdot\tilde{A} - (I+\tilde{L})(I+\tilde{L}^T)$ für eine reelle Zahl $c$ mit $0 < c < 1$ positiv semidefinit ist, gilt*

$$\rho(M) \;\le\; \|\tilde{A}^{\frac{1}{2}} \tilde{M} \tilde{A}^{-\frac{1}{2}}\|_2 \;\le\; \sqrt{1-c}. \tag{71}$$

*Beweis.* (70) und [21], Satz 4.4.21 sowie Lemma 4.4.23. □

Um mit Hilfe dieses Satzes $\rho(M)$ abschätzen zu können, bedarf es der Bestimmung eines geeigneten Wertes $c$. Dazu befassen wir uns in einem vorbereitenden Lemma zunächst mit der Zeilensummennorm der Matrizen $\tilde{L}$ und $\tilde{L}^T$.

Lemma 6.4. *Sei $A$ die Steifigkeitsmatrix zum dünnen Gitter $\tilde{G}_n^{(2)}$, und seien ferner $\tilde{A}$ sowie $\tilde{L}$ wie oben definiert. Dann gelten die Abschätzungen*

$$\|\tilde{L}\|_\infty \;\le\; \frac{39}{8} + \frac{9}{4}\cdot\sqrt{2}\,, \qquad \|\tilde{L}^T\|_\infty \;\le\; \frac{3}{4}\cdot\left(\frac{3}{2} + \sqrt{2}\right)\cdot 2^{\frac{n}{4}}. \tag{72}$$

*Beweis.* Wir betrachten Zeile $i$ der Matrix $A$, $1 \le i \le N_I$. Die dem Gitterpunkt $i$ zugeordnete Basisfunktion $\varphi_i$ erfülle $\varphi_i \in T_{i_1,i_2},\; 1 \le i_1, i_2 \le n,\; i_1 + i_2 \le n+1$. Nach Satz 6.1 gilt für das Diagonalelement $a_{ii}$ der Matrix $A$

$$a_{ii} \;=\; \frac{4}{3}\cdot\left(\frac{h_{i_1}}{h_{i_2}} + \frac{h_{i_2}}{h_{i_1}}\right) \;=\; \frac{4}{3}\cdot\left(2^{i_2-i_1} + 2^{i_1-i_2}\right),$$

für die Einträge $\tilde{a}_{ij}$ der Matrix $\tilde{A}$ ergibt sich nach (67)

$$\tilde{a}_{ii} \;=\; 1, \qquad \tilde{a}_{ij} \;=\; \frac{a_{ij}}{\sqrt{a_{ii}}\cdot\sqrt{a_{jj}}}\,.$$

Sei nun $j$, $1 \leq j \leq N_I$, ein weiterer, von $i$ verschiedener Gitterpunkt mit zugeordneter Basisfunktion $\varphi_j \in T_{j_1,j_2}$. Es gelte $a_{ij} \neq 0$. Aufgrund der Resultate aus Satz 6.1 sowie der hierarchischen Glättungsreihenfolge (Reihenfolge der Variablenmodifikation bei der Gauß-Seidel-Iteration, vgl. (41) und Abschnitt 6.2.2) ergibt sich folgender Zusammenhang:

$$
\begin{array}{lll}
j < i & \Leftrightarrow_{a_{ij}\neq 0} & (j_1 = i_1 \;\wedge\; j_2 < i_2) \;\vee\; (j_1 < i_1 \;\wedge\; j_2 = i_2), \\
j > i & \Leftrightarrow_{a_{ij}\neq 0} & (j_1 = i_1 \;\wedge\; j_2 > i_2) \;\vee\; (j_1 > i_1 \;\wedge\; j_2 = i_2).
\end{array} \tag{73}
$$

Somit enthält die $i$-te Zeile der Matrix $L$ genau die $a_{ij} \neq 0$, die von Gitterpunkten beziehungsweise Basisfunktionen aus hierarchisch höheren Teilräumen $T_{j_1,j_2}$ herrühren (siehe (2)). In der $i$-ten Zeile der Matrix $L^T$ stehen dementsprechend genau die $a_{ij} \neq 0$, die von Gitterpunkten beziehungsweise Basisfunktionen aus hierarchisch niedrigeren Teilräumen $T_{j_1,j_2}$ herrühren.

Tabelle 4

*Matrixeinträge $a_{ij} \neq 0$ in Zeile $i$, $j > i$ $(j_1 > i_1)$.*

| Teilraum | Anzahl der $j$ mit $a_{ij} \neq 0$ | $a_{ij}$ | Summe |
|---|---|---|---|
| $T_{i_1+1,i_2}$ | 2 | $2 \cdot \frac{h_{i_1+1}^2}{h_{i_1} \cdot h_{i_2}}$ | $\frac{h_{i_1}}{h_{i_2}} = 2^{i_2-i_1}$ |
| $T_{i_1+2,i_2}$ | 2 | $2 \cdot \frac{h_{i_1+2}^2}{h_{i_1} \cdot h_{i_2}}$ | $\frac{h_{i_1}}{h_{i_2}} = 2^{i_2-i_1}$ |
| | 2 | $2 \cdot \frac{3 \cdot h_{i_1+2}^2}{h_{i_1} \cdot h_{i_2}}$ | |
| $T_{i_1+3,i_2}$ | 2 | $2 \cdot \frac{h_{i_1+3}^2}{h_{i_1} \cdot h_{i_2}}$ | $\frac{h_{i_1}}{h_{i_2}} = 2^{i_2-i_1}$ |
| | 2 | $2 \cdot \frac{3 \cdot h_{i_1+3}^2}{h_{i_1} \cdot h_{i_2}}$ | |
| | 2 | $2 \cdot \frac{5 \cdot h_{i_1+3}^2}{h_{i_1} \cdot h_{i_2}}$ | |
| | 2 | $2 \cdot \frac{7 \cdot h_{i_1+3}^2}{h_{i_1} \cdot h_{i_2}}$ | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $T_{n+1-i_2,i_2}$ | ... | ... | $\frac{h_{i_1}}{h_{i_2}} = 2^{i_2-i_1}$ |

Tabelle 4 gibt über die $a_{ij} \neq 0$ mit $\varphi_j \in T_{i_1+l,i_2}$ und $l > 0$ Aufschluß. Man beachte dabei, daß für festes $l$ alle Diagonalelemente $a_{jj}$ denselben, durch $T_{i_1+l,i_2}$ festgelegten Wert annehmen. Im Spezialfall $i_1 = 1$ erhalten wir demnach

$$
\sum_{\substack{j\,:\,\varphi_j \in T_{1+l,i_2} \\ 0 < l \leq n-i_2}} \tilde{a}_{ij} \;=\; \sum_{\substack{j\,:\,\varphi_j \in T_{1+l,i_2} \\ 0 < l \leq n-i_2}} \frac{a_{ij}}{\sqrt{a_{ii}} \cdot \sqrt{a_{jj}}}
$$

$$
\begin{aligned}
&= \sum_{l=1}^{n-i_2} \frac{2^{i_2-1}}{\sqrt{\frac{4}{3}\cdot(2^{i_2-1}+2^{1-i_2})}\cdot\sqrt{\frac{4}{3}\cdot(2^{i_2-1-l}+2^{1+l-i_2})}} \\
&= \frac{3}{4}\cdot\frac{2^{i_2-1}}{\sqrt{2^{i_2-1}+2^{1-i_2}}}\cdot\sum_{l=1}^{n-i_2}\frac{1}{\sqrt{2^{i_2-1-l}+2^{1+l-i_2}}} \\
&\leq \frac{3}{4}\cdot\frac{2^{i_2-1}}{\sqrt{2^{i_2-1}}}\cdot\sum_{l=1}^{n-i_2}\frac{1}{\sqrt{2^{i_2-1-l}+2^{1+l-i_2}}} \\
&= \frac{3}{4\cdot\sqrt{2}}\cdot\sqrt{2^{i_2}}\cdot\sum_{l=2-i_2}^{n+1-2i_2}\frac{1}{\sqrt{2^l+2^{-l}}} \\
&\leq \frac{3}{4\cdot\sqrt{2}}\cdot\sum_{l=2-i_2}^{n+1-2i_2}\left(\frac{1}{\sqrt{2}}\right)^{|l|-i_2} \\
&\leq \frac{3}{4\cdot\sqrt{2}}\cdot\left(2^{\frac{n}{4}}\cdot\left(\frac{3}{2}\cdot\sqrt{2}+2\right)-\sqrt{2}-1\right) \\
&= \frac{3}{4}\cdot\left(\frac{3}{2}+\sqrt{2}\right)\cdot 2^{\frac{n}{4}}-\frac{3}{8}\cdot(2+\sqrt{2}),
\end{aligned}
$$

da nach kurzer Rechnung folgt, daß die Summe für $i_2 = \frac{n}{2}$ ($n$ gerade) beziehungsweise $i_2 = \frac{n+1}{2}$ ($n$ ungerade) maximal wird. Analog ergibt sich

$$
\begin{aligned}
\sum_{\substack{j\,:\,\varphi_j\,\in T_{1,i_2+l}\\ 0<l\leq n-i_2}} \tilde{a}_{ij} &= \sum_{\substack{j\,:\,\varphi_j\,\in T_{1,i_2+l}\\ 0<l\leq n-i_2}} \frac{a_{ij}}{\sqrt{a_{ii}}\cdot\sqrt{a_{jj}}} \\
&= \frac{3}{4}\cdot\frac{2^{1-i_2}}{\sqrt{2^{i_2-1}+2^{1-i_2}}}\cdot\sum_{l=1}^{n-i_2}\frac{1}{\sqrt{2^{i_2+l-1}+2^{1-i_2-l}}} \\
&\leq \frac{3}{4}\cdot\frac{2^{1-i_2}}{\sqrt{2^{i_2-1}+2^{1-i_2}}}\cdot\sum_{l=1}^{n-i_2}\frac{1}{\sqrt{2^{i_2+l-1}}} \\
&\leq \frac{3}{4}\cdot\frac{2^{1-i_2}}{\sqrt{2^{i_2-1}+2^{1-i_2}}}\cdot\frac{1}{\sqrt{2^{i_2}}}\cdot\sum_{l=0}^{\infty}\left(\frac{1}{\sqrt{2}}\right)^l \\
&= \frac{3}{4}\cdot(2+\sqrt{2})\cdot\frac{2^{1-i_2}}{\sqrt{2^{i_2-1}+2^{1-i_2}}}\cdot\frac{1}{\sqrt{2^{i_2-1}+2^{i_2-1}}} \\
&\leq \frac{3}{4}\cdot(2+\sqrt{2})\cdot\frac{2^{1-i_2}}{\sqrt{2^{i_2-1}+2^{1-i_2}}}\cdot\frac{1}{\sqrt{2^{i_2-1}+2^{1-i_2}}} \\
&= \frac{3}{4}\cdot(2+\sqrt{2})\cdot\frac{2^{1-i_2}}{2^{i_2-1}+2^{1-i_2}} \\
&\leq \frac{3}{8}\cdot(2+\sqrt{2}).
\end{aligned}
$$

Bei der Berechnung von $\|\tilde{L}^T\|_\infty$ ist es aus Symmetriegründen und wegen der Hierarchie der Teilräume ausreichend, die Suche nach dem Maximum auf diejenigen Zeilen $i$ zu erstrecken, die $\varphi_i \in T_{1,i_2}$ erfüllen. Mit (73) folgt nun

$$
\|\tilde{L}^T\|_\infty = \max_{1\leq i\leq N_I} \sum_{j>i} \tilde{a}_{ij}
$$

$$
\begin{aligned}
&= \max_{1\le i_2\le n} \max_{i:\varphi_i\in T_{1,i_2}} \sum_{j>i} \tilde a_{ij} \\
&= \max_{1\le i_2\le n} \max_{i:\varphi_i\in T_{1,i_2}} \left( \sum_{\substack{j\,:\,\varphi_j\,\in T_{1+l,i_2}\\ 0<l\le n-i_2}} \tilde a_{ij} + \sum_{\substack{j\,:\,\varphi_j\,\in T_{1,i_2+l}\\ 0<l\le n-i_2}} \tilde a_{ij} \right) \\
&\le \max_{1\le i_2\le n} \left( \frac{3}{4}\cdot\left(\frac{3}{2}+\sqrt{2}\right)\cdot 2^{\frac{n}{4}} - \frac{3}{8}\cdot(2+\sqrt{2}) + \frac{3}{8}\cdot(2+\sqrt{2}) \right) \\
&= \frac{3}{4}\cdot\left(\frac{3}{2}+\sqrt{2}\right)\cdot 2^{\frac{n}{4}}.
\end{aligned}
$$

Tabelle 5
*Matrixeinträge $a_{ij}\neq 0$ in Zeile $i$, $j<i$ $(j_1<i_1)$.*

| Teilraum | Anzahl der $j$ mit $a_{ij}\neq 0$ | $a_{ij}$ | Summe |
|---|---|---|---|
| $T_{i_1-1,i_2}$ | 1 | $2\cdot\frac{h_{i_1}^2}{2\cdot h_{i_1}\cdot h_{i_2}}$ | $\frac{h_{i_1}}{h_{i_2}}\le 2\cdot 2^{i_2-i_1}$ |
| $T_{i_1-2,i_2}$ | 1 | $\le 2\cdot\frac{3\cdot h_{i_1}^2}{4\cdot h_{i_1}\cdot h_{i_2}}$ | $\le\frac{3}{2}\cdot\frac{h_{i_1}}{h_{i_2}}\le 2\cdot 2^{i_2-i_1}$ |
| $T_{i_1-3,i_2}$ | 1 | $\le 2\cdot\frac{7\cdot h_{i_1}^2}{8\cdot h_{i_1}\cdot h_{i_2}}$ | $\le\frac{7}{4}\cdot\frac{h_{i_1}}{h_{i_2}}\le 2\cdot 2^{i_2-i_1}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $T_{1,i_2}$ | 1 | $\le 2\cdot\frac{h_{i_1}}{h_{i_2}}$ | $\le 2\cdot\frac{h_{i_1}}{h_{i_2}}=2\cdot 2^{i_2-i_1}$ |

Tabelle 5 zeigt die $a_{ij}\neq 0$ mit $\varphi_j\in T_{i_1-l,i_2}$ und $l>0$. Wir erhalten für $i_1\le\frac{n}{2}$ und $i_2=n+1-i_1$

$$
\begin{aligned}
\sum_{\substack{j\,:\,\varphi_j\,\in T_{i_1-l,n+1-i_1}\\ 0<l<i_1}} \tilde a_{ij} &= \sum_{\substack{j\,:\,\varphi_j\,\in T_{i_1-l,n+1-i_1}\\ 0<l<i_1}} \frac{a_{ij}}{\sqrt{a_{ii}}\cdot\sqrt{a_{jj}}} \\
&\le \sum_{l=1}^{i_1-1} \frac{2\cdot 2^{n+1-i_1-i_1}}{\sqrt{\frac{4}{3}\cdot(2^{n+1-2i_1}+2^{2i_1-n-1})}\cdot\sqrt{\frac{4}{3}\cdot(2^{n+1-2i_1+l}+2^{2i_1-l-n-1})}} \\
&\le \frac{3}{2}\cdot\frac{2^{n+1-2i_1}}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}\cdot\sum_{l=1}^{i_1-1}\frac{1}{\sqrt{2^{n+1-2i_1+l}}} \\
&\le \frac{3}{2}\cdot\frac{2^{n+1-2i_1}}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}\cdot\frac{1}{\sqrt{2^{n+1-2i_1+1}}}\cdot\sum_{l=0}^{\infty}\left(\frac{1}{\sqrt{2}}\right)^l \\
&= \frac{3}{2}\cdot(2+\sqrt{2})\cdot\frac{2^{n+1-2i_1}}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}\cdot\frac{1}{\sqrt{2^{n+1-2i_1}+2^{n+1-2i_1}}} \\
&\le \frac{3}{2}\cdot(2+\sqrt{2})\cdot\frac{2^{n+1-2i_1}}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}\cdot\frac{1}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{3}{2}\cdot(2+\sqrt{2})\cdot\frac{2^{n+1-2i_1}}{2^{n+1-2i_1}+2^{2i_1-n-1}} \\
&\leq \frac{3}{2}\cdot(2+\sqrt{2})
\end{aligned}
$$

und analog

$$
\begin{aligned}
\sum_{\substack{j\,:\,\varphi_j\,\in\,T_{i_1,n+1-i_1-l}\\ 0<l<n+1-i_1}} \tilde{a}_{ij} &= \sum_{\substack{j\,:\,\varphi_j\,\in\,T_{i_1,n+1-i_1-l}\\ 0<l<n+1-i_1}} \frac{a_{ij}}{\sqrt{a_{ii}}\cdot\sqrt{a_{jj}}} \\
&\leq \frac{3}{2}\cdot\frac{2^{2i_1-n-1}}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}\cdot\sum_{l=1}^{n-i_1}\frac{1}{\sqrt{2^{n+1-2i_1-l}+2^{2i_1+l-n-1}}} \\
&\leq \frac{3}{2}\cdot\frac{2^{2i_1-n-1}}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}\cdot\left(\frac{1}{\sqrt{2}}+2\cdot\sum_{l=1}^{\infty}\frac{1}{\sqrt{2^l+2^{-l}}}\right) \\
&\leq \frac{3}{2}\cdot\frac{2^{2i_1-n-1}}{\sqrt{2^{n+1-2i_1}+2^{2i_1-n-1}}}\cdot\left(\frac{1}{\sqrt{2}}+2\cdot\sum_{l=1}^{\infty}\left(\frac{1}{\sqrt{2}}\right)^l\right) \\
&\leq \frac{3}{2}\cdot\left(\frac{1}{\sqrt{2}}+2+2\cdot\sqrt{2}\right)\cdot\frac{2^{2i_1-n-1}}{\sqrt{2^{n+1-2i_1}}} \\
&\leq \frac{3}{4}\cdot\left(\frac{1}{2}+\sqrt{2}+2\right) \\
&= \frac{15}{8}+\frac{3}{4}\cdot\sqrt{2}.
\end{aligned}
$$

Bei der Berechnung von $\|\tilde{L}\|_\infty$ ist es aus Symmetriegründen und wegen der Hierarchie der Teilräume ausreichend, die Suche nach dem Maximum auf diejenigen Zeilen $i$ zu erstrecken, die $\varphi_i \in T_{i_1,n+1-i_1}$ und $i_1 \leq \frac{n}{2}$ erfüllen. Mit (73) folgt

$$
\begin{aligned}
\|\tilde{L}\|_\infty &= \max_{1\leq i\leq N_I}\sum_{j<i}\tilde{a}_{ij} \\
&= \max_{1\leq i_1\leq\frac{n}{2}}\ \max_{i:\varphi_i\in T_{i_1,n+1-i_1}}\sum_{j<i}\tilde{a}_{ij} \\
&= \max_{1\leq i_1\leq\frac{n}{2}}\ \max_{i:\varphi_i\in T_{i_1,n+1-i_1}}\left(\sum_{\substack{j\,:\,\varphi_j\,\in\,T_{i_1-l,n+1-i_1}\\ 0<l<i_1}}\tilde{a}_{ij} + \sum_{\substack{j\,:\,\varphi_j\,\in\,T_{i_1,n+1-i_1-l}\\ 0<l<n+1-i_1}}\tilde{a}_{ij}\right) \\
&\leq \frac{3}{2}\cdot(2+\sqrt{2}) + \frac{15}{8}+\frac{3}{4}\cdot\sqrt{2} \\
&= \frac{39}{8}+\frac{9}{4}\cdot\sqrt{2}.
\end{aligned}
$$

□

Leider hängt die obere Schranke für $\|\tilde{L}^T\|_\infty$ auch nach der Skalierung (67) exponentiell von $n$ ab. Schuld daran sind die extrem schlanken Rechtecke als Basiselemente (d.h. die Teilräume $T_{i_1,i_2}$ mit $i_1 \ll i_2$ beziehungsweise $i_2 \ll i_1$). Bei der Bildung der Zeilensummen reicht hier die Division durch $\sqrt{a_{jj}}$ nicht aus, eine logarithmische

Abhängigkeit oder gar $O(1)$ zu erhalten. Numerische Experimente zeigen dabei, daß die Schranken aus Lemma 6.4 größenordnungsmäßig scharf sind, daß also $\|\tilde{L}^T\|_\infty = O(2^{\frac{n}{4}})$ sowie $\|\tilde{L}\|_\infty = O(1)$ gilt. Auf jeden Fall sind wir jetzt in der Lage, $\rho(M)$ abzuschätzen.

SATZ 6.5. *Sei $A$ die Steifigkeitsmatrix zum dünnen Gitter $\tilde{G}_n^{(2)}$, und seien $\tilde{A}$ und $M$ wie oben definiert. Dann gilt für den Spektralradius $\rho(M)$*

$$\rho(M) \;\leq\; \sqrt{1 - \frac{1}{2^{\frac{n}{4}} \cdot \frac{18}{\lambda_{min}(\tilde{A})} + 1}} \quad , \tag{74}$$

*wobei $\lambda_{min}(\tilde{A})$ den kleinsten Eigenwert der Matrix $\tilde{A}$ bezeichnet.*

*Beweis.* Wir definieren zunächst

$$c \;:=\; \frac{1}{2^{\frac{n}{4}} \cdot \frac{18}{\lambda_{min}(\tilde{A})} + 1} \quad .$$

Mit Lemma 6.4 gilt demzufolge für den größten Eigenwert $\lambda_{max}(\tilde{L}\tilde{L}^T)$ der positiv semidefiniten Matrix $\tilde{L}\tilde{L}^T$

$$\begin{aligned} \lambda_{max}(\tilde{L}\tilde{L}^T) \;=\; \rho(\tilde{L}\tilde{L}^T) \;&\leq\; \|\tilde{L}\tilde{L}^T\|_\infty \\ &\leq\; \|\tilde{L}\|_\infty \cdot \|\tilde{L}^T\|_\infty \\ &\leq\; 18 \cdot 2^{\frac{n}{4}} \\ &=\; \left(\frac{1}{c} - 1\right) \cdot \lambda_{min}(\tilde{A}). \end{aligned}$$

Somit erhalten wir für alle $v \neq 0$

$$\frac{v^T \tilde{L}\tilde{L}^T v}{v^T v} \;\leq\; \lambda_{max}(\tilde{L}\tilde{L}^T) \;\leq\; \left(\frac{1}{c} - 1\right) \cdot \lambda_{min}(\tilde{A}) \;\leq\; \left(\frac{1}{c} - 1\right) \cdot \frac{v^T \tilde{A} v}{v^T v}$$

und insbesondere

$$v^T \left(\left(\frac{1}{c} - 1\right) \cdot \tilde{A} - \tilde{L}\tilde{L}^T\right) v \;\geq\; 0.$$

Aufgrund der Beziehung

$$\begin{aligned} \left(\frac{1}{c} - 1\right) \cdot \tilde{A} - \tilde{L}\tilde{L}^T \;&=\; \frac{1}{c} \cdot \tilde{A} - \tilde{A} - \tilde{L}\tilde{L}^T \\ &=\; \frac{1}{c} \cdot \tilde{A} - (I + \tilde{L} + \tilde{L}^T) - \tilde{L}\tilde{L}^T \\ &=\; \frac{1}{c} \cdot \tilde{A} - (I + \tilde{L})(I + \tilde{L}^T) \end{aligned}$$

ist folglich auch die Matrix $\frac{1}{c} \cdot \tilde{A} - (I + \tilde{L})(I + \tilde{L}^T)$ positiv semidefinit.

Damit sind nun aber mit obiger Wahl von $c$ alle Voraussetzungen des Satzes 6.3 erfüllt, und wir erhalten als Abschätzung

$$\rho(M) \;\leq\; \sqrt{1 - c} \;=\; \sqrt{1 - \frac{1}{2^{\frac{n}{4}} \cdot \frac{18}{\lambda_{min}(\tilde{A})} + 1}} \quad ,$$

womit Satz 6.5 vollständig bewiesen ist. □

Um die Größenordnung der Schranke aus (74) genauer fassen zu können, benötigen wir noch eine Abschätzung des kleinsten Eigenwertes $\lambda_{min}(\tilde{A})$ der skalierten Steifigkeitsmatrix $\tilde{A}$. Zu diesem Zweck verwenden wir ein Resultat aus [29], wo die Existenz einer Konstanten $\gamma$ gezeigt wurde mit

$$\lambda_{min}(\tilde{A}) \;\geq\; \gamma \cdot 2^{-n}. \tag{75}$$

Diese Abschätzung führt jetzt zu

KOROLLAR 6.6. *Es gibt eine positive Zahl $\gamma$, so daß für den Spektralradius der Gauß-Seidel-Iterationsmatrix $M$ bezüglich des dünnen Gitters $\tilde{G}_n^{(2)}$ gilt:*

$$\rho(M) \;\leq\; \sqrt{1 - \frac{1}{2^{\frac{5}{4}n} \cdot \frac{18}{\gamma} + 1}}\;. \tag{76}$$

*Beweis.* Satz 6.5 und (75). □

Wir wollen nun (76) dazu verwenden, eine Aussage über die (algebraische) Konvergenzgeschwindigkeit des vorgestellten Verfahrens zu machen. Unter dem Begriff Konvergenzgeschwindigkeit soll dabei die Anzahl der Iterationen verstanden werden, die zur Verminderung des Fehlers um einen gegebenen Faktor $\epsilon$ benötigt werden.

Aus (66) folgt zunächst

$$\|e^{(m)}\|_2 \;\leq\; \|M\|_2^m \cdot \|e^{(0)}\|_2.$$

Da es aber zu jedem $\delta > 0$ ein $m_\delta$ gibt, so daß

$$\|e^{(m)}\|_2 \;\leq\; (\rho(M) + \delta)^m \cdot \|e^{(0)}\|_2$$

für alle $m > m_\delta$ (siehe z.B. [21]), wird als Maß für die (asymptotische) Konvergenzgeschwindigkeit im oben genannten Sinne üblicherweise der Spektralradius $\rho(M)$ herangezogen. So erhält man aus der Bedingung

$$\rho(M)^{m_\epsilon} \;<\; \epsilon$$

mit $0 < \epsilon < 1$ eine Forderung an die Iterationsanzahl $m_\epsilon$ (vgl. [34], [21]; beachte, daß $0 < \rho(M) < 1$):

$$m_\epsilon \;>\; \frac{\ln(\epsilon)}{\ln(\rho(M))}\;. \tag{77}$$

Wir formulieren das Ergebnis dieser Überlegungen in einem Satz.

SATZ 6.7. *Seien $\gamma$ gemäß (76) und $\epsilon$ mit $0 < \epsilon < 1$ gegeben. $m_\epsilon$ Iterationen mit der Gauß-Seidel-Iterationsmatrix $M$ bezüglich des dünnen Gitters $\tilde{G}_n^{(2)}$,*

$$m_\epsilon \;>\; -\ln(\epsilon) \cdot \left(2^{\frac{5}{4}n} \cdot \frac{36}{\gamma} + 2\right), \tag{78}$$

*reduzieren den Startfehler $\|e^{(0)}\|_2$ mindestens um den Faktor $\epsilon$.*

*Beweis.* Aus Korollar 6.6 folgt mit $y := \left(2^{\frac{5}{4}n} \cdot \frac{18}{\gamma} + 1\right)^{-1}$

$$\begin{aligned}
\frac{\ln(\epsilon)}{\ln(\rho(M))} \leq \frac{\ln(\epsilon)}{\ln\sqrt{1-y}} &= \frac{2 \cdot \ln(\epsilon)}{\ln(1-y)} \\
&= \frac{2 \cdot \ln(\epsilon)}{-\sum_{i=1}^{\infty} \frac{y^i}{i}} \qquad (0 < y < 1) \\
&= \frac{-2 \cdot \ln(\epsilon)}{y} \cdot \frac{1}{1 + \sum_{i=1}^{\infty} \frac{y^i}{i+1}} \\
&\leq \frac{-2 \cdot \ln(\epsilon)}{y} \\
&= -\ln(\epsilon) \cdot \left(2^{\frac{5}{4}n} \cdot \frac{36}{\gamma} + 2\right).
\end{aligned}$$

Zusammen mit (77) ergibt sich die Behauptung. □

Bevor wir nun anhand der Resultate aus diesem Abschnitt unser Verfahren kritisch beurteilen und mit anderen Methoden vergleichen, erwähnen wir noch ein Ergebnis aus [29], wo eine obere Schranke für die (spektrale) Kondition der skalierten Steifigkeitsmatrix $\tilde{A}$ angegeben wird.

SATZ 6.8. *Sei $A$ die Steifigkeitsmatrix zum dünnen Gitter $\tilde{G}_n^{(2)}$, und sei $\tilde{A}$ wie in (67) definiert. Dann gibt es eine positive Konstante $C > 0$, so daß für die spektrale Kondition $\kappa(\tilde{A})$ der Matrix $\tilde{A}$ gilt:*

$$\kappa(\tilde{A}) := \rho(\tilde{A}) \cdot \rho(\tilde{A}^{-1}) = \frac{\lambda_{max}(\tilde{A})}{\lambda_{min}(\tilde{A})} \leq C \cdot n^2 \cdot 2^n. \tag{79}$$

*Beweis.* Siehe [29]. □

Diese Abschätzung wird vor allem dann interessant, wenn wir anstelle der Gauß-Seidel-Iteration das Verfahren der konjugierten Gradienten (cg) zur Lösung des linearen Gleichungssystems heranziehen; denn dort ist bekanntlich die Kondition der Steifigkeitsmatrix die für die Konvergenzgeschwindigkeit relevante Größe. Die Implementierung von cg auf dünnen Gittern ist beabsichtigt, wurde aber bislang noch nicht realisiert.

**6.2.5. Diskussion des vorgestellten Verfahrens.** In unserem Verfahren verwenden wir hierarchische Basen und dünne Gitter, um mittels der Finite-Elemente-Diskretisierung ein lineares Gleichungssystem zu erhalten. Dieses lösen wir dann mit Hilfe der Gauß-Seidel-Iteration.

Die Hierarchie der Teilräume gestattet dabei in natürlicher Weise eine Mehrgittersehweise des Algorithmus', wenn wir alle Teilräume $T_{i_1,i_2}$ mit $i_1 + i_2 = s$ als eine Ebene auffassen. Im Gegensatz zu üblichen Mehrgitterverfahren, bei denen eine Variable auf mehreren Ebenen auftreten kann (nämlich auf der gröbsten Ebene, auf der sie noch existiert, und auf allen feineren), ist in unserer Mehrgittersehweise eine Variable $i$ eindeutig einem Teilraum $T_{i_1,i_2}$ und somit einer Gitterebene $s$ zugeordnet. In diesem Sinne nun kann unser Verfahren als Erweiterung der in [5] vorgestellten HB-MG-Methode von Bank, Dupont und Yserentant angesehen werden. Das dortige Prinzip wird dabei auf rechteckige Elemente, Produktbasen und dünne Gitter verallgemeinert.

Die Verknüpfung von hierarchischen Basen mit dünnen Gittern legt einen Vergleich des hier vorgestellten Verfahrens mit zwei Voll-Gitter-Techniken zur Lösung von aus der Finite-Elemente-Diskretisierung resultierenden linearen Gleichungssystemen nahe. Zum einen wollen wir die Standard-Gauß-Seidel-Iteration als klassische Methode betrachten, zum anderen ziehen wir mit Yserentants hierarchischen Basen als cg-Vorkonditionierer eine moderne Alternative zum Vergleich heran.

Ausgehend von Stützpunktbasen auf vollen (quadratischen) Gittern diskretisieren wir also unser Modellproblem gemäß der Methode der Finiten Elemente und lösen das entstehende System mit dem Gauß-Seidel-Verfahren (im folgenden kurz $GS$). Hier gilt (siehe z.B. [34] und [20])

$$\begin{aligned} \kappa(A^{GS}) &= O\left(4^n\right), \\ \rho(M^{GS}) &= O\left(1 - \frac{1}{4^n}\right), \\ m_\epsilon^{GS} &= O\left(4^n\right). \end{aligned}$$

Führen wir die Finite-Elemente-Diskretisierung auf vollen (Dreiecks-) Gittern bezüglich hierarchischer Basen durch und lösen wir das entstehende System mit Hilfe des (vorkonditionierten) cg-Verfahrens (im folgenden kurz $HB$), so ergibt sich dagegen eine polynomiale Abhängigkeit von $n$ (siehe [36], [5] und [6]):

$$\begin{aligned} \kappa(A^{HB}) &= O\left(n^2\right), \\ m_\epsilon^{HB} &= O\left(n\right). \end{aligned}$$

Fassen wir nun für unser Dünn-Gitter-Verfahren ($DG$) die Ergebnisse aus Korollar 6.6 sowie den Sätzen 6.7 und 6.8 zusammen, so stellen wir für die drei zentralen Größen $\rho$, $m_\epsilon$ und $\kappa$ das folgende Verhalten fest:

$$\begin{aligned} \kappa(\tilde{A}^{DG}) &= O\left(n^2 \cdot 2^n\right), \\ \rho(M^{DG}) &= O\left(1 - \frac{1}{2^{\frac{5}{4}n}}\right), \\ m_\epsilon^{DG} &= O\left(2^{\frac{5}{4}n}\right). \end{aligned}$$

Dabei ist zu beachten, daß aufgrund der für sehr kleine positive $\delta$ geltenden Beziehung $\sqrt{1-\delta} \approx 1 - \frac{\delta}{2}$ die Wurzel bei der Abschätzung (76) des Spektralradius' $\rho(M)$ keinen Einfluß auf die Ordnung hat und deshalb bei Ordnungsbetrachtungen für große $n$ vernachlässigt werden kann.

Wir fügen als Bemerkung an, daß das Konvergenzverhalten unseres Verfahrens in den numerischen Experimenten (siehe Abschnitt 9) erheblich besser ist, als es die hier ermittelte Schranke für den Spektralradius der Iterationsmatrix vermuten läßt. Zudem verschlechtert sich die Konvergenzgeschwindigkeit im dreidimensionalen Fall nur leicht.

Um ein Maß für den Gesamtaufwand der drei diskutierten Verfahren zu erhalten, multiplizieren wir schließlich die Anzahl $m_\epsilon$ der erforderlichen Iterationen mit der Anzahl der Unbekannten und bezeichnen das Ergebnis mit $W$. Es ergibt sich

$$\begin{aligned}
W^{GS} &= m_\epsilon^{GS} \cdot O\left(4^n\right) &&= O\left(16^n\right), \\
W^{HB} &= m_\epsilon^{HB} \cdot O\left(4^n\right) &&= O\left(n \cdot 4^n\right), \\
W^{DG} &= m_\epsilon^{DG} \cdot O\left(n \cdot 2^n\right) &&= O\left(n \cdot 4^n \cdot 2^{\frac{n}{4}}\right).
\end{aligned}$$

Dies zeigt, daß unser Dünn-Gitter-Verfahren sowohl bezüglich der Konvergenzgeschwindigkeit als auch im Hinblick auf den Gesamtaufwand zwischen der Standard-Gauß-Seidel-Iteration bei Stützpunktbasen und den hierarchischen Basen als cg-Vorkonditionierer anzusiedeln ist. Man sieht aber auch, daß die Konvergenzgeschwindigkeit unseres Verfahrens verbessert werden muß , wenn wir zu einer leistungs- und konkurrenzfähigen Methode auf dünnen Gittern gelangen wollen. Hier liegt der Übergang zu „echten“ Mehrgitterverfahren auf dünnen Gittern nahe, worauf wir abschließend noch kurz eingehen wollen.

Das reine Gauß-Seidel-Verfahren zum Vergleich heranzuziehen erscheint insofern als problematisch, als dieses heute kaum mehr als Löser, sondern nur noch als Glätter — integriert in einen Mehrgitteralgorithmus — auftritt, wodurch bekanntlich für den Spektralradius der Mehrgitter-Iterationsmatrix und damit für die Konvergenzgeschwindigkeit $O(1)$ erzielt werden kann. Dieses Prinzip kann jedoch auch auf unser Verfahren angewandt werden. Dazu müssen wir unsere obige Mehrgittersehweise nur dahingehend ändern, daß jede Variable $i$ mit zugeordneter Basisfunktion $\varphi_i \in T_{i_1,i_2}$ nicht nur auf der Ebene $s := i_1 + i_2$, sondern auch auf allen „feineren“ Ebenen $s+1, ..., n+1$ relaxiert wird. Die dafür nötigen Änderungen an den Datenstrukturen wurden in [12] und [32] untersucht, Implementierungen für Parallel- und Vektorrechner wurden in [16] sowie [17] vorgestellt.

Vergleichen wir nun Konvergenzgeschwindigkeit und Gesamtaufwand des Standard-Mehrgitterverfahrens auf vollen Gittern $(MG-GS)$ mit der erwähnten Mehrgittererweiterung unseres Dünn-Gitter-Verfahrens $(MG-DG)$, so ergibt sich

$$\begin{aligned}
m_\epsilon^{MG-GS} &= O(1), \\
m_\epsilon^{MG-DG} &= O(1), \\
W^{MG-GS} &= O(4^n), \\
W^{MG-DG} &= O(n \cdot 2^n).
\end{aligned}$$

Wir halten als Ergebnis fest, daß die Mehrgittervariante unseres Verfahrens die Speichervorteile dünner Gitter mit der von $n$ unabhängigen Mehrgitter-Konvergenzgeschwindigkeit verbindet. Somit bieten dünne Gitter sehr effiziente Möglichkeiten zur Lösung von Poisson-Problemen. Die Attraktivität dünner Gitter steigt dabei mit wachsender Dimensionalität $d$ und wachsender Auflösung $n$.

**7. Dreidimensionale Poisson-Probleme auf dünnen Gittern.** Im vorigen Abschnitt haben wir ein Verfahren zur Lösung zweidimensionaler Poisson-Probleme auf dünnen Gittern vorgestellt. Ein wesentlicher Vorteil der dabei verwendeten rechteckigen Elemente und der damit verbundenen Produktbasen gegenüber den gebräuchlicheren Triangulationen besteht in der leichteren Übertragbarkeit der gewonnenen Methoden und Erkenntnisse auf drei- und höherdimensionale Probleme. So führt die Verallgemeinerung von Dreiecksnetzen im dreidimensionalen Fall zu Netzen von Tetraedern, deren Verfeinerung bereits ein nicht-triviales Problem darstellt und die räumliche Vorstellungskraft des Menschen zumindest strapaziert. Wie wir in Abschnitt 2 gesehen haben, gestattet dagegen der Zugang über kartesische Produkte von Intervallen als Basiselemente eine einfache Darstellung von Gittern, Räumen und Basisfunktionen bei beliebiger Dimensionalität $d$. Insbesondere kann die Verfeinerung von Quadern im dreidimensionalen Falle direkt aus der Verfeinerung von Rechtecken im Falle $d = 2$ abgeleitet werden. Zwar erweisen sich Triangulationen als flexibler in bezug auf die zugrundeliegende Geometrie, doch läßt sich dieser Nachteil durch Adaption am Rand ausgleichen. So wurde in [10] unser (zweidimensionales) Dünn-Gitter-Verfahren erfolgreich bei Problemen mit krummlinig berandeten Gebieten eingesetzt.

Während die theoretischen Betrachtungen zu dünnen Gittern im ersten Teil dieser Arbeit unabhängig von der Dimensionalität $d$ durchgeführt wurden, haben wir uns bei der Vorstellung des Poisson-Lösers aus Gründen der Übersichtlichkeit auf den zweidimensionalen Fall beschränkt. Das produktorientierte Konzept gestattet jedoch auch hier die Verallgemeinerung auf den Fall $d > 2$. Dabei bleibt der grundsätzliche Ablauf des Algorithmus' erhalten. Wir verzichten deshalb an dieser Stelle auf eine detaillierte Beschreibung des dreidimensionalen Verfahrens und zeigen nur die wesentlichen Gesichtspunkte sowie neu zu beachtende Aspekte auf.

In diesem Abschnitt sei also folgendes Poisson-Modellproblem auf dem Einheitskubus gegeben:

$$\begin{aligned} \Delta u(x_1, x_2, x_3) &= f(x_1, x_2, x_3) \;\text{ für }\; (x_1, x_2, x_3) \in \Omega^{(3)} = (0,1)^3, \\ u(x_1, x_2, x_3) &= g(x_1, x_2, x_3) \;\text{ für }\; (x_1, x_2, x_3) \in \delta\bar{\Omega}^{(3)}. \end{aligned} \tag{80}$$

Wiederum betrachten wir zunächst die Gittergenerierung. Hierbei gehen wir insbesondere auf die Darstellung der Dirichlet'schen Randfunktion $g$ ein. Denn im Gegensatz zum zweidimensionalen Fall, wo jede Seite des Einheitsquadrats eindimensional ist (weswegen nach (15) auf dem Rand kein Unterschied zwischen vollem und dünnem Gitter besteht), setzt sich der Rand des Einheitswürfels aus Flächen — also zweidimensionalen Gebilden — zusammen. Das resultierende dreidimensionale dünne Gitter nimmt auf den Seiten des Einheitswürfels folglich die Gestalt eines zweidimensionalen dünnen Gitters an.

Sodann befassen wir uns mit dem eigentlichen Lösungsverfahren. Hierbei stehen wiederum die Diskretisierung des Differentialgleichungsoperators, die Diskretisierung der rechten Seite $f$ sowie die iterative Lösung des linearen Gleichungssystems im Vordergrund. Schließlich erörtern wir noch Konvergenzfragen.

**7.1. Die sukzessive Erzeugung des dünnen Gitters.** Analog zu unserer Vorgehensweise im zweidimensionalen Fall verwenden wir als Startnäherung für die Lösung $u$ von (80) jetzt den trilinearen Interpolanten aus den acht Eckwerten der Dirichlet'schen Randfunktion $g$. Konstruiert wird anfangs also ein Baum, in dem nur die acht Eckpunkte des Einheitswürfels durch Knoten repräsentiert sind. Knoten für weitere Punkte auf dem Rand von $\bar{\Omega}^{(3)}$ werden erst dann während des Iterationsprozesses erzeugt und in den Baum eingehängt, wenn die Berechnung des hierarchischen Überschusses in einem inneren Gitterpunkt dies erzwingt.

Im zweidimensionalen Fall ist nun das Einfügen von Randknoten während der Iteration unproblematisch. Satz 6.1 stellt sicher, daß es für einen beliebigen Gitterpunkt $i$ auf jeder Randseite von $\bar{\Omega}^{(2)}$ genau einen Randpunkt $j$ mit

$$\int_0^1 \int_0^1 \langle \nabla\zeta_j, \nabla\varphi_i \rangle \; dx_2 \; dx_1 \quad \neq \quad 0$$

gibt. Dies ist für $d = 3$ nicht mehr erfüllt. Zu einem beliebigen Gitterpunkt $i$ kann es auf jeder der sechs Randflächen von $\bar{\Omega}^{(3)}$ mehrere Randpunkte $j$ geben, für die das dreidimensionale Analogon des obigen Integrals,

$$\int_0^1 \int_0^1 \int_0^1 \langle \nabla\zeta_j, \nabla\varphi_i \rangle \; dx_3 \; dx_2 \; dx_1 \quad ,$$

einen Wert ungleich Null annimmt. Da diese Einflüsse aus Effizienzgründen nicht in jedem Iterationsschritt neu berechnet, sondern in einem Summationsprozeß akkumuliert und anschließend im Baum gespeichert werden sollen, bedeutet jedes nachträgliche Einfügen eines Randpunktes im dreidimensionalen Fall einen Neustart dieses Rand-Akkumulationsprozesses. Nach jedem Iterationsschritt, in dessen Verlauf ein Knoten für einen zusätzlichen Randpunkt erzeugt worden ist, müssen folglich die entsprechenden akkumulierten Größen neu bestimmt werden. Der hierfür erforderliche Rechenaufwand fällt jedoch kaum ins Gewicht, da die Aktualisierung der jeweiligen Größen nur die Gitterpunkte auf dem Rand des Einheitswürfels betrifft.

Analog zum zweidimensionalen Fall besteht ein Iterationsschritt aus einem Durchlauf durch den Teil des Baumes, der das Innere des Einheitswürfels repräsentiert. Dieser Baumdurchlauf erfolgt wiederum in hierarchischer Ordnung. Er beginnt somit beim Mittelpunkt $x_1 = x_2 = x_3 = 0.5$. Dabei wird in jedem Knoten der Funktionswert der Näherungslösung im entsprechenden Gitterpunkt gemäß der noch zu behandelnden Iterationsvorschrift aktualisiert.

Wird ein Blatt des Baumes erreicht, so stellt sich die Frage der Verfeinerung. Im dreidimensionalen Fall unterscheiden wir jetzt drei Klassen von Knoten: Zu den zwei bereits eingeführten Klassen (Knoten, die einen einzelnen Gitterpunkt $x_1 = \bar{x}_1$ beziehungsweise eine Gitterlinie $x_2 = \bar{x}_2$ repräsentieren, vgl. die Abschnitte 5.2 und 6.1), kommen nun noch die Knoten im Baum von Gitterebenen, die eine ganze Ebene $x_3 = \bar{x}_3$ repräsentieren (siehe dazu die Abbildungen 10, 11 sowie 15). Ein Blatt aus dieser Klasse erhält bei positiver Entscheidung über die Verfeinerung zwei Söhne, die im Baum dann für die Gitterebenen $x_3 = \bar{x}_3 - h_{i_3}/2$ beziehungsweise $x_3 = \bar{x}_3 + h_{i_3}/2$ stehen. Als

Entscheidungskriterium verwendet man dabei im Falle regulärer dünner Gitter die Zugehörigkeit des entsprechenden Teilraumes $T_{i_1,i_2,i_3+1}$ zu $\tilde{S}_n^{(3)}$ beziehungsweise $\hat{S}_n^{(3)}$. Falls hingegen adaptiv verfeinert werden soll, so kommt eines der Kriterien (38) oder (39) aus Abschnitt 5.1 zum Tragen.

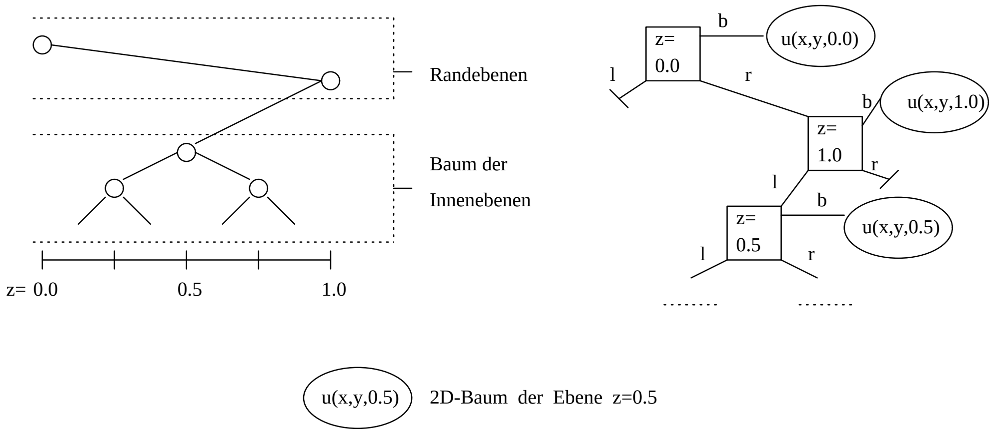



ABB. 15. *Baum zur Darstellung einer Funktion $u(x, y, z)$ auf dünnen Gittern.*

**7.2. Das Iterationsverfahren zur Lösung des Randwertproblems.** Zur Lösung von (80) wählen wir wiederum einen Ansatz gemäß der Methode der Finiten Elemente. Hierzu betrachten wir ein beliebig gegebenes, reguläres oder adaptiv erzeugtes dünnes Gitter mit $N_I$ Innen- und $N_R$ Randpunkten. $\{\varphi_j, 1 \leq j \leq N_I\} \cup \{\zeta_j, 1 \leq j \leq N_R\}$ sei die zugehörige hierarchische Basis, bestehend aus den stückweise trilinearen hierarchischen Basisfunktionen aus Abschnitt 2 im Innern ($\varphi_j$) und auf dem Rand ($\zeta_j$). Beispielsweise erhalten wir bei Betrachtung des regulären dünnen Gitters $\tilde{G}_n^{(3)}$ als Basisfunktionen $\varphi_j$ mit (6)

$$\{\varphi_j, 1 \leq j \leq N_I\} \;=\; \left\{ g_{k_1,k_2,k_3}^{(i_1,i_2,i_3)} \in \tilde{B}_n^{(3)}, 1 \leq k_l \leq 2^{i_l-1}, 1 \leq l \leq 3 \right\},$$

wobei der Index $j$ erneut die Hierarchie der Teilräume widerspiegeln soll. Analog zu (41) gelte somit für $\varphi_i \in T_{i_1,i_2,i_3}$ und $\varphi_j \in T_{j_1,j_2,j_3}$ (vgl. (2))

$$T_{i_1,i_2,i_3} \;<\; T_{j_1,j_2,j_3} \quad \Rightarrow \quad i < j. \tag{81}$$

Die $\zeta_j$ werden entsprechend als hierarchische Basisfunktionen auf dem Rand definiert.

Die Herleitung des linearen Gleichungssystems $Av = b$ verläuft nun wie im zweidimensionalen Fall. Es gilt folglich

$$\begin{aligned}
A &= (a_{ij})_{1 \leq i,j \leq N_I} \quad \in \mathbb{R}^{N_I \times N_I}, \\
a_{ij} &= \int_0^1 \int_0^1 \int_0^1 \langle \nabla\varphi_i, \nabla\varphi_j \rangle \; dx_3 \; dx_2 \; dx_1, \\
v &= (v_1, ..., v_{N_I})^T \quad \in \mathbb{R}^{N_I}, \\
b &= (b_1, ..., b_{N_I})^T \quad \in \mathbb{R}^{N_I}, \\
b_i &= -\int_0^1 \int_0^1 \int_0^1 \Big( \langle \nabla\varphi_0, \nabla\varphi_i \rangle + f \cdot \varphi_i \Big) \; dx_3 \; dx_2 \; dx_1.
\end{aligned} \tag{82}$$

Für zwei beliebige Indizes $i$ und $j$ mit $\varphi_i \in T_{i_1,i_2,i_3}$ und $\varphi_j \in T_{j_1,j_2,j_3}$ tritt dabei $a_{ij} \neq 0$ genau dann auf, wenn es eine Ebene $x_k = c,\ 1 \leq k \leq 3,\ 0 < c < 1$, gibt, die die entsprechenden Gitterpunkte — die Mittelpunkte der Träger von $\varphi_i$ und $\varphi_j$ — beide enthält, und wenn die Träger von $\varphi_i$ und $\varphi_j$ nicht disjunkt sind. Im Falle des Mittelpunktes des Einheitswürfels beispielsweise ($i = 1,\ \varphi_1 \in T_{1,1,1}$) gilt also $a_{1j} \neq 0$ für alle Gitterpunkte $j$, welche die Bedingung

$$(x_1 = 0.5)\ \vee\ (x_2 = 0.5)\ \vee\ (x_3 = 0.5)$$

erfüllen.

Das aus dem Finite-Elemente-Ansatz resultierende lineare Gleichungssystem $Av = b$ wird nun iterativ mit Hilfe des Gauß-Seidel-Verfahrens gelöst. Unter Beachtung der hierarchischen Anordnung der Teilräume $T_{i_1,i_2,i_3}$ werden die Werte $v_i$ in jedem Iterationsschritt gemäß (54) korrigiert. So wird zunächst der den Teilraum $T_{1,1,1}$ repräsentierende Mittelpunkt $(0.5, 0.5, 0.5)$ behandelt. Anschließend werden die Gitterpunkte zu $T_{i_1,i_2,i_3}$ mit $i_1+i_2+i_3 = 4$ mit neuen Näherungswerten versehen, gefolgt von den Gitterpunkten zu Teilräumen $T_{i_1,i_2,i_3}$ mit $i_1 + i_2 + i_3 = 5$, usw.. Da sich die Baumstruktur nach Abschnitt 7.1 ebenfalls an der Hierarchie der Teilräume orientiert, kann ein vollständiger Gauß-Seidel-Schritt (einmalige Korrektur jeder Variablen $v_i$) auf einem dünnen Gitter während eines Durchlaufs durch den gesamten Baum erledigt werden.

Um einen solchen Iterationsschritt mit $O(N_I)$ Additionen und Multiplikationen durchführen zu können, modifizieren wir eine Variable $v_i$ mit $\varphi_i \in T_{i_1,i_2,i_3}$ jedoch nicht mittels expliziter Berechnung der Matrixeinträge $a_{ij}$. Vielmehr wird in jedem Gitterpunkt die zugehörige Gauß-Seidel-Gleichung (54) unter Verwendung während des Baumdurchlaufs akkumulierter Größen aufgestellt und gelöst. Mit den Vereinbarungen

$$R\ :=\ h_{i_1} \cdot h_{i_2} \cdot R_i^{(1,2)}\ +\ h_{i_1} \cdot h_{i_3} \cdot R_i^{(1,3)}\ +\ h_{i_2} \cdot h_{i_3} \cdot R_i^{(2,3)}$$

und

$$r\ :=\ h_{i_1} \cdot h_{i_2} \cdot r_i^{(1,2)}\ +\ h_{i_1} \cdot h_{i_3} \cdot r_i^{(1,3)}\ +\ h_{i_2} \cdot h_{i_3} \cdot r_i^{(2,3)}$$

erhalten wir folgende Gestalt der Iterationsgleichung (54):

$$v_i\ :=\ \frac{-9 \cdot (R\ +\ r\ -\ h_{i_1} \cdot h_{i_2} \cdot h_{i_3} \cdot b_i)}{8 \cdot (h_{i_1}^2 \cdot h_{i_2}^2 + h_{i_1}^2 \cdot h_{i_3}^2 + h_{i_2}^2 \cdot h_{i_3}^2)}\ . \tag{83}$$

$R_i^{(1,2)}$ enthält dabei die Einflüsse von den Gitterpunkten $j$, die dieselbe $x_3$-Koordinate wie $i$ haben und die auf dem Rand oder außerhalb des Trägers von $\varphi_i$ liegen. Entsprechendes gilt für die Größen $R_i^{(1,3)}$ und $R_i^{(2,3)}$. Mit der Bezeichnung $Inn(Tr\ddot{a}ger(\varphi_i))$ für das Innere des Trägers von $\varphi_i$ erhalten wir somit

$$R_i^{(1,2)}\ =\ h_{i_3} \cdot \sum_{\substack{x_3(j)\, =\, x_3(i) \\ j\, \notin \mathrm{Inn}(\mathrm{Träger}(\varphi_i))}} \omega_{ij} \cdot a_{ij} \cdot v_j\ ,$$

$$
\begin{aligned}
R_i^{(1,3)} &= h_{i_2} \cdot \sum_{\substack{x_2(j)=x_2(i)\\ j\notin \mathrm{Inn}(\text{Träger}(\varphi_i))}} \omega_{ij} \cdot a_{ij} \cdot v_j \ , \\
R_i^{(2,3)} &= h_{i_1} \cdot \sum_{\substack{x_1(j)=x_1(i)\\ j\notin \mathrm{Inn}(\text{Träger}(\varphi_i))}} \omega_{ij} \cdot a_{ij} \cdot v_j \ .
\end{aligned}
$$

Man beachte, daß die Matrixeinträge $a_{ij}$ zu den Gitterpunkten $j$, die mit $i$ zum Beispiel in der $x_1$- *und* $x_2$-Koordinate übereinstimmen, in geeigneter Weise auf $R_i^{(1,3)}$ und $R_i^{(2,3)}$ verteilt werden müssen. Hierzu dienen die Gewichte $\omega_{ij}$.

Analog werden in den Variablen $r_i^{(1,2)}$, $r_i^{(1,3)}$ sowie $r_i^{(2,3)}$ die Einflüsse akkumuliert, die von Gitterpunkten $j$ im Inneren des Trägers von $\varphi_i$ mit gleicher $x_3$-, $x_2$- beziehungsweise $x_1$-Komponente wie $i$ herrühren. Die Gewichte $\rho_{ij}$ berücksichtigen wiederum den Fall, daß $i$ und $j$ in mehr als einer Koordinate übereinstimmen.

$$
\begin{aligned}
r_i^{(1,2)} &= h_{i_3} \cdot \sum_{\substack{x_3(j)=x_3(i)\\ j\in \mathrm{Inn}(\text{Träger}(\varphi_i))}} \rho_{ij} \cdot a_{ij} \cdot v_j \ , \\
r_i^{(1,3)} &= h_{i_2} \cdot \sum_{\substack{x_2(j)=x_2(i)\\ j\in \mathrm{Inn}(\text{Träger}(\varphi_i))}} \rho_{ij} \cdot a_{ij} \cdot v_j \ , \\
r_i^{(2,3)} &= h_{i_1} \cdot \sum_{\substack{x_1(j)=x_1(i)\\ j\in \mathrm{Inn}(\text{Träger}(\varphi_i))}} \rho_{ij} \cdot a_{ij} \cdot v_j \ .
\end{aligned}
$$

Wie im zweidimensionalen Fall können wir die in (83) auftretenden Summen $R_i^{(k,l)}$ und $r_i^{(k,l)}$ während des Baumdurchlaufs mit konstantem Rechenaufwand je Gitterpunkt bilden. Hierbei stützt sich die Berechnung obiger sechs Größen für einen Gitterpunkt $i$ auf die entsprechenden Größen hierarchisch tieferer Gitterpunkte. Somit folgt wiederum für unsere Gauß-Seidel-Iteration, daß — ungeachtet der Gestalt der hierarchischen Dünn-Gitter-Matrix $A$ — die Anzahl der arithmetischen Operationen in einem Iterationsschritt proportional zur Anzahl $N_I$ der Unbekannten ist.

Was die Diskretisierung der rechten Seite $f$ der Poisson-Gleichung angeht, so können die Ergebnisse des zweidimensionalen Falles aus Abschnitt 6.2.3 direkt übertragen werden. Nach (82) benötigen wir zur Ermittlung der Komponenten $b_i$ des Vektors $b$ in unserem Gleichungssystem die $N_I$ Integrale $\mathcal{I}_i$ $(1 \le i \le N_I)$,

$$
\mathcal{I}_i \ := \ \int_0^1 \int_0^1 \int_0^1 f \cdot \varphi_i \ dx_3 \ dx_2 \ dx_1 \ .
$$

Als Näherung hierfür berechnen wir mit Hilfe eines geeignet gewählten Dünn-Gitter-Interpolanten $\tilde{f}_I$ von $f$ in einer Vorlaufrechnung vor Beginn der eigentlichen Iteration wiederum die Integrale $\tilde{\mathcal{I}}_i$,

$$
\tilde{\mathcal{I}}_i \ := \ \int_0^1 \int_0^1 \int_0^1 \tilde{f}_I \cdot \varphi_i \ dx_3 \ dx_2 \ dx_1 \ ,
$$

und entkoppeln so die Auflösung beziehungsweise den Adaptionsprozeß von $f$ und der Näherungslösung $u$. Die anschließend startende Iteration benutzt dann einen geeigneten

Ausschnitt des für $\tilde{f}_I$ erzeugten Baumes zur Darstellung von $u$. Zur Vermeidung eines Neustarts der Vorlaufrechnung aufgrund während der Iteration neu eingefügter Knoten verwenden wir als Kriterium für die Wahl der Gitterparameter $n_f$ und $n_u$ (reguläre dünne Gitter) beziehungsweise $\epsilon_f$ und $\epsilon_u$ (adaptiv erzeugte Gitter) erneut (60):

$$n_f > n_u \quad \text{beziehungsweise} \quad \epsilon_f < \epsilon_u.$$

Zum Abschluß dieses Abschnitts über dreidimensionale Poisson-Probleme auf dünnen Gittern wollen wir noch kurz auf Konvergenzfragen eingehen. Zunächst halten wir fest, daß die im Beweis von Satz 6.2 verwendete Minimaleigenschaft der Finite-Elemente-Lösung auch im dreidimensionalen Fall gilt: Von allen Funktionen aus einem Dünn-Gitter-Raum weist die von uns berechnete Näherungslösung bezüglich der Energienorm den kleinsten Abstand zur exakten Lösung von (80) auf. Zusammen mit den Resultaten über den in der Energienorm gemessenen Interpolationsfehler aus Satz 4.4 und Satz 4.5 folgt daraus sofort

Satz 7.1. *Sei $u^{(3)} \in X_0^{(3)}$ die exakte Lösung des Problems (80) mit $g \equiv 0$. Seien ferner $\tilde{u}_n^{(3)} \in \tilde{S}_n^{(3)}$ und $\hat{u}_n^{(3)} \in \hat{S}_n^{(3)}$ die von unserem Verfahren auf den dünnen Gittern $\tilde{G}_n^{(3)}$ beziehungsweise $\hat{G}_n^{(3)}$ produzierten Näherungslösungen. Dann gilt:*

$$\begin{aligned} \|u^{(3)} - \tilde{u}_n^{(3)}\|_E &\leq h_n \cdot \frac{\sqrt{3} \cdot |u^{(3)}|_\infty}{32}, \\ \|u^{(3)} - \hat{u}_n^{(3)}\|_E &\leq h_n \cdot \frac{5 \cdot \sqrt{3} \cdot |u^{(3)}|_\infty}{64}. \end{aligned} \tag{84}$$

Wir erhalten also auch im dreidimensionalen Fall eine Dünn-Gitter-Lösung, die bezüglich der Energienorm mit einem Fehler der Ordnung $O(h_n)$ behaftet ist, was dem Verhalten bei Verwendung eines vollen Gitters entspricht. Die Anzahl der Gitterpunkte ist dagegen nur von der Ordnung $O(h_n^{-1} \cdot (\mathrm{ld}(h_n^{-1}))^2)$ im Falle des dünnen Gitters $\tilde{G}_n^{(3)}$ beziehungsweise $O(h_n^{-1})$ im Falle des dünnen Gitters $\hat{G}_n^{(3)}$ und damit signifikant kleiner als beim vollen Gitter $G_n^{(3)}$.

Aussagen über die Konvergenzgeschwindigkeit des dreidimensionalen Verfahrens, d.h. insbesondere über den Spektralradius der Gauß-Seidel-Iterationsmatrix $M$, können analog zur Vorgehensweise in Abschnitt 6.2.4 hergeleitet werden. Derartige Untersuchungen wurden jedoch bislang nicht durchgeführt. Da die in [16] und [17] für den zweidimensionalen Fall beschriebene Technik den Weg hin zu einer Mehrgittererweiterung auch für dreidimensionale Probleme weist, in der die Gauß-Seidel-Iteration nur noch als Glätter auftritt, verliert eine detaillierte Untersuchung der Konvergenzgeschwindigkeit des reinen Gauß-Seidel-Verfahrens ohnehin an Bedeutung.

Mit diesen Bemerkungen wollen wir die Diskussion unseres Verfahrens zur Lösung dreidimensionaler Poisson-Probleme beenden und uns nun noch einigen algorithmischen Varianten der beschriebenen Dünn-Gitter-Technik zuwenden.

**8. Zwei Varianten des Poisson-Lösers im zweidimensionalen Fall.** Drei Aspekte standen bei der bisherigen Analyse dünner Gitter und des darauf definierten Verfahrens zur Lösung von Poisson-Problemen im Vordergrund: die Anzahl der Unbekannten im Gitter und damit der Speicherplatzbedarf, die Genauigkeit des Interpolanten beziehungsweise der erzielten Näherungslösung sowie die Konvergenzgeschwindigkeit des Iterationsverfahrens. Wie die Resultate aus Abschnitt 6.2.4 zeigen, ist dabei eine Verbesserung unseres Verfahrens hinsichtlich der Konvergenzgeschwindigkeit am dringlichsten. Dem wurde mit der bereits erwähnten Möglichkeit einer Mehrgittererweiterung Rechnung getragen. Im folgenden stellen wir nun eine Technik zur weiteren Speicherplatzersparnis vor und zeigen außerdem Wege zur Erzielung genauerer Interpolanten beziehungsweise genauerer Näherungslösungen für unser Modellproblem auf, wobei wir uns jeweils auf den zweidimensionalen Fall beziehen.

**8.1. Weitere Reduzierung des Speicheraufwands.** Die Anzahl der im regulären dünnen Gitter $\tilde{G}_n^{(2)}$ auftretenden Gitterpunkte ist nach den Ergebnissen aus Abschnitt 3 von der Ordnung $O\left(h_n^{-1} \cdot \mathrm{ld}(h_n^{-1})\right)$. Für das Gitter $\hat{G}_n^{(2)}$ ergaben die Betrachtungen in Abschnitt 4 sogar die Ordnung $O(h_n^{-1})$. Korollar 3.2 und Korollar 4.3 liefern jedoch nicht nur die Größenordnung betreffende Resultate, sondern geben die Anzahl der Gitterpunkte sogar exakt an. Damit ist sichergestellt, daß die Ordnungsaussagen nicht nur asymptotische Bedeutung haben. Auch für kleinere $n$ sind die Vorteile dünner Gitter bereits spürbar.

Bei derartigen Betrachtungen darf aber nicht übersehen werden, daß die Anzahl der Gitterpunkte nicht direkt als Maß für den Speicheraufwand herangezogen werden sollte. Vielmehr müssen einige andere Aspekte mit berücksichtigt werden:

- Muß der diskretisierte Differetialgleichungsoperator — die Matrix $A$ — ebenfalls gespeichert werden?
- Falls ja — wieviele von Null verschiedene Matrixeinträge stehen im Mittel in jeder Zeile von $A$?
- Wieviele Variablen werden pro Gitterpunkt benötigt?
- Erfordert das Verfahren organisatorischen Aufwand (z.B. Zeiger)?

Daß diese Fragestellungen nicht unerheblich sind, wird anhand zweier Beispiele für Verfahren zur Lösung partieller Differentialgleichungen schnell klar. Lösen wir etwa ein aus dem Finite-Elemente-Ansatz bezüglich Stützpunktbasen, rechteckiger Elemente und regulärer voller Gitter resultierendes lineares Gleichungssystem mit Hilfe der Gauß-Seidel-Iteration, so benötigen wir dazu lediglich zwei Gleitpunktzahlen pro Unbekannter $i$. Denn die Matrix $A$ hat eine sehr regelmäßige Gestalt und braucht deshalb nicht explizit abgespeichert zu werden, so daß wir je eine Gleitpunktzahl für die Komponente $b_i$ der rechten Seite und die Komponente $v_i$ der Näherungslösung benötigen. Da in diesem Falle Vektoren als Datenstruktur verwendet werden, fällt auch kein organisatorischer Aufwand an, so daß man bei Rechnung mit doppelter Genauigkeit mit 16 Byte pro Gitterpunkt auskommt.

Betrachten wir dagegen auf Triangulationen arbeitende adaptive Techniken, so fällt hier oftmals ein Vielfaches an Speicheraufwand an. Zum Teil werden — vor allem bei komplizierteren Differentialgleichungen — die von Null verschiedenen Matrixeinträge explizit gespeichert, die Koordinaten und Nachbarschaftsbeziehungen der Gitterpunkte müssen erfaßt werden, und überhaupt gestaltet sich die Verwaltung des Dreiecksnetzes recht aufwendig. Ein Gesamtbedarf von mehreren hundert Byte pro Gitterpunkt ist hier keine Seltenheit.

Bei unserem Verfahren bestimmen wir die Einträge der Steifigkeitsmatrix nicht explizit. Insbesondere entsteht somit auch kein Speicherplatzbedarf für die Matrixelemente. In seiner bislang beschriebenen Form verwendet der Algorithmus sechs Gleitpunkt-Variablen für jeden Gitterpunkt: je eine Variable für den absoluten Wert der Näherungslösung sowie den hierarchischen Überschuß, zwei für die Größen $r_i^{(1)}$ und $r_i^{(2)}$ (vgl. (55)), eine für die beiden Größen $R_i^{(1)}$ und $R_i^{(2)}$ gemäß (55) sowie die Einflüsse der Dirichlet-Randfunktion und schließlich eine Variable für die diskretisierte rechte Seite der Poisson-Gleichung. Zusammen mit einer ganzzahligen Variablen für die Tiefe des Knotens im Baum und den Zeigern für die Realisierung unseres Binärbaumes ergibt sich bei Rechnung in doppelter Genauigkeit ein Speicheraufwand von 72 Byte je Gitterpunkt.

Dieser Aufwand kann jetzt aber noch weiter reduziert werden. Da sowohl die Baumstruktur als auch die Durchlaufreihenfolge bei der Gauß-Seidel-Iteration auf die hierarchische Ordnung der Teilräume ausgerichtet sind, kann man den Algorithmus so modifizieren, daß lediglich zwei Gleitpunkt-Variablen in jedem Gitterpunkt benötigt werden. Eine dieser Variablen enthält wie bisher die diskretisierte rechte Seite. In der anderen werden nun abwechselnd absoluter Wert (Funktionswert), hierarchischer Überschuß, akkumulierte Einflüsse aus dem Inneren sowie akkumulierte Einflüsse vom Rand oder von außerhalb des Trägers der jeweiligen Basisfunktion gespeichert. Unmittelbar vor der Modifikation einer Variablen $i$ mit $\varphi_i \in T_{i_1,i_2}$ enthalten dabei $v_i$ sowie alle $v_j$, deren jeweils zugeordnete Basisfunktion $\varphi_j$ in einem hierarchisch niedrigeren Teilraum $T_{j_1,j_2}$ liegt ($T_{j_1,j_2} > T_{i_1,i_2}$), die akkumulierten Einflüsse aus dem Inneren des Trägers von $\varphi_i$ beziehungsweise $\varphi_j$. In denjenigen Variablen $v_j$, für die $\varphi_j$ in einem hierarchisch höheren Teilraum $T_{j_1,j_2}$ liegt ($T_{j_1,j_2} < T_{i_1,i_2}$), stehen zu diesem Zeitpunkt die akkumulierten Einflüsse vom Rand oder von außerhalb des Trägers der jeweiligen Basisfunktion. Direkt im Anschluß an die Modifikation der Variablen $i$ enthält $v_i$ nun einen verbesserten Wert für den hierarchischen Überschuß im Gitterpunkt $i$, dividiert durch den Skalierungsfaktor $h_{i_1} \cdot h_{i_2}$.

Wir erhalten schließlich folgende Gestalt der Iterationsgleichung (55), in der anstelle der Größen $r_i^{(1)}$, $r_i^{(2)}$, $R_i^{(1)}$, $R_i^{(2)}$ und $b_i$ auch auf der rechten Seite nur noch $b_i$ sowie Variablen $v_j$ auftreten. Man beachte jedoch, daß die einzelnen $v_j$ völlig verschiedenartige Information tragen.

$$(85) \qquad v_i := \frac{-3 \cdot \Big( (v_{süd} + v_{nord}) \cdot h_{i_2} + (v_{west} + v_{ost}) \cdot h_{i_1} + v_i - b_i \Big)}{4 \cdot (h_{i_1}^2 + h_{i_2}^2)}.$$

Dabei sind die vier Gitterpunkte *süd*, *nord*, *west* und *ost* die Schnittpunkte der achsenparallelen Geraden durch $i$ mit dem Rand des Trägers von $\varphi_i$.

Da die Werte der Variablen $v_j$ jetzt an verschiedenen Stellen im Algorithmus unterschiedlich interpretiert werden, muß zwischen den einzelnen Bedeutungen hin- und hergeschaltet werden. Die dazu erforderlichen Additionen und Multiplikationen erhöhen die Gesamtzahl der arithmetischen Operationen pro Gitterpunkt und Iterationsschritt jedoch lediglich um etwa zwanzig Prozent, so daß die Rechenzeit zur Durchführung eines Gauß-Seidel-Schrittes nur geringfügig zunimmt.

Bei Rechnung mit doppelter Genauigkeit geht der Gesamtbedarf an Speicherplatz pro Gitterpunkt durch die beschriebene Modifikation unseres Verfahrens von 72 auf 40 Byte zurück. Der jetzt noch im Vergleich zu den anfangs erwähnten 16 Byte (reguläres volles Gitter, Finite-Elemente-Diskretisierung, Stützpunktbasis, Gauß-Seidel) für einen Gitterpunkt zusätzlich benötigte Speicherplatz ist dabei ausschließlich auf die zur Realisierung der Binärbaum-Datenstruktur verwendeten Zeiger zurückzuführen.

**8.2. Erhöhung der Ordnung.** In Abschnitt 3 haben wir die Approximationseigenschaften des Dünn-Gitter-Raumes $\tilde{S}_n^{(2)}$ untersucht und dabei für den Interpolationsfehler sowohl bezüglich der $L_2$-Norm als auch bezüglich der Maximumsnorm Schranken der Größenordnung $O(h_n^2 \cdot \mathrm{ld}(h_n^{-1}))$ angegeben. Die mangelnde Übertragbarkeit dieser Resultate auf den Fehler bei der Lösung unseres Modellproblems (40) führte dann in Abschnitt 4 zu den Energienormbetrachtungen. Numerische Experimente zeigen jedoch, daß auch die $L_2$- beziehungsweise Maximumsnorm des Diskretisierungsfehlers von der Größenordnung $O(h_n^2 \cdot \mathrm{ld}(h_n^{-1}))$ ist. Wir können deshalb unseren Dünn-Gitter-Algorithmus auch als Verfahren höherer Ordnung interpretieren. Sieht man einmal von den logarithmischen Faktoren ab, so produziert er mit $O(h_n^{-1})$ Gitterpunkten Näherungslösungen, die mit einem Fehler der Ordnung $O(h_n^2)$ behaftet sind, wozu auf vollen Gittern bereits ein Verfahren der Ordnung $O(h_n^4)$ erforderlich ist. Dieser Effekt nimmt mit wachsender Dimensionalität noch zu: Im dreidimensionalen Fall leistet unser Verfahren bereits soviel wie ein Verfahren der Ordnung $O(h_n^6)$. Dabei werden jedoch keine ausgedehnten Sterne verwendet, die bei gängigen Techniken zur Ordnungserhöhung oftmals zu Schwierigkeiten am Rand führen.

Das regelmäßige Verhalten des Diskretisierungsfehlers legt es nahe, durch extrapolationsähnliche Techniken oder — wie bei der $p$-Version der Methode der Finiten Elemente — durch die Verwendung von Polynomen höheren Grades als Basisfunktionen die Fehlerordnung weiter zu erhöhen. Wir stellen hier zwei solche Ansätze vor.

Eine erste Idee knüpft an Lemma 2.4 an (siehe (12)). Dort wurde für die hierarchischen Überschüsse $c_{k_1,k_2}^{(i_1,i_2)}$ die Abschätzung

$$|c_{k_1,k_2}^{(i_1,i_2)}| \;\leq\; \frac{|u^{(2)}|_\infty}{4} \cdot h_{i_1}^2 \cdot h_{i_2}^2 \;=\; \frac{|u^{(2)}|_\infty}{4} \cdot 4^{-i_1-i_2} \tag{86}$$

gezeigt. Der Dünn-Gitter-Interpolant $\tilde{u}_{n,I}^{(2)} \in \tilde{S}_n^{(2)}$ von $u^{(2)} \in X_0^{(2)}$ berücksichtigt solche Überschüsse jedoch nur bis zu Teilräumen $T_{i_1,i_2}$ der Tiefe $i_1 + i_2 = n+1$. Bezeichne $\mathcal{B}_n$ den $\tilde{S}_n^{(2)}$ entsprechenden Baum zur Darstellung von $\tilde{u}_{n,I}^{(2)}$. Ausgehend vom Baum $\mathcal{B}_n$ konstruieren wir eine Folge von Bäumen $\mathcal{B}_{n+k}$, $k = 1, 2, \ldots$ . Jeder Baum $\mathcal{B}_{n+k}$ — geeignet zur Darstellung von Funktionen aus $\tilde{S}_{n+k}^{(2)}$ — soll dabei wiederum unseren Dünn-Gitter-Interpolanten $\tilde{u}_{n,I}^{(2)}$ repräsentieren. In seinem $\tilde{S}_n^{(2)}$ entsprechenden Teilbaum stimmt er also mit $\mathcal{B}_n$ überein. In allen zusätzlichen Knoten zu Gitterpunkten $i$, deren zugeordnete

Basisfunktion in einem Teilraum $T_{i_1,i_2}$ mit $i_1 + i_2 > n + 1$ liegt, wird der hierarchische Überschuß auf Null gesetzt. Ausgehend von dieser Gedankenkonstruktion wollen wir zu einer besseren Approximation für $u^{(2)}$ als $\tilde{u}^{(2)}_{n,I}$ gelangen, ohne den Approximationsraum $\tilde{S}^{(2)}_n$ explizit verlassen zu müssen. Zu diesem Zwecke tragen wir nun bei den Söhnen eines Blattes im Baum $\mathcal{B}_n$, die ja gerade die Blätter im Baum $\mathcal{B}_{n+1}$ darstellen und Teilräumen $T_{i_1,i_2}$ mit $i_1 + i_2 = n + 2$ entsprechen, als hierarchischen Überschuß nicht Null, sondern ein Viertel des Wertes aus diesem Blatt ein. Dementsprechend tragen wir bei den Enkeln des betreffenden Blattes ein Sechzehntel des Wertes ein und fahren rekursiv fort. Dieses Abfallen der hierarchischen Überschüsse simuliert das durch (86) beschriebene Verhalten und sollte deshalb zu einem besseren Interpolanten für $u^{(2)}$ beziehungsweise — im Algorithmus zur Lösung des Modellproblems (40) — zu einer besseren Näherungslösung führen.

Zur praktischen Realisierung der beschriebenen Idee setzen wir in einem Gitterpunkt $i$ mit zugeordneter Basisfunktion $\varphi_i \in T_{i_1,i_2}$ $(i_1 + i_2 = n + 1)$ — also einem Blatt des Baumes $\mathcal{B}_n$ — den hierarchischen Überschuß $v_i$ gemäß (86) auf das Achsenkreuz des Trägers von $\varphi_i$ fort, summieren die daraus resultierenden Einflüsse in den Variablen $r_i^{(1)}$ sowie $r_i^{(2)}$ auf und lösen die Iterationsgleichung (55) im Punkt $i$ mit diesen Werten. Abbildung 16 veranschaulicht die Ausbreitung des hierarchischen Überschusses im Punkt $i$ auf den gesamten Träger von $\varphi_i$.

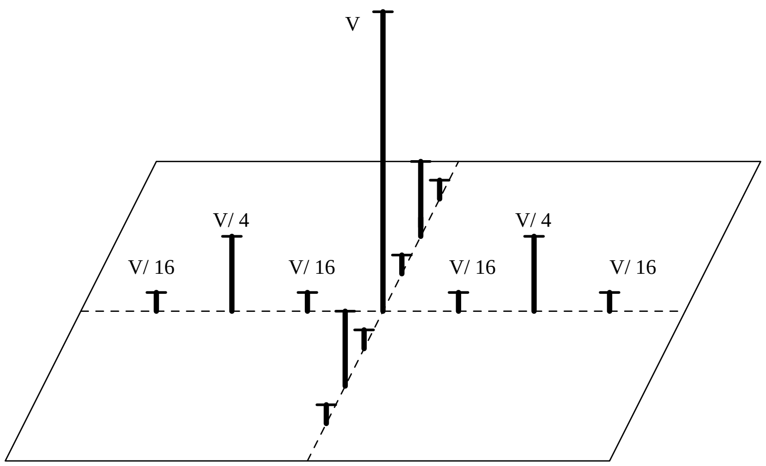



ABB. 16. *Ausbreitung des hierarchischen Überschusses $V = v_i$ im Punkt $i$ auf den Träger von $\varphi_i$.*

Abbildung 17 zeigt, daß diese Vorgehensweise auf jeder Achse des Trägers von $\varphi_i$ dem Übergang von stückweise linearer zu quadratischer Interpolation entspricht.

Bislang waren die Größen $r_i^{(1)}$ sowie $r_i^{(2)}$, die ja nach (55) die Einflüsse aus dem Inneren des Trägers von $\varphi_i$ enthalten, für einen Gitterpunkt $i$ mit $\varphi_i \in T_{i_1,i_2}$ und $i_1 + i_2 = n + 1$ gleich Null, da es im Inneren des Trägers von $\varphi_i$ keine Gitterpunkte gab. Mit unserer Modifikation simulieren wir jetzt die Existenz solcher Gitterpunkte. Insbesondere müssen wir also die Einflüsse $(\frac{v_i}{4}, \frac{v_i}{16}, \ldots)$ akkumulieren und in $r_i^{(1)}$ beziehungsweise $r_i^{(2)}$ bereitstellen. Wir erhalten nach kurzer Rechnung

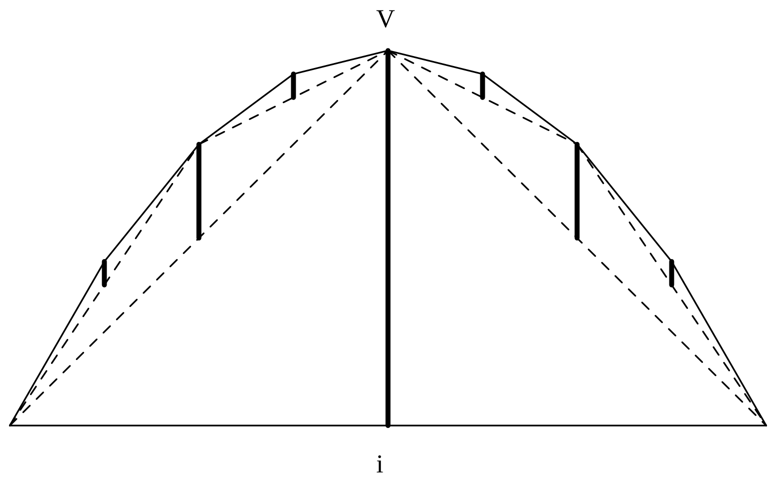



ABB. 17. *Übergang von stückweise linearer zu quadratischer Interpolation (Hutfunktion → Parabel).*

$$r_i^{(1)} \;=\; \frac{v_i \cdot h_{i_1}}{3}$$

sowie

$$r_i^{(2)} \;=\; \frac{v_i \cdot h_{i_2}}{3}.$$

Da sich die Berechnung der Größen $r_j^{(1)}$ sowie $r_j^{(2)}$ für Gitterpunkte $j$, die hierarchisch höheren Teilräumen entsprechen und für die $a_{ij} \neq 0$ gilt, auf die Größen $r_i^{(1)}$ und $r_i^{(2)}$ stützt (siehe Abschnitt 6.2.2), ist mit dieser Vorbesetzung ebenfalls sichergestellt, daß die neuen Einflüsse auch in die Iterationsgleichungen zu solchen Punkten $j$ korrekt einfließen. Somit genügt eine geringfügige Modifikation — die Vorbesetzung von $r_i^{(1)}$ und $r_i^{(2)}$ in den Blättern des Baumes $\mathcal{B}_n$ mit den obigen Werten — zur Implementierung unserer Idee.

Zu einer anderen Interpretation der beschriebenen Vorgehensweise gelangt man, wenn man die Iterationsgleichung (55) auf der niedrigsten Hierarchieebene — also in Punkten $i$ mit $\varphi_i \in T_{i_1,i_2}$ und $i_1 + i_2 = n + 1$ — betrachtet. Diese entspricht zunächst natürlich der Anwendung des gewöhnlichen Finite-Elemente-Sternes

$$\frac{1}{6 \cdot h_{i_1}^2} \cdot \begin{bmatrix} 1 & -2 & 1 \\ 4 & -8 & 4 \\ 1 & -2 & 1 \end{bmatrix} + \frac{1}{6 \cdot h_{i_2}^2} \cdot \begin{bmatrix} 1 & 4 & 1 \\ -2 & -8 & -2 \\ 1 & 4 & 1 \end{bmatrix}$$

auf unsere Näherungslösung. Die Modifikation dieses Abschnitts führt demgegenüber auf der feinsten Ebene zu dem Stern

$$\frac{1}{12 \cdot h_{i_1}^2} \cdot \begin{bmatrix} 1 & -2 & 1 \\ 10 & -20 & 10 \\ 1 & -2 & 1 \end{bmatrix} + \frac{1}{12 \cdot h_{i_2}^2} \cdot \begin{bmatrix} 1 & 10 & 1 \\ -2 & -20 & -2 \\ 1 & 10 & 1 \end{bmatrix},$$

der für quadratische Elemente ($h_{i_1} = h_{i_2}$) gerade dem Collatz-Stern entspricht. Wir können die Ausbreitung der hierarchischen Überschüsse von den Blättern des Baumes aus also auch als andere Diskretisierung in den Blättern deuten. Anstelle der sonst verwendeten Finite-Elemente-Diskretisierung benutzen wir auf der feinsten Ebene eine Verallgemeinerung des Collatz-Sternes auf rechteckige Elemente.

Numerische Experimente zeigen, daß auf diese Weise ein Fehler bezüglich der $L_2$- beziehungsweise Maximumsnorm der Ordnung $O(h_n^2)$ erzielt werden kann (siehe Abschnitt 9.4, Abbildung 37). Der als Dünn-Gitter-Effekt aufgetretene logarithmische Faktor läßt sich also beseitigen.

Ein zweiter Ansatz liefert nun Verfahren, welche in numerischen Experimenten zu Näherungslösungen mit einem Fehler der Ordnung $O(h_n^3 \cdot \mathrm{ld}(h_n^{-1}))$ beziehungsweise $O(h_n^4 \cdot \mathrm{ld}(h_n^{-1}))$ führen (siehe Abschnitt 9.4). Die zugrundeliegende Idee dabei ist, aus den nach wie vor verwendeten stückweise bilinearen Basisfunktionen die zur Rechnung mit quadratischen beziehungsweise kubischen Polynomen als Basisfunktionen erforderlichen Informationen zu gewinnen, um so den Einsatz solcher Basisfunktionen gewissermaßen zu simulieren. Das soeben beschriebene Prinzip der Ausbreitung der hierarchischen Überschüsse von den Blättern des Baumes aus gelangt dabei ebenfalls zur Anwendung.

Zur Veranschaulichung betrachten wir den eindimensionalen Fall. Sei $i$ der Gitterpunkt mit der $x_1$-Koordinate $a + 4h$. Ohne Einschränkung gelte dabei $a = 0$. Die zugeordnete Basisfunktion $\varphi_i$ habe als Träger das Intervall $[0, 8h]$. In diesem Intervall sollen neben $i$ selbst zumindest die Gitterpunkte zu $x_1 \in \{h, 2h, 3h\}$ existieren. Wir bestimmen den hierarchischen Überschuß in den Punkten $x_1 = h$ und $x_1 = 3h$ bei stückweise linearer, quadratischer und kubischer Interpolation. Abbildung 18 zeigt die Hierarchiebeziehungen zwischen den Gitterpunkten und verdeutlicht die Zusammenhänge bei der Interpolation.

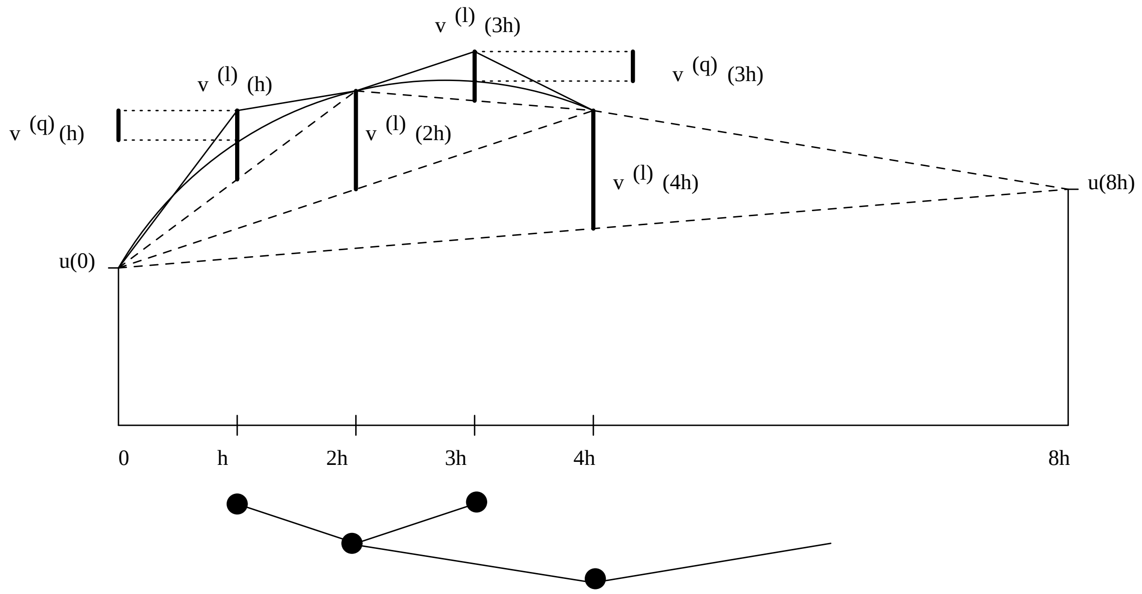



Abb. 18. *Stückweise lineare, quadratische und kubische Interpolation.*

Im linearen Fall erhalten wir als hierarchischen Überschuß $v^{(l)}(h)$ beziehungsweise $v^{(l)}(3h)$ nach Abschnitt 2

$$\begin{aligned} v^{(l)}(h) &:= u(h) - l(h) &&= u(h) - \frac{1}{2} \cdot \Big(u(0) + u(2h)\Big), \\ v^{(l)}(3h) &:= u(3h) - l(3h) &&= u(3h) - \frac{1}{2} \cdot \Big(u(2h) + u(4h)\Big). \end{aligned}$$

Hierbei bezeichnet $l(x)$ den Streckenzug, der $u$ an den drei Punkten 0, $2h$ sowie $4h$ stückweise linear interpoliert.

Im quadratischen Fall stellt der hierarchische Überschuß $v^{(q)}(h)$ gerade die Differenz $u(h)-q(h)$ dar, wobei $q(x)$ das quadratische Polynom bezeichnet, welches $u$ an den drei Punkten 0, $2h$ sowie $4h$ interpoliert. Entsprechendes gilt für $v^{(q)}(3h)$. Es ergibt sich

$$\begin{aligned} v^{(q)}(h) &:= u(h) - q(h) \\ &= v^{(l)}(h) - \frac{1}{4} \cdot v^{(l)}(2h), \\ v^{(q)}(3h) &:= u(3h) - q(3h) \\ &= v^{(l)}(3h) - \frac{1}{4} \cdot v^{(l)}(2h). \end{aligned}$$

Folglich können wir den *quadratischen* hierarchischen Überschuß $v^{(q)}$ in den Punkten $h$ und $3h$ allein durch die *linearen* hierarchischen Überschüsse $v^{(l)}$ im betreffenden Gitterpunkt selbst sowie in seinem Vater ausdrücken. Die explizite Verwendung oder Angabe quadratischer Polynome ist dazu nicht erforderlich.

Im kubischen Fall schließlich ist der hierarchische Überschuß $v^{(k)}(h)$ die Differenz $u(h)-k(h)$, wobei $k(x)$ das kubische Polynom bezeichnet, welches $u$ an den vier Punkten 0, $2h$, $4h$ sowie $8h$ interpoliert. Entsprechendes gilt wiederum für $v^{(k)}(3h)$. Eine kurze Rechnung ergibt

$$\begin{aligned} v^{(k)}(h) &:= u(h) - k(h) \\ &= v^{(q)}(h) - \frac{1}{8} \cdot v^{(q)}(2h) \\ &= v^{(l)}(h) - \frac{3}{8} \cdot v^{(l)}(2h) + \frac{1}{32} \cdot v^{(l)}(4h), \\ v^{(k)}(3h) &:= u(3h) - k(3h) \\ &= v^{(q)}(3h) + \frac{1}{8} \cdot v^{(q)}(2h) \\ &= v^{(l)}(3h) - \frac{1}{8} \cdot v^{(l)}(2h) - \frac{1}{32} \cdot v^{(l)}(4h). \end{aligned}$$

Wir können also *kubische* hierarchische Überschüsse $v^{(k)}$ ohne Angabe kubischer Interpolationspolynome berechnen. Hierzu sind nur die *linearen* hierarchischen Überschüsse $v^{(l)}$ im betreffenden Gitterpunkt selbst, in seinem Vater und in seinem Großvater erforderlich.

Diese Erkenntnis benutzen wir nun zur Erhöhung der Ordnung in unserem zweidimensionalen Poisson-Algorithmus. Da dabei nach wie vor die hierarchische Basis aus stückweise bilinearen Basisfunktionen (6) verwendet wird, bleibt insbesondere die Gestalt der Iterationsgleichung (55) erhalten. Bei der Berechnung der Größen $r_i^{(1)}$ sowie $r_i^{(2)}$ gemäß Abschnitt 6.2.2 setzen wir jetzt allerdings für $v_j$ anstelle des bisher herangezogenen *linearen* zweidimensionalen hierarchischen Überschusses $v_j^{(l)}$ den *quadratischen* beziehungsweise *kubischen* zweidimensionalen hierarchischen Überschuß $v_j^{(q)}$ beziehungsweise $v_j^{(k)}$ ein. In Analogie zu obigen Überlegungen im eindimensionalen Fall nutzen wir auch hier aus, daß sich $v_j^{(q)}$ und $v_j^{(k)}$ direkt aus den linearen Überschüssen im Punkt $j$ selbst sowie in einem beziehungsweise zwei hierarchisch höheren Punkten berechnen lassen. Insbesondere ist die Kenntnis der jeweiligen Polynome nicht erforderlich. Während des gesamten Akkumulationsprozesses wird also die Verwendung von Polynomen höheren Grades als Basisfunktionen simuliert. Bevor allerdings die Iterationsgleichung (55) für den Punkt $i$ gelöst wird, müssen die aus dem modifizierten Akkumulationsprozeß resultierenden Größen $r^{(1)}$ und $r^{(2)}$ wieder auf lineare zweidimensionale hierarchische Überschüsse „umgeschaltet“ werden, da (55) von stückweise bilinearen Basisfunktionen ausgeht.

Somit läßt sich mit einigen wenigen Modifikationen an unserem Algorithmus die Genauigkeit der von ihm produzierten Näherungslösung erhöhen. An dieser Stelle sei allerdings angemerkt, daß sich diese Erhöhung der Ordnung nur dann auswirkt, wenn die in Abschnitt 6.2.3 beschriebene Integration der rechten Seite $f$ der Poisson-Gleichung ebenfalls mit erhöhter Genauigkeit durchgeführt wird.

Die hier aufgezeigten Wege zur Erhöhung der Ordnung unseres Verfahrens zur Lösung von Poisson-Problemen auf dünnen Gittern beruhen auf heuristischen Ansätzen. Mit den in diesem Abschnitt vorgestellten Techniken soll in erster Linie auf die interessanten Möglichkeiten, die dünne Gitter im Zusammenhang mit Verfahren höherer Ordnung bieten, sowie auf deren einfache Realisierung hingewiesen werden.

**9. Numerische Ergebnisse.** Wir wollen nun noch die Wirkungsweise des in dieser Arbeit vorgestellten Verfahrens zur Lösung von Poisson-Problemen anhand einiger numerischer Beispiele demonstrieren. Dabei befassen wir uns zunächst mit dem zweidimensionalen Fall und $L_2$-normorientierten dünnen Gittern. Neben Beispielen von Laplace- beziehungsweise Poisson-Problemen mit glatter Lösung betrachten wir auch ein Problem mit einer Randsingularität. Anschließend wenden wir uns dreidimensionalen Aufgabenstellungen zu. Als feinstes reguläres dünnes Gitter verwenden wir hierbei das Gitter $\tilde{G}_{11}^{(3)}$ mit 262145 Gitterpunkten. Man beachte, daß das entsprechende volle Gitter $G_{11}^{(3)}$ demgegenüber bereits mehr als $8.6 \cdot 10^9$ Gitterpunkte enthält. Zur Veranschaulichung der Resultate aus Abschnitt 4 geben wir Beispiele zu zwei- und dreidimensionalen energieorientierten dünnen Gittern an. Den Abschluß bilden Experimente zu den ordnungserhöhenden Techniken aus dem vorigen Abschnitt.

Im folgenden bezeichne $u^{(d)}$ die exakte Lösung des jeweils betrachteten Modellproblems. $\tilde{u}_n^{(d)}$, $\hat{u}_n^{(d)}$, $\tilde{u}_\epsilon^{(d)}$ sowie $\hat{u}_\epsilon^{(d)}$ seien die Näherungslösungen, die unser Verfahren auf den regulären dünnen Gittern $\tilde{G}_n^{(d)}$ oder $\hat{G}_n^{(d)}$ beziehungsweise auf den mittels der Kriterien (38) oder (39) bei gegebenem Adaptionsparameter $\epsilon$ adaptiv erzeugten dünnen Gittern liefert. Dann definieren wir als Fehler die Differenzen

$$\begin{aligned}
\tilde{e}_n^{(d)} &:= \tilde{u}_n^{(d)} - \tilde{u}_{n,I}^{(d)} \;, \\
\hat{e}_n^{(d)} &:= \hat{u}_n^{(d)} - \hat{u}_{n,I}^{(d)} \;, \\
\tilde{e}_\epsilon^{(d)} &:= \tilde{u}_\epsilon^{(d)} - \tilde{u}_{\epsilon,I}^{(d)} \;, \\
\hat{e}_\epsilon^{(d)} &:= \hat{u}_\epsilon^{(d)} - \hat{u}_{\epsilon,I}^{(d)} \;,
\end{aligned}$$

wobei die Subtrahenden gerade die Interpolanten der exakten Lösung $u^{(d)}$ auf die jeweiligen dünnen Gitter darstellen.

Zur Beurteilung der Qualität der erhaltenen Näherungslösungen betrachten wir die folgenden diskreten Normen

$$\begin{aligned}
\|\tilde{e}_n^{(d)}\|_\infty &:= \max_{i \in \tilde{G}_n^{(d)} \cap \Omega^{(d)}} |\tilde{e}_n^{(d)}(i)| \;, \\
\|\tilde{e}_n^{(d)}\|_2 &:= \left( \frac{1}{\tilde{N}_n^{(d)}} \cdot \sum_{i \in \tilde{G}_n^{(d)} \cap \Omega^{(d)}} |\tilde{e}_n^{(d)}(i)|^2 \right)^{\frac{1}{2}} \;.
\end{aligned}$$

Für $\hat{e}_n^{(d)}$, $\tilde{e}_\epsilon^{(d)}$ und $\hat{e}_\epsilon^{(d)}$ werden die entsprechenden Normen analog definiert. Während die diskrete $L_2$-Norm (gewichtete euklidische Norm) und die Maximumsnorm unmittelbar auf dem dünnen Gitter berechnet werden, interpolieren wir aus rechentechnischen Gründen die Dünn-Gitter-Lösung $\tilde{u}_n^{(d)}$ sowie die exakte Lösung $u^{(d)}$ zunächst auf ein volles Gitter $G_m^{(d)}$ und berechnen anschließend dort die Energienorm der Differenz $e_m^{(d)}$:

$$\|\tilde{e}_n^{(d)}\|_{E_m} := \|e_m^{(d)}\|_E \;.$$

Im Falle adaptiv verfeinerter dünner Gitter wählen wir $m := 10$. Bei regulären dünnen Gittern gilt $m := \min(n, 10)$. An den Punkten von $G_m^{(d)}$, die nicht im dünnen Gitter enthalten sind, werden die entsprechenden Werte mittels Interpolation gebildet, wodurch ein zusätzlicher Interpolationsfehler entsteht. Enthält das dünne Gitter dagegen Punkte, die nicht zu $G_m^{(d)}$ gehören, so ist die gemessene Energie zu klein. Dieser Effekt führt insbesondere ab $n = 11$ zu schneller fallenden Werten von $\|\tilde{e}_n^{(d)}\|_{E_m}$.

Ferner verfolgen wir die Fehlerentwicklung im Mittelpunkt $P$ des betrachteten Gebietes, also $P = (\frac{1}{2}, \frac{1}{2})$ beziehungsweise $P = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$.

Als Maß für die Abnahme der $L_2$-, Maximums- und Energienorm des Fehlers sowie seines Wertes im Mittelpunkt $P$ bei wachsendem $n$ beziehungsweise abnehmendem $\epsilon$ verwenden wir das Verhältnis der jeweiligen Absolutwerte bei $n$ und $n+1$ beziehungsweise $\epsilon$ und $\epsilon/4$. Die resultierenden Quotienten nennen wir — je nach betrachteter Größe — $\rho_\infty$, $\rho_2$, $\rho_E$ oder $\rho_P$. Die Dimensionalität $d$, der Typ des betrachteten dünnen Gitters (regulär oder adaptiv verfeinert, $L_2$- oder energienormorientiert) sowie die Größe des Gitterparameters $n$ beziehungsweise $\epsilon$ ergeben sich dabei jeweils aus dem Kontext. Bei regulären dünnen Gittern zeigen die Quotienten $\rho_\infty$, $\rho_2$, $\rho_E$ und $\rho_P$ die Ordnung des Fehlers beziehungsweise der entsprechenden Fehlernorm. Bei quadratischer Ordnung liegen die Werte beispielsweise ungefähr bei 4.

In den nachfolgenden Rechenbeispielen soll auch die (algebraische) Konvergenzgeschwindigkeit unseres Verfahrens untersucht werden, d.h. die Anzahl an Iterationen, die zur Reduzierung des Startfehlers um einen vorgegebenen Faktor $\delta$ erforderlich sind. Da der Startfehler in Ermangelung der Lösung zu Beginn des Iterationsprozesses nicht berechnet werden kann, ermitteln wir stattdessen, wie oft iteriert werden muß, bis die aus einem Iterationsschritt resultierende Änderung der aktuellen Näherung weniger als das $\delta$-fache des Unterschieds zwischen Startnäherung und Näherung nach dem ersten Iterationsschritt beträgt. Mit der Bezeichnung $u_j$ für die Näherung nach $j$ Iterationsschritten und $u_\infty$ für das Ergebnis des Iterationsprozesses heißt dies, daß $m_\delta$ anstelle von

$$\|u_{m_\delta} - u_\infty\|_2 \;\leq\; \delta \cdot \|u_0 - u_\infty\|_2$$

der Bedingung

$$\|u_{m_\delta+1} - u_{m_\delta}\|_2 \;\leq\; \delta \cdot \|u_1 - u_0\|_2$$

zu genügen hat. Eine derartige Vorgehensweise legt zum Beispiel der Fixpunktsatz von Banach-Weissinger nahe, mit dessen Hilfe sich für lineare Iterationsverfahren mit Iterationsmatrix $M$ unter gewissen Voraussetzungen aus $\|u_{m_\delta+1} - u_{m_\delta}\|_2 \;\leq\; \delta \cdot \|u_1 - u_0\|_2$

$$\begin{aligned}
\|u_{m_\delta+1} - u_\infty\|_2 &\leq \|u_{m_\delta+1} - u_{m_\delta}\|_2 \cdot \sum_{k=1}^{\infty} \|M^k\|_2 \\
&\leq \delta \cdot \|u_1 - u_0\|_2 \cdot \sum_{k=1}^{\infty} \|M^k\|_2 \\
&\leq \delta \cdot \Big(\|u_1 - u_\infty\|_2 + \|u_0 - u_\infty\|_2\Big) \cdot \sum_{k=1}^{\infty} \|M^k\|_2 \\
&\leq \left(\delta \cdot \Big(\|M\|_2 + 1\Big) \cdot \sum_{k=1}^{\infty} \|M^k\|_2\right) \cdot \|u_0 - u_\infty\|_2
\end{aligned}$$

$$=: \quad \tilde{\delta} \cdot \|u_0 - u_\infty\|_2$$

folgern läßt. In sämtlichen numerischen Beispielen dieses Abschnitts betrachten wir stets die Situation im Falle $\delta := 10^{-6}$.

Schließlich geben wir noch in jedem Rechenbeispiel die Anzahl $K_n$ beziehungsweise $K_\epsilon$ der Knoten im Baum an, die zur Realisierung des betreffenden dünnen Gitters erforderlich sind. Für den Fall des regulären dünnen Gitters $\tilde{G}_n^{(2)}$ gilt beispielsweise

$$K_n \;=\; |\tilde{G}_n^{(2)}| + (2^n + 1) \;=\; 2^n \cdot (n+4) + 2,$$

da nach Abschnitt 5.2 der entsprechende Baum je einen Knoten für jeden Gitterpunkt und jede Gitterzeile enthält (siehe auch (22)).

Alle im folgenden dokumentierten numerischen Berechnungen wurden am Institut für Informatik der Technischen Universität München auf einer mit 64 MByte Hauptspeicher ausgestatteten DECstation 5000/200 von Digital Equipment durchgeführt. Sämtliche Programme wurden in C geschrieben.

**9.1. Zweidimensionale Probleme.** Wir betrachten zunächst ein Laplace-Problem mit glatter Lösung:

Beispiel 1:

$$\begin{array}{rcll} \Delta u(x,y) &=& 0 & \text{in } \ \Omega^{(2)}, \\ u(x,y) &=& \sin(\pi y) \cdot \frac{\sinh(\pi(1-x))}{\sinh(\pi)} & \text{auf } \ \delta\bar{\Omega}^{(2)}. \end{array}$$

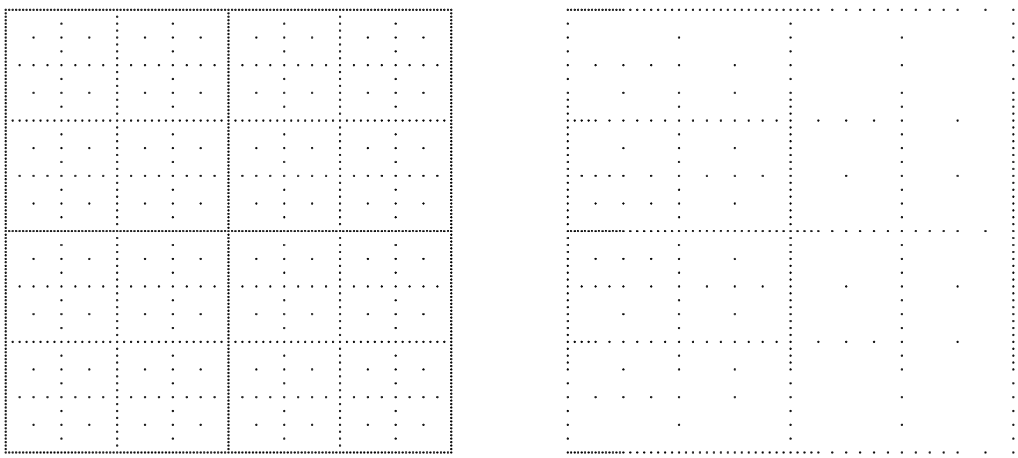


ABB. 19. *Beispiel 1: Reguläres und adaptives dünnes Gitter ($\tilde{G}_7^{(2)}$ bzw. $\epsilon = 4^{-5}$).*

Abbildung 19 zeigt das reguläre dünne Gitter $\tilde{G}_7^{(2)}$ sowie das aus der Wahl $\epsilon = 4^{-5}$ resultierende, adaptiv verfeinerte dünne Gitter. In Abbildung 20 sehen wir die auf dem regulären dünnen Gitter $\tilde{G}_{10}^{(2)}$ berechnete Näherungslösung $\tilde{u}_{10}^{(2)}$ sowie ihren Fehler $\tilde{e}_{10}^{(2)}$. Zur graphischen Ausgabe wurde hierbei das volle Gitter $G_5^{(2)}$ verwendet, dessen Punkte

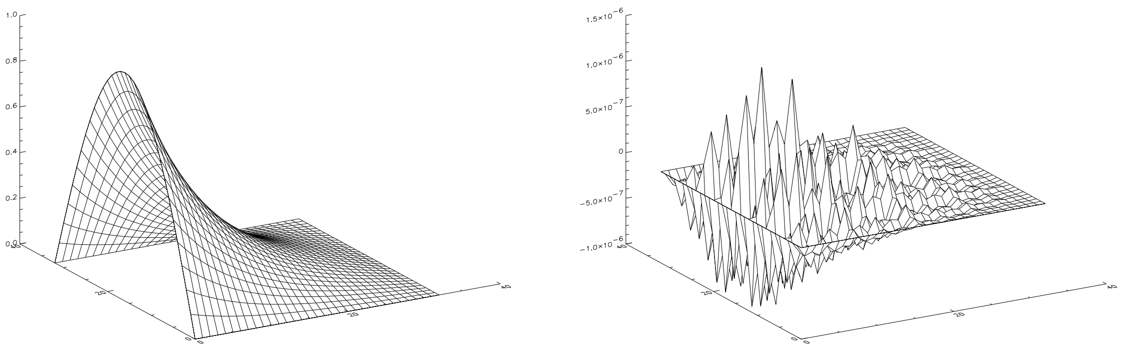


ABB. 20. *Beispiel 1: Die Lösung $\tilde{u}_{10}^{(2)}$ und ihr Fehler $\tilde{e}_{10}^{(2)}$.*

jedoch alle auch in $\tilde{G}_{10}^{(2)}$ enthalten sind. Man beachte, daß im Gegensatz zu auf vollen Gittern ermittelten Näherungslösungen der Fehler $\tilde{e}_{10}^{(2)}$ stark oszilliert.

Tabelle 6 zeigt die Resultate im regulären Fall. Das feinste verwendete Gitter $\tilde{G}_{14}^{(2)}$ führt zu einem Baum mit 294914 Knoten. Nach Satz 3.3 entfallen davon 278529 auf die Darstellung der Gitterpunkte, $2^{14} + 1 = 16385$ Knoten werden gemäß Abschnitt 5.2 zur Darstellung der Gitterlinien benötigt. Da ein Knoten im Baum bei Laplace-Problemen und Rechnung in doppelter Genauigkeit 64 Byte Speicherplatz erfordert, ergibt sich der Gesamtspeicherbedarf im Falle $n = 14$ zu 18.9 MByte. Mit der Variante unseres Verfahrens aus Abschnitt 8.1, die bei Laplace-Problemen mit 32 Byte pro Knoten auskommt, beträgt der Gesamtspeicherbedarf im Falle $n = 14$ sogar nur 9.5 MByte. Demgegenüber benötigt man bei Verwendung des vollen Gitters $G_{14}^{(2)}$ bereits mehr als 2.1 GByte. Bei Poisson-Problemen (siehe Beispiel 3) ist in jedem Gitterpunkt eine weitere reelle Variable für die rechte Seite der Differentialgleichung erforderlich, so daß sich hier ein Gesamtspeicherbedarf für $\tilde{G}_{14}^{(2)}$ von 21.2 MByte beziehungsweise 11.8 MByte gegenüber 4.3 GByte für $G_{14}^{(2)}$ ergibt.

TABELLE 6
*Beispiel 1, reguläre dünne Gitter $\tilde{G}_n^{(2)}$.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\Vert\tilde{e}_n^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\tilde{e}_n^{(2)}\Vert_2$ | $\rho_2$ | $\Vert\tilde{e}_n^{(2)}\Vert_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $-7.427 \cdot 10^{-2}$ | | $7.427 \cdot 10^{-2}$ | | $7.427 \cdot 10^{-2}$ | | $1.213 \cdot 10^{-1}$ | | 1 | 12 |
| 2 | $-5.917 \cdot 10^{-3}$ | 12.55 | $3.502 \cdot 10^{-2}$ | 2.12 | $2.018 \cdot 10^{-2}$ | 3.68 | $6.716 \cdot 10^{-2}$ | 1.81 | 7 | 26 |
| 3 | $5.756 \cdot 10^{-4}$ | 10.28 | $1.352 \cdot 10^{-2}$ | 2.59 | $5.461 \cdot 10^{-3}$ | 3.70 | $3.649 \cdot 10^{-2}$ | 1.84 | 22 | 58 |
| 4 | $4.739 \cdot 10^{-4}$ | 1.21 | $4.528 \cdot 10^{-3}$ | 2.99 | $1.449 \cdot 10^{-3}$ | 3.77 | $1.879 \cdot 10^{-2}$ | 1.94 | 26 | 130 |
| 5 | $2.079 \cdot 10^{-4}$ | 2.28 | $1.372 \cdot 10^{-3}$ | 3.30 | $3.782 \cdot 10^{-4}$ | 3.83 | $9.457 \cdot 10^{-3}$ | 1.99 | 36 | 290 |
| 6 | $7.588 \cdot 10^{-5}$ | 2.74 | $4.256 \cdot 10^{-4}$ | 3.22 | $9.767 \cdot 10^{-5}$ | 3.87 | $4.732 \cdot 10^{-3}$ | 2.00 | 47 | 642 |
| 7 | $2.505 \cdot 10^{-5}$ | 3.03 | $1.282 \cdot 10^{-4}$ | 3.32 | $2.506 \cdot 10^{-5}$ | 3.90 | $2.366 \cdot 10^{-3}$ | 2.00 | 62 | 1410 |
| 8 | $7.801 \cdot 10^{-6}$ | 3.21 | $3.701 \cdot 10^{-5}$ | 3.46 | $6.401 \cdot 10^{-6}$ | 3.92 | $1.183 \cdot 10^{-3}$ | 2.00 | 81 | 3074 |
| 9 | $2.336 \cdot 10^{-6}$ | 3.34 | $1.050 \cdot 10^{-5}$ | 3.52 | $1.631 \cdot 10^{-6}$ | 3.92 | $5.912 \cdot 10^{-4}$ | 2.00 | 113 | 6658 |
| 10 | $6.803 \cdot 10^{-7}$ | 3.43 | $2.925 \cdot 10^{-6}$ | 3.59 | $4.147 \cdot 10^{-7}$ | 3.93 | $2.956 \cdot 10^{-4}$ | 2.00 | 151 | 14338 |
| 11 | $1.942 \cdot 10^{-7}$ | 3.50 | $8.166 \cdot 10^{-7}$ | 3.58 | $1.053 \cdot 10^{-7}$ | 3.94 | $6.671 \cdot 10^{-5}$ | 4.43 | 213 | 30722 |
| 12 | $5.456 \cdot 10^{-8}$ | 3.56 | $2.232 \cdot 10^{-7}$ | 3.66 | $2.673 \cdot 10^{-8}$ | 3.94 | $1.512 \cdot 10^{-5}$ | 4.41 | 287 | 65538 |
| 13 | $1.509 \cdot 10^{-8}$ | 3.62 | $6.103 \cdot 10^{-8}$ | 3.66 | $6.777 \cdot 10^{-9}$ | 3.94 | $3.508 \cdot 10^{-6}$ | 4.31 | 403 | 139266 |
| 14 | $4.164 \cdot 10^{-9}$ | 3.62 | $1.643 \cdot 10^{-8}$ | 3.71 | $1.718 \cdot 10^{-9}$ | 3.94 | $8.372 \cdot 10^{-7}$ | 4.19 | 554 | 294914 |

Man erkennt in Tabelle 6 deutlich das $h_n^2 \cdot \mathrm{ld}(h_n^{-1})$-Verhalten der $L_2$- und der Maximumsnorm sowie das lineare Abfallen der Energienorm. Die für die beiden erstgenannten Normen des Interpolationsfehlers in Satz 3.5 bewiesene Ordnung $O(h_n^2 \cdot \mathrm{ld}(h_n^{-1}))$ gilt hier also auch für die entsprechenden Normen des Fehlers der berechneten Näherungslösung für unser Modellproblem. Der sprunghafte Anstieg der Werte $\rho_E$ ab $n = 11$ liegt daran, daß wir die Energie auf dem regulären vollen Gitter $G_m^{(2)}$, $m = \min(n, 10)$, messen und deshalb für $n > 10$ gewisse Punkte des dünnen Gitters $\tilde{G}_n^{(2)}$ nicht in die Berechnung der Energienorm eingehen. Was nun die Konvergenzgeschwindigkeit betrifft, so nehmen die Werte von $m_\delta$ beim Übergang von $n$ zu $n+1$ jeweils höchstens um einen Faktor $\sqrt{2}$ zu. Es stellt sich also heraus, daß das tatsächliche Konvergenzverhalten erheblich besser ist, als es das (auf groben Abschätzungen beruhende) Resultat $m_\delta = O(2^{\frac{5}{4}n})$ aus Abschnitt 6.2.4 erwarten ließ.

In Abbildung 21 ist das Verhalten des durch $h_n^2$ dividierten Fehlers im Mittelpunkt dargestellt. Der ab $n = 5$ nahezu lineare Verlauf bestätigt die Vermutung, daß auch der punktweise Fehler durch einen Ausdruck der Ordnung $O(h_n^2 \cdot \mathrm{ld}(h_n^{-1}))$ beschränkt ist. Darüber hinaus scheint es sogar eine Reihenentwicklung des Fehlers zu geben mit $h_n^2 \cdot \mathrm{ld}(h_n^{-1})$ als führendem Glied.

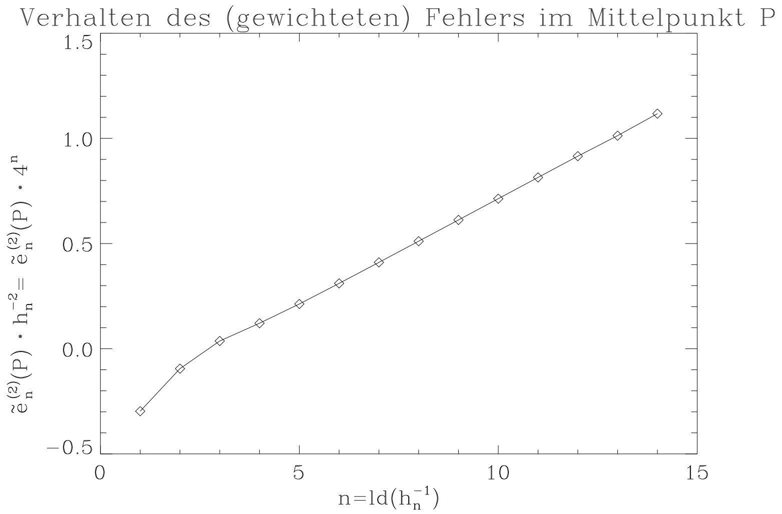



ABB. 21. *Mit $h_n^{-2}$ multiplizierter Fehler $\tilde{e}_n^{(2)}(P)$ im Mittelpunkt $P = (\frac{1}{2}, \frac{1}{2})$.*

In Tabelle 7 sind die Ergebnisse bei adaptiver Verfeinerung zusammengefaßt. Als Kriterium für die Erzeugung neuer Gitterpunkte gelangt dabei (38) zur Anwendung. Wir betrachten hier also $L_2$-normorientierte Gitter. Da (38) in jedem Knoten den hierarchischen Überschuß mit dem Adaptionsparameter $\epsilon$ vergleicht und da ferner der hierarchische Überschuß nach (12) durch einen Ausdruck der Ordnung $O(h_1^2 \cdot h_2^2)$ beschränkt ist, benutzen wir als adaptives Analogon des Übergangs von $n$ zu $n+1$ den Übergang von $\epsilon$ zu $\frac{\epsilon}{4}$. Die ersten neun Zeilen in Tabelle 7 enthalten die Werte, die sich bei Verwendung des entsprechenden $\epsilon$ vom ersten Iterationsschritt an ergeben. Für die Einträge in der letzten Zeile haben wir mit $\epsilon = 4^{-1}$ begonnen, einige Iterationsschritte durchgeführt und anschließend $\epsilon$ auf $4^{-2}$ gesetzt. Diese Vorgehensweise haben wir bis

$\epsilon = 4^{-9}$ wiederholt. Es stellt sich heraus, daß eine solche sukzessive Gitterverfeinerung wesentlich zurückhaltender in der Erzeugung neuer Gitterpunkte ist.

TABELLE 7
*Beispiel 1, adaptiv erzeugte dünne Gitter (Kriterium (38)).*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_\epsilon^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_\epsilon^{(2)}\|_2$ | $\rho_2$ | $\|\tilde{e}_\epsilon^{(2)}\|_{E_{10}}$ | $\rho_E$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-5.917 \cdot 10^{-3}$ | | $3.501 \cdot 10^{-2}$ | | $2.018 \cdot 10^{-2}$ | | $2.871 \cdot 10^{-1}$ | | 26 |
| $4^{-2}$ | $-9.206 \cdot 10^{-4}$ | 6.43 | $1.237 \cdot 10^{-2}$ | 2.83 | $6.416 \cdot 10^{-3}$ | 3.15 | $1.272 \cdot 10^{-1}$ | 2.26 | 58 |
| $4^{-3}$ | $3.091 \cdot 10^{-4}$ | 2.98 | $4.537 \cdot 10^{-3}$ | 2.73 | $2.239 \cdot 10^{-3}$ | 2.87 | $7.694 \cdot 10^{-2}$ | 1.65 | 100 |
| $4^{-4}$ | $4.122 \cdot 10^{-4}$ | 0.75 | $1.132 \cdot 10^{-3}$ | 4.01 | $4.693 \cdot 10^{-4}$ | 4.77 | $3.402 \cdot 10^{-2}$ | 2.26 | 246 |
| $4^{-5}$ | $1.266 \cdot 10^{-4}$ | 3.26 | $2.992 \cdot 10^{-4}$ | 3.78 | $1.282 \cdot 10^{-4}$ | 3.66 | $1.776 \cdot 10^{-2}$ | 1.92 | 514 |
| $4^{-6}$ | $2.338 \cdot 10^{-5}$ | 5.41 | $7.451 \cdot 10^{-5}$ | 4.02 | $3.103 \cdot 10^{-5}$ | 4.13 | $8.912 \cdot 10^{-3}$ | 1.99 | 1150 |
| $4^{-7}$ | $6.774 \cdot 10^{-6}$ | 3.45 | $1.942 \cdot 10^{-5}$ | 3.84 | $7.268 \cdot 10^{-6}$ | 4.27 | $4.256 \cdot 10^{-3}$ | 2.09 | 2812 |
| $4^{-8}$ | $1.819 \cdot 10^{-6}$ | 3.72 | $5.290 \cdot 10^{-6}$ | 3.67 | $1.714 \cdot 10^{-6}$ | 4.24 | $1.920 \cdot 10^{-3}$ | 2.22 | 6838 |
| $4^{-9}$ | $4.310 \cdot 10^{-7}$ | 4.22 | $1.252 \cdot 10^{-6}$ | 4.23 | $3.991 \cdot 10^{-7}$ | 4.29 | $4.952 \cdot 10^{-4}$ | 3.88 | 20266 |
| $4^{-9}$ | $1.184 \cdot 10^{-6}$ | – | $1.496 \cdot 10^{-6}$ | – | $5.373 \cdot 10^{-7}$ | – | $5.165 \cdot 10^{-4}$ | – | 10794 |

Bereits bei diesem Beispiel einer glatten Lösung zeigt sich, daß der Einsatz adaptiver Verfeinerung von Vorteil ist. So erreichen wir mit weniger als 11000 Knoten einen maximalen Fehler von $1.5 \cdot 10^{-6}$, wofür im regulären Falle bereits das Gitter $\tilde{G}_{11}^{(2)}$ mit über 30000 Knoten erforderlich ist. Noch deutlicher werden die Vorzüge adaptiv verfeinerter dünner Gitter am folgenden Beispiel, in welchem die Lösung eine Randsingularität aufweist.

Beispiel 2:

$$\begin{aligned} \Delta u(x,y) &= 0 && \text{in } \ \Omega := (0,1) \times (-\tfrac{1}{2}, \tfrac{1}{2}), \\ u(x,y) &= \mathrm{Re}(z^{\frac{1}{2}}) \\ &= (x^2+y^2)^{\frac{1}{4}} \cdot \cos\left(\tfrac{\arctan(x/y)}{2}\right) && \text{auf } \ \delta\bar{\Omega}. \end{aligned}$$

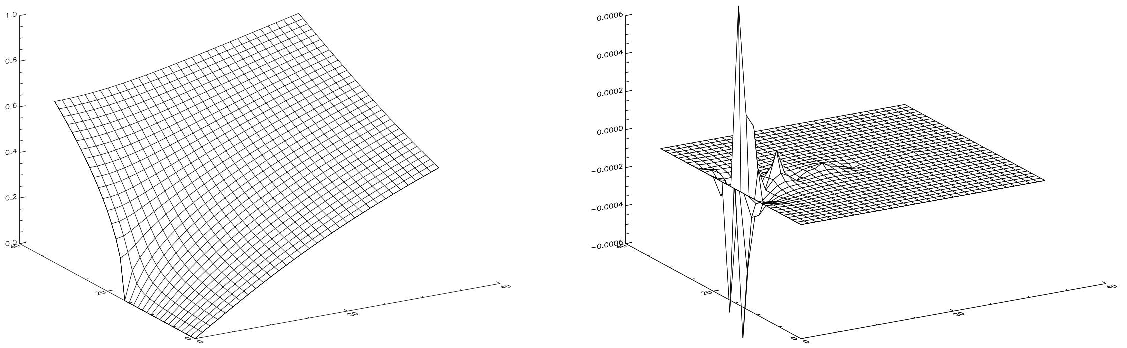



ABB. 22. *Beispiel 2: Die Lösung $\tilde{u}_{10}^{(2)}$ und ihr Fehler $\tilde{e}_{10}^{(2)}$.*

Die Lösung $u(x,y)$ ist singulär im Punkt $(0,0)$ und erfüllt somit nicht die Glattheitsanforderung $u(x,y) \in X^{(2)}$. In Abbildung 22 sind die Näherungslösung $\tilde{u}_{10}^{(2)}$ sowie ihr Fehler $\tilde{e}_{10}^{(2)}$ auf dem vollen Gitter $G_5^{(2)}$ dargestellt. Deutlich erkennt man wiederum das stark oszillierende Verhalten des Fehlers $\tilde{e}_{10}^{(2)}$. Abbildung 23 zeigt das reguläre dünne Gitter $\tilde{G}_7^{(2)}$ sowie das aus der Wahl $\epsilon = 4^{-4}$ resultierende, adaptiv verfeinerte dünne Gitter.

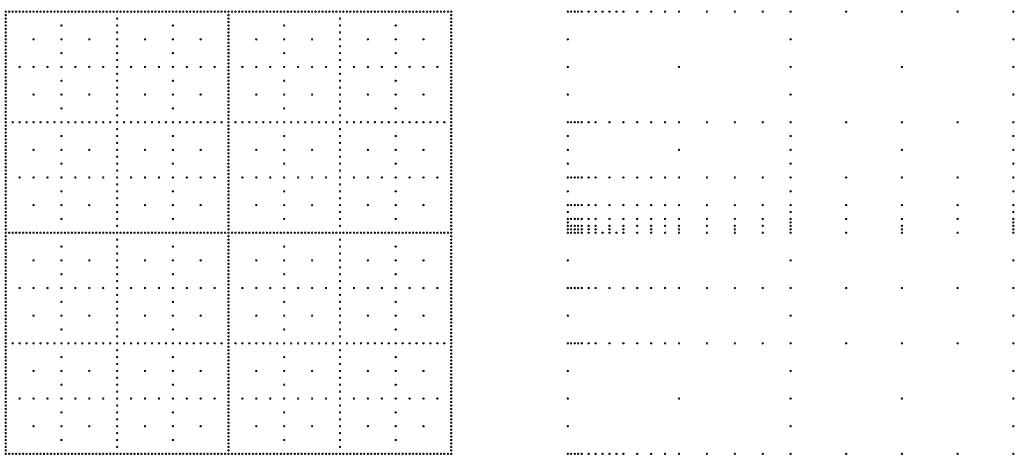


Abb. 23. *Beispiel 2: Reguläres und adaptives dünnes Gitter ($\tilde{G}_7^{(2)}$ bzw. $\epsilon = 4^{-4}$).*

Tabelle 8
*Beispiel 2, reguläre dünne Gitter $\tilde{G}_n^{(2)}$.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\Vert\tilde{e}_n^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\tilde{e}_n^{(2)}\Vert_2$ | $\rho_2$ | $\Vert\tilde{e}_n^{(2)}\Vert_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $-3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $6.482 \cdot 10^{-3}$ | | 1 | 12 |
| 2 | $8.294 \cdot 10^{-3}$ | 0.48 | $3.987 \cdot 10^{-2}$ | 0.10 | $2.215 \cdot 10^{-2}$ | 0.18 | $9.846 \cdot 10^{-2}$ | 0.07 | 7 | 26 |
| 3 | $4.830 \cdot 10^{-3}$ | 1.72 | $5.668 \cdot 10^{-2}$ | 0.70 | $1.625 \cdot 10^{-2}$ | 1.36 | $1.469 \cdot 10^{-1}$ | 0.67 | 23 | 58 |
| 4 | $9.807 \cdot 10^{-4}$ | 4.93 | $5.907 \cdot 10^{-2}$ | 0.96 | $1.062 \cdot 10^{-2}$ | 1.53 | $1.772 \cdot 10^{-1}$ | 0.83 | 25 | 130 |
| 5 | $-1.366 \cdot 10^{-4}$ | 7.18 | $5.403 \cdot 10^{-2}$ | 1.09 | $6.814 \cdot 10^{-3}$ | 1.56 | $1.957 \cdot 10^{-1}$ | 0.91 | 32 | 290 |
| 6 | $-2.328 \cdot 10^{-4}$ | 0.59 | $4.588 \cdot 10^{-2}$ | 1.18 | $4.402 \cdot 10^{-3}$ | 1.55 | $2.068 \cdot 10^{-1}$ | 0.95 | 44 | 642 |
| 7 | $-1.749 \cdot 10^{-4}$ | 1.33 | $3.790 \cdot 10^{-2}$ | 1.21 | $2.852 \cdot 10^{-3}$ | 1.54 | $2.133 \cdot 10^{-1}$ | 0.97 | 58 | 1410 |
| 8 | $-9.344 \cdot 10^{-5}$ | 1.87 | $3.234 \cdot 10^{-2}$ | 1.17 | $1.870 \cdot 10^{-3}$ | 1.53 | $2.170 \cdot 10^{-1}$ | 0.98 | 76 | 3074 |
| 9 | $-4.359 \cdot 10^{-5}$ | 2.14 | $2.738 \cdot 10^{-2}$ | 1.18 | $1.232 \cdot 10^{-3}$ | 1.52 | $2.196 \cdot 10^{-1}$ | 0.99 | 102 | 6658 |
| 10 | $-1.893 \cdot 10^{-5}$ | 2.30 | $2.289 \cdot 10^{-2}$ | 1.20 | $8.203 \cdot 10^{-4}$ | 1.50 | $2.212 \cdot 10^{-1}$ | 0.99 | 131 | 14338 |
| 11 | $-7.891 \cdot 10^{-6}$ | 2.40 | $1.929 \cdot 10^{-2}$ | 1.19 | $5.478 \cdot 10^{-4}$ | 1.50 | $1.549 \cdot 10^{-1}$ | 1.43 | 184 | 30772 |
| 12 | $-3.195 \cdot 10^{-6}$ | 2.47 | $1.623 \cdot 10^{-2}$ | 1.19 | $3.691 \cdot 10^{-4}$ | 1.48 | $1.083 \cdot 10^{-1}$ | 1.43 | 245 | 65538 |

Tabelle 8 liefert die Ergebnisse für reguläre dünne Gitter $\tilde{G}_n^{(2)}$, Tabelle 9 stellt die Resultate im Falle adaptiver Verfeinerung zusammen, wobei als Kriterium für die Erzeugung neuer Gitterpunkte wiederum (38) verwendet wird. Auf regulären vollen Gittern erhält man in der Nähe der Singularität bekanntlich punktweise Konvergenzraten der Ordnung $O(h_n^{1/2})$, wohingegen in ausreichender Entfernung von der Singularität punktweise Konvergenzraten der Ordnung $O(h_n^{3/2})$ erzielt werden können (siehe [27]). Auf regulären vollen Gittern können wir folglich $\rho_\infty \approx (h_n/h_{n+1})^{1/2} = \sqrt{2}$ sowie $\rho_P \approx (h_n/h_{n+1})^{3/2} = \sqrt{8}$ erwarten. Die entsprechenden Werte von $\rho_\infty$ und $\rho_P$ in Tabelle 8 sind dagegen kleiner als $\sqrt{2}$ beziehungsweise $\sqrt{8}$. Über die genauen Auswirkungen der Verwendung dünner Gitter liegen jedoch für den Fall nicht glatter Funktionen bislang keine Erkenntnisse vor. Zur Erläuterung der Werte von $\rho_E$ in Tabelle 8 sei wiederum darauf verwiesen, daß wir die Energie jeweils auf dem vollen Gitter $G_m^{(2)}$, $m = \min(n, 10)$, berechnen. Bezüglich der Konvergenzgeschwindigkeit sind keine wesentlichen Unterschiede zum glatten Fall im ersten Beispiel auszumachen.

Das Auftreten der Singularität in der Lösung unseres zweiten Modellproblems führt unter anderem dazu, daß sich die adaptive Gitterverfeinerung stärker als bei glatten Funktionen niederschlägt. Bei vergleichbarer Knotenzahl liefern adaptiv verfeinerte Gitter deutlich kleinere maximale Fehler als reguläre dünne Gitter. Wie man der letzten Zeile aus Tabelle 9 entnehmen kann, gilt dies insbesondere dann, wenn $\epsilon$ im Laufe des Iterationsprozesses schrittweise verkleinert wird.

TABELLE 9
*Beispiel 2, adaptiv erzeugte dünne Gitter (Kriterium (38)).*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\Vert\tilde{e}_\epsilon^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\tilde{e}_\epsilon^{(2)}\Vert_2$ | $\rho_2$ | $\Vert\tilde{e}_\epsilon^{(2)}\Vert_{E_{10}}$ | $\rho_E$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $2.874 \cdot 10^{-1}$ | | 12 |
| $4^{-2}$ | $7.268 \cdot 10^{-3}$ | 0.55 | $3.402 \cdot 10^{-2}$ | 0.12 | $1.176 \cdot 10^{-2}$ | 0.34 | $2.153 \cdot 10^{-1}$ | 1.33 | 56 |
| $4^{-3}$ | $2.130 \cdot 10^{-3}$ | 3.41 | $3.554 \cdot 10^{-2}$ | 0.96 | $9.130 \cdot 10^{-3}$ | 1.29 | $1.935 \cdot 10^{-1}$ | 1.11 | 174 |
| $4^{-4}$ | $6.196 \cdot 10^{-4}$ | 3.44 | $2.257 \cdot 10^{-2}$ | 1.57 | $7.091 \cdot 10^{-3}$ | 1.29 | $8.238 \cdot 10^{-2}$ | 2.35 | 472 |
| $4^{-5}$ | $1.342 \cdot 10^{-4}$ | 4.62 | $1.574 \cdot 10^{-2}$ | 1.43 | $4.257 \cdot 10^{-3}$ | 1.67 | $5.784 \cdot 10^{-2}$ | 1.42 | 1220 |
| $4^{-6}$ | $3.536 \cdot 10^{-5}$ | 3.80 | $1.115 \cdot 10^{-2}$ | 1.41 | $2.461 \cdot 10^{-3}$ | 1.73 | $3.742 \cdot 10^{-2}$ | 1.55 | 3434 |
| $4^{-7}$ | $8.241 \cdot 10^{-6}$ | 4.29 | $7.861 \cdot 10^{-3}$ | 1.42 | $1.465 \cdot 10^{-3}$ | 1.68 | $2.455 \cdot 10^{-2}$ | 1.52 | 9782 |
| $4^{-8}$ | $2.208 \cdot 10^{-6}$ | 3.73 | $5.559 \cdot 10^{-3}$ | 1.41 | $9.753 \cdot 10^{-4}$ | 1.50 | $1.547 \cdot 10^{-2}$ | 1.59 | 27862 |
| $4^{-9}$ | $6.359 \cdot 10^{-7}$ | 3.47 | $3.932 \cdot 10^{-3}$ | 1.41 | $6.330 \cdot 10^{-4}$ | 1.54 | $8.864 \cdot 10^{-3}$ | 1.75 | 78748 |
| $4^{-9}$ | $-1.018 \cdot 10^{-7}$ | – | $3.934 \cdot 10^{-3}$ | – | $6.245 \cdot 10^{-4}$ | – | $8.868 \cdot 10^{-3}$ | – | 27154 |

Beispiel 3:

$$\begin{aligned} \Delta u(x,y) &= -2\pi^2 \cdot \sin(\pi x) \cdot \sin(\pi y) && \text{in } \Omega^{(2)}, \\ u(x,y) &= 0 && \text{auf } \delta\bar{\Omega}^{(2)}. \end{aligned}$$

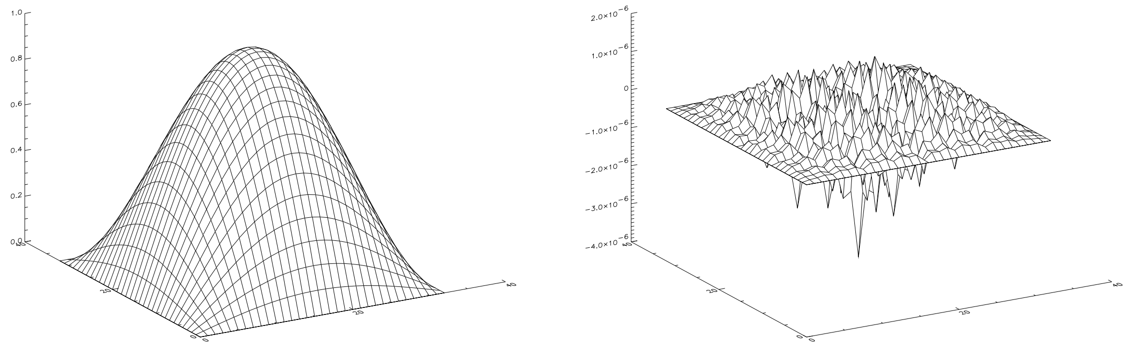



ABB. 24. *Beispiel 3: Die Lösung $\tilde{u}_{10}^{(2)}$ und ihr Fehler $\tilde{e}_{10}^{(2)}$ ($n_f = 14$).*

In Abbildung 24 sind wieder die Näherungslösung $\tilde{u}_{10}^{(2)}$ und ihr Fehler $\tilde{e}_{10}^{(2)}$ auf dem vollen Gitter $G_5^{(2)}$ dargestellt, wobei für die Vorlaufrechnung das reguläre dünne Gitter

$\tilde{G}_{14}^{(2)}$ gewählt wurde. Abbildung 25 zeigt das aus der Wahl $\epsilon = 4^{-5}$ resultierende, adaptiv verfeinerte dünne Gitter sowie zum Vergleich wieder das reguläre dünne Gitter $\tilde{G}_{7}^{(2)}$. Dabei sind nur solche Gitterpunkte eingetragen, die zur Darstellung der Näherungslösung erforderlich sind. Gitterpunkte, die ausschließlich während der Vorlaufrechnung zur Integration der rechten Seite der Poisson-Gleichung erzeugt worden sind (siehe Abschnitt 6.2.3), tauchen in Abbildung 25 nicht auf.

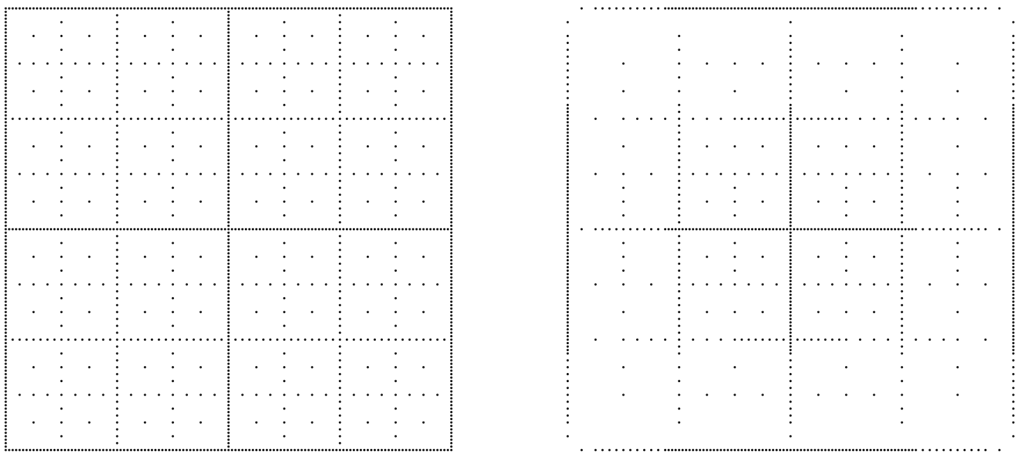


ABB. 25. *Beispiel 3: Reguläres und adaptives dünnes Gitter ($\tilde{G}_{7}^{(2)}$ bzw. $\epsilon = 4^{-5}$).*

Tabelle 10 zeigt die Resultate unseres zweidimensionalen Poisson-Beispiels im regulären Fall. Für die Vorlaufrechnung wurde dabei das Gitter $\tilde{G}_{14}^{(2)}$ (d.h. $n_f = 14$) benutzt. Solange nun das zur Berechnung der Näherungslösung herangezogene $n = n_u$ deutlich kleiner als 14 ist, nehmen die verschiedenen Fehlernormen sowie der Fehler im Mittelpunkt in erwartetem Maße ab. Die Größe $\rho_2$ zeigt hier sogar $h_n^2$-Charakteristik. Ab $n = 13$ ist das Konvergenzverhalten dagegen merklich schlechter, da die Genauigkeit des Integrationsprozesses während der Vorlaufrechnung jetzt nicht mehr ausreicht, die Genauigkeit der für unser betrachtetes Poisson-Problem produzierten Näherungslösung wie bisher weiter zu steigern. Für das Verhalten von $\rho_E$ gilt wiederum das bereits bei den vorangegangenen Beispielen Gesagte.

Tabelle 11 stellt die Ergebnisse bei adaptiver Verfeinerung mittels des Kriteriums (38) zusammen. Der während der Vorlaufrechnung mit $\epsilon_f = 10^{-6}$ erzeugte Baum besteht dabei aus 140398 Knoten. Die Größe $K_\epsilon^{neu}$ in Tabelle 11 gibt die Anzahl der während des Iterationsprozesses zusätzlich erzeugten Knoten an. Da die Vorlaufrechnung im Falle $K_\epsilon^{neu} \neq 0$ jedoch aus Effizienzgründen nicht neu gestartet wird, sondern vielmehr für die entsprechenden Integrale grobe Näherungswerte eingetragen werden, wirken sich solche Knoten ungünstig auf die Qualität der erzielten Lösung aus. Eine sukzessive Verfeinerungsstrategie mit allmählich abnehmendem $\epsilon$ kann dazu beitragen, die Generierung neuer Knoten während des Iterationsprozesses zu vermeiden (siehe die letzte Zeile in Tabelle 11).

TABELLE 10

*Beispiel 3, reguläre dünne Gitter $\tilde{G}_n^{(2)}$ $(n_f = 14)$.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_n^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_n^{(2)}\|_2$ | $\rho_2$ | $\|\tilde{e}_n^{(2)}\|_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $2.159 \cdot 10^{-1}$ | | $2.159 \cdot 10^{-1}$ | | $2.159 \cdot 10^{-1}$ | | $3.525 \cdot 10^{-1}$ | | 1 | 12 |
| 2 | $2.950 \cdot 10^{-2}$ | 7.32 | $5.612 \cdot 10^{-2}$ | 3.85 | $5.190 \cdot 10^{-2}$ | 4.16 | $1.458 \cdot 10^{-1}$ | 2.42 | 7 | 26 |
| 3 | $-7.220 \cdot 10^{-4}$ | 40.86 | $1.888 \cdot 10^{-2}$ | 2.97 | $1.274 \cdot 10^{-2}$ | 4.07 | $6.922 \cdot 10^{-2}$ | 2.11 | 22 | 58 |
| 4 | $-2.206 \cdot 10^{-3}$ | 0.33 | $5.161 \cdot 10^{-3}$ | 3.66 | $3.125 \cdot 10^{-3}$ | 4.08 | $2.969 \cdot 10^{-2}$ | 2.33 | 27 | 130 |
| 5 | $-1.051 \cdot 10^{-3}$ | 2.10 | $1.450 \cdot 10^{-3}$ | 3.56 | $7.731 \cdot 10^{-4}$ | 4.04 | $1.458 \cdot 10^{-2}$ | 2.04 | 37 | 290 |
| 6 | $-3.869 \cdot 10^{-4}$ | 2.72 | $3.978 \cdot 10^{-4}$ | 3.65 | $1.927 \cdot 10^{-4}$ | 4.01 | $7.253 \cdot 10^{-3}$ | 2.01 | 48 | 642 |
| 7 | $-1.277 \cdot 10^{-4}$ | 3.03 | $1.277 \cdot 10^{-4}$ | 3.12 | $4.832 \cdot 10^{-5}$ | 3.99 | $3.622 \cdot 10^{-3}$ | 2.00 | 63 | 1410 |
| 8 | $-3.970 \cdot 10^{-5}$ | 3.22 | $3.970 \cdot 10^{-5}$ | 3.22 | $1.215 \cdot 10^{-5}$ | 3.98 | $1.810 \cdot 10^{-3}$ | 2.00 | 84 | 3074 |
| 9 | $-1.188 \cdot 10^{-5}$ | 3.34 | $1.188 \cdot 10^{-5}$ | 3.34 | $3.054 \cdot 10^{-6}$ | 3.98 | $9.050 \cdot 10^{-4}$ | 2.00 | 118 | 6658 |
| 10 | $-3.476 \cdot 10^{-6}$ | 3.42 | $3.476 \cdot 10^{-6}$ | 3.42 | $7.620 \cdot 10^{-7}$ | 4.01 | $4.525 \cdot 10^{-4}$ | 2.00 | 159 | 14338 |
| 11 | $-1.012 \cdot 10^{-6}$ | 3.43 | $1.012 \cdot 10^{-6}$ | 3.43 | $1.836 \cdot 10^{-7}$ | 4.15 | $9.587 \cdot 10^{-5}$ | 4.72 | 223 | 30722 |
| 12 | $-3.051 \cdot 10^{-7}$ | 3.32 | $3.051 \cdot 10^{-7}$ | 3.32 | $3.867 \cdot 10^{-8}$ | 4.75 | $2.295 \cdot 10^{-5}$ | 4.18 | 301 | 65538 |
| 13 | $-1.057 \cdot 10^{-7}$ | 2.89 | $1.057 \cdot 10^{-7}$ | 2.89 | $1.055 \cdot 10^{-8}$ | 3.67 | $5.678 \cdot 10^{-6}$ | 4.04 | 422 | 139266 |
| 14 | $-5.018 \cdot 10^{-8}$ | 2.11 | $5.018 \cdot 10^{-8}$ | 2.11 | $1.373 \cdot 10^{-8}$ | 0.77 | $1.419 \cdot 10^{-6}$ | 4.00 | 579 | 294914 |

TABELLE 11

*Beispiel 3, adaptiv erzeugte dünne Gitter (Kriterium (38), $\epsilon_f = 10^{-6}$).*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_\epsilon^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_\epsilon^{(2)}\|_2$ | $\rho_2$ | $\|\tilde{e}_\epsilon^{(2)}\|_{E_{10}}$ | $\rho_E$ | $K_\epsilon^{neu}$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $2.950 \cdot 10^{-2}$ | | $5.612 \cdot 10^{-2}$ | | $5.190 \cdot 10^{-2}$ | | $5.119 \cdot 10^{-1}$ | | 0 |
| $4^{-2}$ | $-7.243 \cdot 10^{-3}$ | 4.07 | $1.030 \cdot 10^{-2}$ | 5.45 | $6.431 \cdot 10^{-3}$ | 8.07 | $1.632 \cdot 10^{-1}$ | 3.14 | 0 |
| $4^{-3}$ | $-3.845 \cdot 10^{-3}$ | 1.88 | $3.845 \cdot 10^{-3}$ | 2.68 | $1.836 \cdot 10^{-3}$ | 3.50 | $8.242 \cdot 10^{-2}$ | 1.98 | 0 |
| $4^{-4}$ | $-3.865 \cdot 10^{-4}$ | 9.95 | $8.198 \cdot 10^{-4}$ | 4.69 | $3.841 \cdot 10^{-4}$ | 4.78 | $4.057 \cdot 10^{-2}$ | 2.03 | 0 |
| $4^{-5}$ | $-8.510 \cdot 10^{-5}$ | 4.54 | $2.203 \cdot 10^{-4}$ | 3.72 | $1.015 \cdot 10^{-4}$ | 3.78 | $1.945 \cdot 10^{-2}$ | 2.09 | 0 |
| $4^{-6}$ | $-2.050 \cdot 10^{-5}$ | 4.15 | $6.110 \cdot 10^{-5}$ | 3.61 | $2.199 \cdot 10^{-5}$ | 4.62 | $9.720 \cdot 10^{-3}$ | 2.00 | 0 |
| $4^{-7}$ | $-3.532 \cdot 10^{-6}$ | 5.80 | $1.592 \cdot 10^{-5}$ | 3.84 | $5.185 \cdot 10^{-6}$ | 4.24 | $4.490 \cdot 10^{-3}$ | 2.16 | 8 |
| $4^{-8}$ | $-1.178 \cdot 10^{-6}$ | 3.00 | $4.250 \cdot 10^{-6}$ | 3.75 | $1.208 \cdot 10^{-6}$ | 4.29 | $1.402 \cdot 10^{-3}$ | 3.20 | 864 |
| $4^{-9}$ | $-1.834 \cdot 10^{-6}$ | 0.64 | $1.834 \cdot 10^{-6}$ | 2.32 | $2.661 \cdot 10^{-7}$ | 4.54 | $2.545 \cdot 10^{-4}$ | 5.51 | 7120 |
| $4^{-9}$ | $-1.170 \cdot 10^{-6}$ | – | $1.485 \cdot 10^{-6}$ | – | $3.689 \cdot 10^{-7}$ | – | $3.249 \cdot 10^{-4}$ | – | 0 |

**9.2. Dreidimensionale Probleme.** Als erstes betrachten wir erneut ein Laplace-Problem mit glatter Lösung:

Beispiel 4:

$$\begin{aligned} \Delta u(x,y,z) &= 0 && \text{in } \Omega^{(3)}, \\ u(x,y,z) &= \sin(\pi x)\cdot\sin(\pi y)\cdot\frac{\sinh(\sqrt{2}\pi z)}{\sinh(\sqrt{2}\pi)} && \text{auf } \delta\bar{\Omega}^{(3)}. \end{aligned}$$

In Abbildung 26 sind die auf dem regulären dünnen Gitter $\tilde{G}_8^{(3)}$ berechnete Näherungslösung $\tilde{u}_8^{(3)}(x,y,\frac{1}{2})$ sowie ihr Fehler $\tilde{e}_8^{(3)}(x,y,\frac{1}{2})$ beziehungsweise deren Interpolanten auf das zur graphischen Ausgabe verwendete volle Gitter $G_5^{(2)}$ dargestellt. Abbildung 27 zeigt die beiden Gitterebenen $z=\frac{1}{4}$ und $z=\frac{1}{2}$ des dreidimensionalen dünnen Gitters, welches man bei adaptiver Verfeinerung mit $\epsilon = 4^{-6}$ und (38) als Kriterium für die Knotenerzeugung erhält.

Tabelle 12 enthält nun die Ergebnisse im regulären Fall. Das feinste hier betrachtete dünne Gitter $\tilde{G}_{11}^{(3)}$ besteht nach Satz 3.3 aus 262145 Gitterpunkten. Der entsprechende Baum setzt sich gemäß der Ausführungen in Abschnitt 5.2 aus je einem Knoten für je-

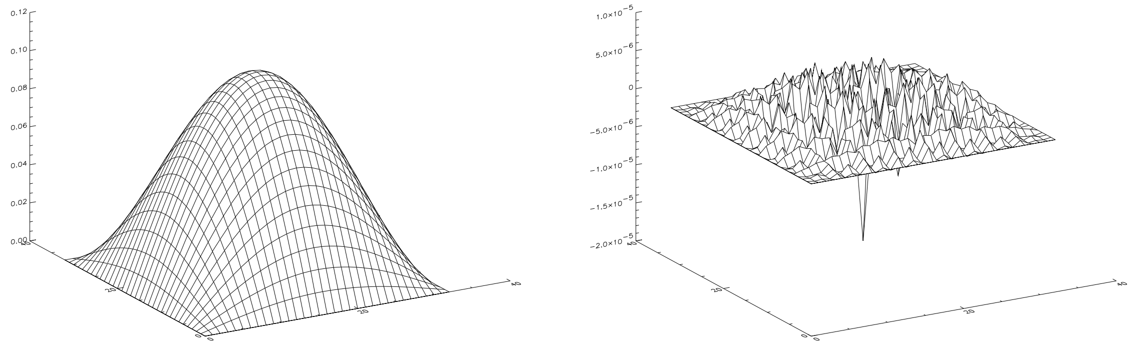



ABB. 26. *Beispiel 4: Die Lösung $\tilde{u}_8^{(3)}$ und ihr Fehler $\tilde{e}_8^{(3)}$ für $z = \frac{1}{2}$.*

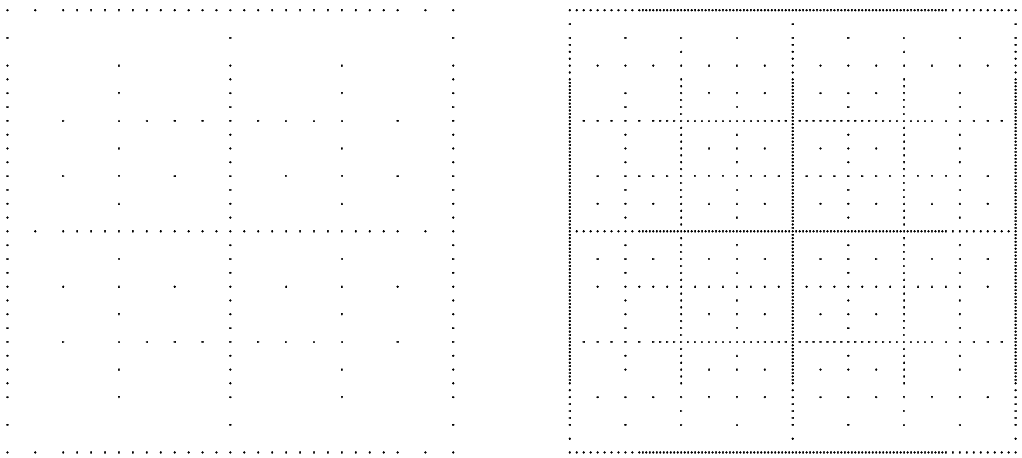


ABB. 27. *Beispiel 4: Punkte im adaptiven Gitter ($\epsilon = 4^{-6}$) für $z = \frac{1}{4}$ (links) und $z = \frac{1}{2}$ (rechts).*

den Gitterpunkt, jede Gitterzeile sowie jede Gitterebene zusammen und enthält somit 292867 Knoten. Bei dreidimensionalen Laplace-Problemen benötigt unser Verfahren in jedem Gitterpunkt acht reelle und drei ganzzahlige Variablen. Zusammen mit dem durch die Datenstruktur bedingten organisatorischen Aufwand (Zeiger) sind bei Rechnung in doppelter Genauigkeit für jeden Knoten 88 Byte erforderlich. Bei Poisson-Problemen erhöht sich diese Zahl um weitere 8 Byte für die rechte Seite der Differentialgleichung. Wir erhalten folglich für $\tilde{G}_{11}^{(3)}$ einen Gesamtspeicherbedarf von 25.8 MByte beziehungsweise 28.1 MByte gegenüber 68.8 GByte beziehungsweise 137.6 GByte für das volle Gitter $G_{11}^{(3)}$. Betrachten wir nun die Werte von $\rho_\infty$ und $\rho_2$ in Tabelle 12, so übertreffen sie — nach gewissen Anlaufschwierigkeiten bei kleinem $n$ — die Erwartungen noch. Sie liegen näher an 4.0, als wir dies aufgrund der vermuteten $h_n^2 \cdot (\mathrm{ld}(h_n^{-1}))^2$-Eigenschaft der entsprechenden Fehlernormen vermuten. Lediglich die Reduzierung des Fehlers im Mittelpunkt verläuft etwas langsamer als erhofft. Die Konvergenzgeschwindigkeit schließlich sinkt nur leicht im Vergleich zu den zweidimensionalen Beispielen. Wuchsen die Werte von $m_\delta$ beim Übergang von $n$ zu $n+1$ etwa um einen Faktor $\sqrt{2}$, so liegt die Zunahme jetzt ungefähr bei einem Faktor 1.55.

TABELLE 12
*Beispiel 4, reguläre dünne Gitter $\tilde{G}_n^{(3)}$.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\\|\tilde{e}_n^{(2)}\\|_\infty$ | $\rho_\infty$ | $\\|\tilde{e}_n^{(2)}\\|_2$ | $\rho_2$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $-1.072 \cdot 10^{-1}$ | | $1.072 \cdot 10^{-1}$ | | $1.072 \cdot 10^{-1}$ | | 2 | 39 |
| 2 | $9.691 \cdot 10^{-4}$ | 110.62 | $9.145 \cdot 10^{-2}$ | 1.17 | $3.718 \cdot 10^{-2}$ | 2.88 | 10 | 107 |
| 3 | $6.043 \cdot 10^{-4}$ | 1.60 | $4.553 \cdot 10^{-2}$ | 2.01 | $1.127 \cdot 10^{-2}$ | 3.30 | 27 | 283 |
| 4 | $-4.006 \cdot 10^{-4}$ | 1.51 | $1.664 \cdot 10^{-2}$ | 2.74 | $3.112 \cdot 10^{-3}$ | 3.62 | 51 | 723 |
| 5 | $-2.423 \cdot 10^{-4}$ | 1.65 | $5.133 \cdot 10^{-3}$ | 3.24 | $8.186 \cdot 10^{-4}$ | 3.80 | 65 | 1795 |
| 6 | $-1.069 \cdot 10^{-4}$ | 2.27 | $1.441 \cdot 10^{-3}$ | 3.56 | $2.099 \cdot 10^{-4}$ | 3.90 | 95 | 4355 |
| 7 | $-4.214 \cdot 10^{-5}$ | 2.54 | $3.946 \cdot 10^{-4}$ | 3.65 | $5.304 \cdot 10^{-5}$ | 3.96 | 149 | 10371 |
| 8 | $-1.539 \cdot 10^{-5}$ | 2.74 | $1.062 \cdot 10^{-4}$ | 3.72 | $1.330 \cdot 10^{-5}$ | 3.99 | 213 | 24323 |
| 9 | $-5.321 \cdot 10^{-6}$ | 2.89 | $2.751 \cdot 10^{-5}$ | 3.86 | $3.327 \cdot 10^{-6}$ | 4.00 | 327 | 56323 |
| 10 | $-1.764 \cdot 10^{-6}$ | 3.02 | $6.982 \cdot 10^{-6}$ | 3.94 | $8.324 \cdot 10^{-7}$ | 4.00 | 504 | 129027 |
| 11 | $-5.657 \cdot 10^{-7}$ | 3.12 | $1.859 \cdot 10^{-6}$ | 3.76 | $2.090 \cdot 10^{-7}$ | 3.98 | 781 | 292867 |

In Tabelle 13 sind die Ergebnisse bei adaptiver Verfeinerung zusammengefaßt. Als Kriterium für die Erzeugung neuer Gitterpunkte gelangt dabei (38) zur Anwendung. Um zu vermeiden, daß zu Beginn des Iterationsprozesses Knoten in den Baum eingehängt werden, die sich später als unwichtig herausstellen, wählen wir hier die bereits beschriebene sukzessive Verfeinerungsstrategie. Der Adaptionsparameter $\epsilon$ wird also erst dann weiter verringert, wenn eine ausreichende Anzahl an Iterationsschritten durchgeführt worden ist.

TABELLE 13
*Beispiel 4, adaptiv erzeugte dünne Gitter (Kriterium (38)).*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\\|\tilde{e}_\epsilon^{(2)}\\|_\infty$ | $\rho_\infty$ | $\\|\tilde{e}_\epsilon^{(2)}\\|_2$ | $\rho_2$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-9.864 \cdot 10^{-3}$ | | $4.436 \cdot 10^{-2}$ | | $1.705 \cdot 10^{-2}$ | | 147 |
| $4^{-2}$ | $1.984 \cdot 10^{-3}$ | 4.97 | $2.153 \cdot 10^{-2}$ | 2.06 | $1.107 \cdot 10^{-2}$ | 1.54 | 283 |
| $4^{-3}$ | $2.460 \cdot 10^{-4}$ | 8.07 | $6.753 \cdot 10^{-3}$ | 3.19 | $2.601 \cdot 10^{-3}$ | 4.26 | 659 |
| $4^{-4}$ | $-8.966 \cdot 10^{-5}$ | 2.74 | $2.155 \cdot 10^{-3}$ | 3.13 | $8.241 \cdot 10^{-4}$ | 3.16 | 1723 |
| $4^{-5}$ | $1.192 \cdot 10^{-6}$ | 75.22 | $7.156 \cdot 10^{-4}$ | 3.01 | $2.345 \cdot 10^{-4}$ | 3.51 | 4075 |
| $4^{-6}$ | $-1.701 \cdot 10^{-5}$ | 0.07 | $2.241 \cdot 10^{-4}$ | 3.19 | $7.057 \cdot 10^{-5}$ | 3.32 | 9227 |
| $4^{-7}$ | $-1.782 \cdot 10^{-5}$ | 0.95 | $6.360 \cdot 10^{-5}$ | 3.52 | $2.030 \cdot 10^{-5}$ | 3.48 | 21703 |
| $4^{-8}$ | $-6.621 \cdot 10^{-6}$ | 2.69 | $1.859 \cdot 10^{-5}$ | 3.42 | $5.681 \cdot 10^{-6}$ | 3.57 | 49059 |
| $4^{-9}$ | $-2.106 \cdot 10^{-6}$ | 3.14 | $5.121 \cdot 10^{-6}$ | 3.63 | $1.544 \cdot 10^{-6}$ | 3.68 | 111707 |

Bei diesem Beispiel führt der Einsatz adaptiver Verfeinerung zu keinen besseren Resultaten. Demgegenüber erweist sie sich im folgenden Problem als äußerst hilfreich.

<u>Beispiel 5:</u>

$$\begin{aligned} \Delta u(x,y,z) &= 0 && \text{in } \Omega^{(3)}, \\ u(x,y,z) &= \tfrac{\cos(4\pi(x-y))\cdot\sinh(4\pi(x+y))}{\sinh(8\pi)} \cdot z && \text{auf } \delta\bar{\Omega}^{(3)}. \end{aligned}$$

Die Lösung $u(x,y,z)$ erfüllt die Glattheitsanforderung $u \in X^{(3)}$, da wegen $\frac{\partial^2 u}{\partial z^2} \equiv 0$ auch $\frac{\partial^6 u}{\partial x^2 \partial y^2 \partial z^2} \equiv 0$ gilt. Somit ist die Funktion glatt im Sinne dieser Arbeit. Betrachtet man dagegen beispielsweise $v(x,y) := u(x,y,\frac{1}{2})$, so ist $\frac{\partial^4 v}{\partial x^2 \partial y^2}$ zwar beschränkt und stetig, nimmt jedoch nahe $(1,1)$ sehr große Werte an. Man erkennt auch in Abbildung 28, in der die Näherungslösung $\tilde{u}_8^{(3)}(x,y,\frac{1}{2})$ sowie ihr Fehler $\tilde{e}_8^{(3)}(x,y,\frac{1}{2})$ dargestellt sind, daß $u$ in der Nähe des Punktes $(1,1,\frac{1}{2})$ sehr steil ansteigt. Deshalb darf man bei adaptiver Verfeinerung erheblich verbesserte Resultate im Vergleich zu regulären dünnen Gittern erwarten. Man beachte, daß bei diesem Beispiel innere Gitterpunkte lediglich aufgrund der Tatsache erforderlich sind, daß man die Lösung $u$ nur näherungsweise berechnet. Denn für $u$ selbst sind wegen $\frac{\partial^2 u}{\partial z^2} \equiv 0$ die hierarchischen Überschüsse an sämtlichen Gitterpunkten im Inneren des Einheitswürfels Null. Abbildung 29 zeigt die beiden Gitterebenen $z = \frac{1}{4}$ und $z = \frac{1}{2}$ des dreidimensionalen dünnen Gitters, welches man bei adaptiver Verfeinerung mit $\epsilon = 4^{-5}$ und (38) als Kriterium für die Knotenerzeugung erhält.

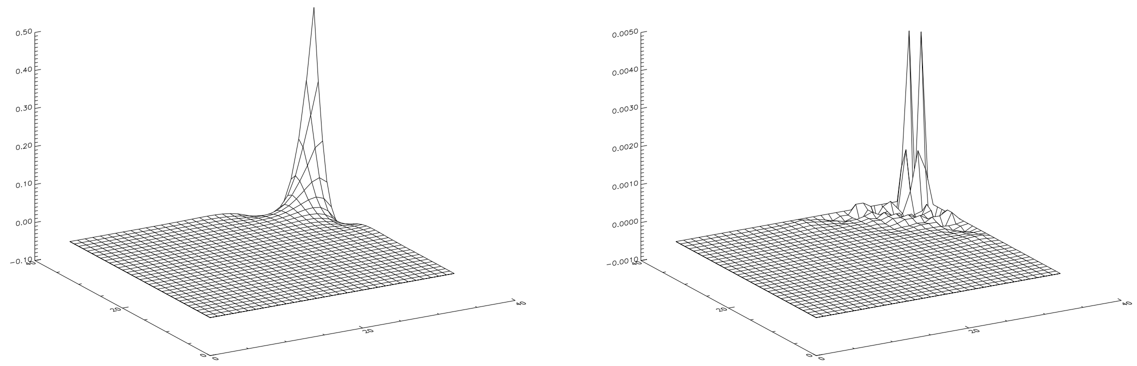



ABB. 28. *Beispiel 5: Die Lösung* $\tilde{u}_8^{(3)}$ *und ihr Fehler* $\tilde{e}_8^{(3)}$ *für* $z = \frac{1}{2}$.

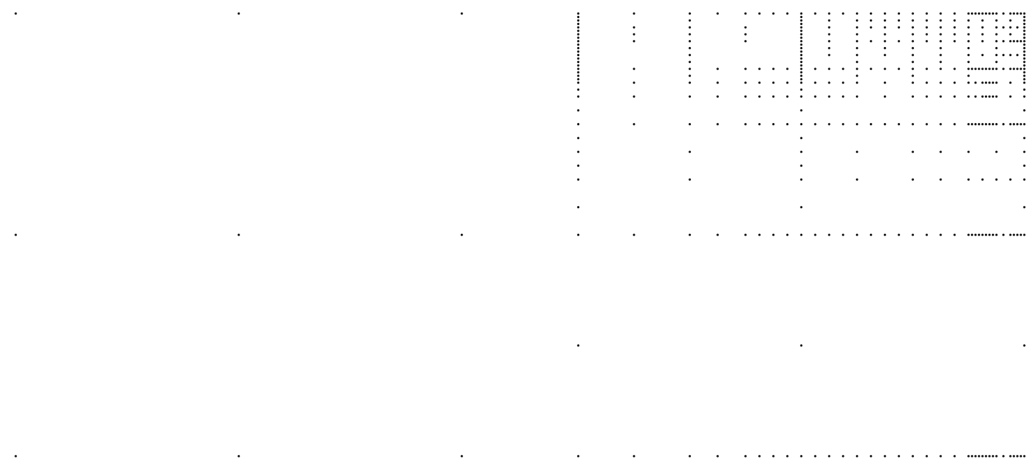


ABB. 29. *Beispiel 5: Punkte im adaptiven Gitter* ($\epsilon = 4^{-5}$) *für* $z = \frac{1}{4}$ *(links) und* $z = \frac{1}{2}$ *(rechts).*

Tabelle 14 enthält die Ergebnisse im regulären Fall, Tabelle 15 stellt die Resultate bei adaptiver Verfeinerung zusammen, wobei der Wert des Adaptionsparameters $\epsilon$

wieder schrittweise verkleinert wird. Bei Verwendung regulärer dünner Gitter sind die Quotienten $\rho_\infty$ und $\rho_2$ zunächst deutlich kleiner als 4. Erst mit wachsendem $n$ erreichen $\rho_\infty$ und $\rho_2$ wie $\rho_P$ Werte im Bereich knapp unter 4.

TABELLE 14
*Beispiel 5, reguläre dünne Gitter $\tilde{G}_n^{(3)}$.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\lVert\tilde{e}_n^{(2)}\rVert_\infty$ | $\rho_\infty$ | $\lVert\tilde{e}_n^{(2)}\rVert_2$ | $\rho_2$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | 2 | 39 |
| 2 | $-3.657 \cdot 10^{-2}$ | 1.72 | $6.744 \cdot 10^{-2}$ | 0.93 | $3.984 \cdot 10^{-2}$ | 1.57 | 9 | 107 |
| 3 | $6.846 \cdot 10^{-3}$ | 5.34 | $7.094 \cdot 10^{-2}$ | 0.95 | $2.322 \cdot 10^{-2}$ | 1.72 | 23 | 283 |
| 4 | $-1.803 \cdot 10^{-3}$ | 3.80 | $6.324 \cdot 10^{-2}$ | 1.12 | $1.036 \cdot 10^{-2}$ | 2.24 | 47 | 723 |
| 5 | $-8.675 \cdot 10^{-5}$ | 20.78 | $4.416 \cdot 10^{-2}$ | 1.43 | $3.895 \cdot 10^{-3}$ | 2.66 | 65 | 1795 |
| 6 | $-3.314 \cdot 10^{-5}$ | 2.62 | $2.426 \cdot 10^{-2}$ | 1.82 | $1.315 \cdot 10^{-3}$ | 2.96 | 94 | 4355 |
| 7 | $-8.766 \cdot 10^{-6}$ | 3.78 | $1.231 \cdot 10^{-2}$ | 1.97 | $4.091 \cdot 10^{-4}$ | 3.21 | 114 | 10371 |
| 8 | $-2.412 \cdot 10^{-6}$ | 3.63 | $4.519 \cdot 10^{-3}$ | 2.72 | $1.174 \cdot 10^{-4}$ | 3.48 | 184 | 24323 |
| 9 | $-6.648 \cdot 10^{-7}$ | 3.63 | $1.620 \cdot 10^{-3}$ | 2.79 | $3.240 \cdot 10^{-5}$ | 3.62 | 294 | 56323 |
| 10 | $-1.824 \cdot 10^{-7}$ | 3.64 | $4.954 \cdot 10^{-4}$ | 3.27 | $8.787 \cdot 10^{-6}$ | 3.69 | 468 | 129027 |
| 11 | $-4.967 \cdot 10^{-8}$ | 3.67 | $1.495 \cdot 10^{-4}$ | 3.31 | $2.383 \cdot 10^{-6}$ | 3.69 | 737 | 292867 |

TABELLE 15
*Beispiel 5, adaptiv erzeugte dünne Gitter (Kriterium (38)).*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\lVert\tilde{e}_\epsilon^{(2)}\rVert_\infty$ | $\rho_\infty$ | $\lVert\tilde{e}_\epsilon^{(2)}\rVert_2$ | $\rho_2$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | 39 |
| $4^{-2}$ | $8.968 \cdot 10^{-4}$ | 69.95 | $3.789 \cdot 10^{-2}$ | 1.66 | $1.167 \cdot 10^{-2}$ | 5.38 | 227 |
| $4^{-3}$ | $9.822 \cdot 10^{-4}$ | 0.91 | $1.151 \cdot 10^{-2}$ | 3.29 | $3.914 \cdot 10^{-3}$ | 2.98 | 377 |
| $4^{-4}$ | $4.570 \cdot 10^{-5}$ | 21.49 | $4.167 \cdot 10^{-3}$ | 2.76 | $1.429 \cdot 10^{-3}$ | 2.74 | 635 |
| $4^{-5}$ | $-7.020 \cdot 10^{-6}$ | 6.51 | $1.997 \cdot 10^{-3}$ | 2.09 | $4.473 \cdot 10^{-4}$ | 3.19 | 1379 |
| $4^{-6}$ | $-7.195 \cdot 10^{-5}$ | 0.10 | $5.121 \cdot 10^{-4}$ | 3.90 | $1.357 \cdot 10^{-4}$ | 3.30 | 2879 |
| $4^{-7}$ | $4.859 \cdot 10^{-6}$ | 14.81 | $1.294 \cdot 10^{-4}$ | 3.96 | $3.276 \cdot 10^{-5}$ | 4.14 | 6581 |
| $4^{-8}$ | $-4.045 \cdot 10^{-7}$ | 12.01 | $4.880 \cdot 10^{-5}$ | 2.65 | $8.342 \cdot 10^{-6}$ | 3.93 | 17483 |
| $4^{-9}$ | $-5.433 \cdot 10^{-8}$ | 7.45 | $1.435 \cdot 10^{-5}$ | 3.40 | $3.091 \cdot 10^{-6}$ | 2.70 | 38639 |

Vergleicht man nun die Zahlen aus Tabelle 14 mit denen aus Tabelle 15, so sieht man klar die Vorteile des Einsatzes adaptiver Verfeinerung. Beispielsweise ist der maximale Fehler bei $\epsilon = 4^{-9}$ fast um einen Faktor 100 kleiner als der auf dem regulären dünnen Gitter $\tilde{G}_9^{(3)}$ erzielte maximale Fehler, obwohl $\tilde{G}_9^{(3)}$ mehr Punkte enthält.

Als drittes dreidimensionales Beispiel betrachten wir jetzt ein Poisson-Problem.

Beispiel 6:

$$\begin{aligned} \Delta u(x,y,z) &= -3\pi^2 \cdot \sin(\pi x)\cdot\sin(\pi y)\cdot\sin(\pi z) \quad && \text{in } \Omega^{(3)}, \\ u(x,y,z) &= 0 && \text{auf } \delta\bar{\Omega}^{(3)}. \end{aligned}$$

Tabelle 16 zeigt die Resultate unseres dreidimensionalen Poisson-Beispiels im regulären Fall, wobei die Vorlaufrechnung auf dem Gitter $\tilde{G}_{11}^{(3)}$ durchgeführt wurde (d.h. $n_f = 11$). Wie zuvor in Beispiel 3 verschlechtern sich die Werte von $\rho_P$, $\rho_\infty$ und $\rho_2$, wenn das zur Berechnung der Näherungslösung herangezogene $n = n_u$ nicht mehr deutlich kleiner als $n_f$ ist. Führen wir dagegen Vorlaufrechnung und Iteration immer mit demselben Parameter $n_f = n_u$ aus, so stellt sich wiederum das erwartete Konvergenzverhalten ein (siehe Tabelle 17).

TABELLE 16

*Beispiel 6, reguläre dünne Gitter $\tilde{G}_n^{(3)}$ $(n_f = 11)$.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_n^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_n^{(2)}\|_2$ | $\rho_2$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $4.783 \cdot 10^{-1}$ | | $4.783 \cdot 10^{-1}$ | | $4.783 \cdot 10^{-1}$ | | 1 | 39 |
| 2 | $3.013 \cdot 10^{-2}$ | 15.87 | $1.067 \cdot 10^{-1}$ | 4.48 | $9.947 \cdot 10^{-2}$ | 4.81 | 9 | 107 |
| 3 | $-1.571 \cdot 10^{-2}$ | 1.92 | $2.785 \cdot 10^{-2}$ | 3.83 | $2.009 \cdot 10^{-2}$ | 4.95 | 27 | 283 |
| 4 | $-8.254 \cdot 10^{-3}$ | 1.90 | $8.254 \cdot 10^{-3}$ | 3.37 | $4.072 \cdot 10^{-3}$ | 4.93 | 51 | 723 |
| 5 | $-2.793 \cdot 10^{-3}$ | 2.96 | $2.793 \cdot 10^{-3}$ | 2.96 | $8.527 \cdot 10^{-4}$ | 4.78 | 73 | 1795 |
| 6 | $-8.037 \cdot 10^{-4}$ | 3.48 | $8.037 \cdot 10^{-4}$ | 3.48 | $1.910 \cdot 10^{-4}$ | 4.46 | 101 | 4355 |
| 7 | $-2.115 \cdot 10^{-4}$ | 3.80 | $2.283 \cdot 10^{-4}$ | 3.52 | $4.772 \cdot 10^{-5}$ | 4.00 | 152 | 10371 |
| 8 | $-5.424 \cdot 10^{-5}$ | 3.90 | $7.024 \cdot 10^{-5}$ | 3.25 | $1.420 \cdot 10^{-5}$ | 3.36 | 221 | 24323 |
| 9 | $-1.623 \cdot 10^{-5}$ | 3.34 | $2.470 \cdot 10^{-5}$ | 2.84 | $5.893 \cdot 10^{-6}$ | 2.41 | 342 | 56323 |
| 10 | $-7.951 \cdot 10^{-6}$ | 2.04 | $1.139 \cdot 10^{-5}$ | 2.17 | $3.786 \cdot 10^{-6}$ | 1.56 | 527 | 129027 |
| 11 | $-6.412 \cdot 10^{-6}$ | 1.24 | $7.650 \cdot 10^{-6}$ | 1.49 | $3.181 \cdot 10^{-6}$ | 1.19 | 812 | 292867 |

TABELLE 17

*Beispiel 6, reguläre dünne Gitter $\tilde{G}_n^{(3)}$ $(n_f = n_u)$.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_n^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_n^{(2)}\|_2$ | $\rho_2$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $1.775 \cdot 10^{-1}$ | | $1.775 \cdot 10^{-1}$ | | $1.775 \cdot 10^{-1}$ | | 1 | 39 |
| 2 | $-1.413 \cdot 10^{-1}$ | 1.26 | $1.413 \cdot 10^{-1}$ | 1.26 | $6.854 \cdot 10^{-2}$ | 2.59 | 9 | 107 |
| 3 | $-7.838 \cdot 10^{-2}$ | 1.80 | $7.838 \cdot 10^{-2}$ | 1.80 | $3.052 \cdot 10^{-2}$ | 2.25 | 27 | 283 |
| 4 | $-3.081 \cdot 10^{-2}$ | 2.54 | $3.081 \cdot 10^{-2}$ | 2.54 | $1.173 \cdot 10^{-2}$ | 2.60 | 51 | 723 |
| 5 | $-1.042 \cdot 10^{-2}$ | 2.96 | $1.042 \cdot 10^{-2}$ | 2.96 | $4.075 \cdot 10^{-3}$ | 2.88 | 73 | 1795 |
| 6 | $-3.270 \cdot 10^{-3}$ | 3.19 | $3.270 \cdot 10^{-3}$ | 3.19 | $1.328 \cdot 10^{-3}$ | 3.07 | 101 | 4355 |
| 7 | $-9.827 \cdot 10^{-4}$ | 3.33 | $9.827 \cdot 10^{-4}$ | 3.33 | $4.157 \cdot 10^{-4}$ | 3.19 | 152 | 10371 |
| 8 | $-2.869 \cdot 10^{-4}$ | 3.43 | $2.984 \cdot 10^{-4}$ | 3.29 | $1.264 \cdot 10^{-4}$ | 3.29 | 221 | 24323 |
| 9 | $-8.201 \cdot 10^{-5}$ | 3.50 | $8.865 \cdot 10^{-5}$ | 3.37 | $3.762 \cdot 10^{-5}$ | 3.36 | 342 | 56323 |
| 10 | $-2.308 \cdot 10^{-5}$ | 3.55 | $2.623 \cdot 10^{-5}$ | 3.38 | $1.101 \cdot 10^{-5}$ | 3.42 | 527 | 129027 |
| 11 | $-6.412 \cdot 10^{-6}$ | 3.60 | $7.650 \cdot 10^{-6}$ | 3.43 | $3.181 \cdot 10^{-6}$ | 3.46 | 812 | 292867 |

Tabelle 18 enthält schließlich die Ergebnisse bei adaptiver Verfeinerung mittels des Kriteriums (38). Der während der Vorlaufrechnung mit $\epsilon_f = 10^{-5}$ erzeugte Baum besteht dabei aus 238179 Knoten. Ausgehend von $\epsilon = \epsilon_u = 4^{-1}$ wird $\epsilon$ dabei schrittweise verkleinert, bis es den Wert $4^{-8} \approx 1.5 \cdot 10^{-5}$ erreicht hat. Dadurch kann die Erzeugung neuer Knoten während des Iterationsprozesses vermieden werden ($K_\epsilon^{neu} = 0$). Auch hier nehmen die Werte von $\rho_\infty$ und $\rho_2$ ab, wenn $\epsilon$ nicht mehr deutlich größer als $\epsilon_f$ ist.

TABELLE 18

*Beispiel 6, adaptiv erzeugte dünne Gitter (Kriterium (38), $\epsilon_f = 10^{-5}$).*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_\epsilon^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_\epsilon^{(2)}\|_2$ | $\rho_2$ | $K_\epsilon^{neu}$ |
|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-1.571 \cdot 10^{-2}$ | | $2.785 \cdot 10^{-2}$ | | $2.009 \cdot 10^{-2}$ | | 0 |
| $4^{-2}$ | $-2.494 \cdot 10^{-2}$ | 0.63 | $2.494 \cdot 10^{-2}$ | 1.12 | $9.619 \cdot 10^{-3}$ | 2.09 | 0 |
| $4^{-3}$ | $-1.794 \cdot 10^{-3}$ | 13.90 | $3.533 \cdot 10^{-3}$ | 7.06 | $1.290 \cdot 10^{-3}$ | 7.46 | 0 |
| $4^{-4}$ | $-6.780 \cdot 10^{-5}$ | 26.46 | $1.332 \cdot 10^{-3}$ | 2.65 | $3.866 \cdot 10^{-4}$ | 3.34 | 0 |
| $4^{-5}$ | $-2.374 \cdot 10^{-5}$ | 2.86 | $3.490 \cdot 10^{-4}$ | 3.82 | $9.536 \cdot 10^{-5}$ | 4.05 | 0 |
| $4^{-6}$ | $-6.948 \cdot 10^{-6}$ | 3.42 | $1.140 \cdot 10^{-4}$ | 3.06 | $3.065 \cdot 10^{-5}$ | 3.11 | 0 |
| $4^{-7}$ | $-3.046 \cdot 10^{-6}$ | 2.28 | $3.587 \cdot 10^{-5}$ | 3.18 | $1.076 \cdot 10^{-5}$ | 2.85 | 0 |
| $4^{-8}$ | $-3.651 \cdot 10^{-6}$ | 0.83 | $1.240 \cdot 10^{-5}$ | 2.89 | $5.583 \cdot 10^{-6}$ | 1.93 | 0 |

**9.3. Beispiele zum energienormorientierten dünnen Gitter.** Nun wenden wir uns den energieorientierten dünnen Gittern gemäß Abschnitt 4 zu. Abbildung 30 stellt die beiden regulären dünnen Gitter $\tilde{G}_7^{(2)}$ und $\hat{G}_7^{(2)}$ gegenüber. Deutlich erkennt man, daß $\hat{G}_7^{(2)}$ weniger Gitterpunkte enthält. Die im Vergleich zu $\tilde{G}_7^{(2)}$ bei $\hat{G}_7^{(2)}$ weggefallenen Punkte entsprechen nach (29) Basisfunktionen aus den Teilräumen $T_{3,5}$, $T_{4,4}$ sowie $T_{5,3}$.

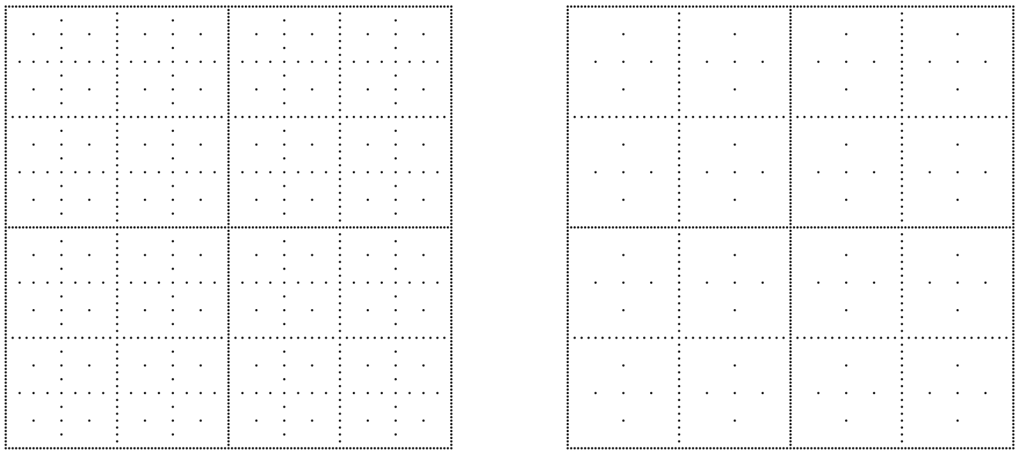


ABB. 30. *Reguläre dünne Gitter $\tilde{G}_7^{(2)}$ (links) und $\hat{G}_7^{(2)}$ (rechts).*

Wir betrachten zunächst wieder Beispiel 1. Abbildung 31 zeigt auf dem vollen Gitter $G_5^{(2)}$ die Fehler $\tilde{e}_{10}^{(2)}$ und $\hat{e}_{10}^{(2)}$ der auf den dünnen Gittern $\tilde{G}_{10}^{(2)}$ beziehungsweise $\hat{G}_{10}^{(2)}$ berechneten Näherungslösungen $\tilde{u}_{10}^{(2)}$ beziehungsweise $\hat{u}_{10}^{(2)}$. Die beiden Fehler weisen dabei eine sehr ähnliche Gestalt auf. Sie sind jedoch in Abbildung 31 unterschiedlich skaliert, da $\hat{e}_{10}^{(2)}$ aufgrund der geringeren Anzahl an Gitterpunkten des Gitters $\hat{G}_{10}^{(2)}$ mit einem größeren maximalen Fehler behaftet ist als $\tilde{e}_{10}^{(2)}$. Es gilt $\|\tilde{e}_{10}^{(2)}\|_\infty = 2.925 \cdot 10^{-6}$ und $\|\hat{e}_{10}^{(2)}\|_\infty = 1.415 \cdot 10^{-5}$.

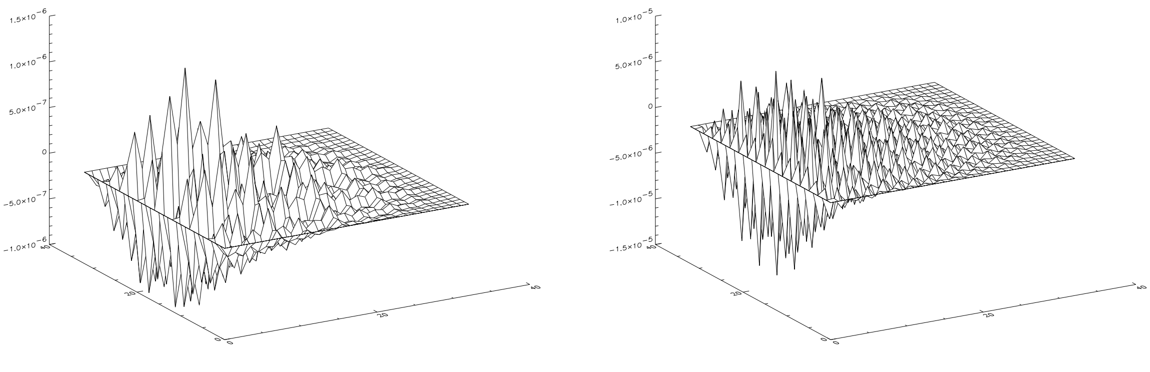



ABB. 31. *Beispiel 1: Die Fehler $\tilde{e}_{10}^{(2)}$ (links) und $\hat{e}_{10}^{(2)}$ (rechts).*

Tabelle 19
*Beispiel 1, reguläre dünne Gitter $\hat{G}_n^{(2)}$.*

| $n$ | $\hat{e}_n^{(2)}(P)$ | $\rho_P$ | $\Vert\hat{e}_n^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\hat{e}_n^{(2)}\Vert_2$ | $\rho_2$ | $\Vert\hat{e}_n^{(2)}\Vert_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $-7.427 \cdot 10^{-2}$ | | $7.427 \cdot 10^{-2}$ | | $7.427 \cdot 10^{-2}$ | | $1.213 \cdot 10^{-1}$ | | 1 | 12 |
| 2 | $-5.917 \cdot 10^{-3}$ | 12.55 | $3.502 \cdot 10^{-2}$ | 2.12 | $2.018 \cdot 10^{-2}$ | 3.68 | $6.716 \cdot 10^{-2}$ | 1.81 | 7 | 26 |
| 3 | $5.756 \cdot 10^{-4}$ | 10.28 | $1.352 \cdot 10^{-2}$ | 2.59 | $5.461 \cdot 10^{-3}$ | 3.70 | $3.649 \cdot 10^{-2}$ | 1.84 | 22 | 58 |
| 4 | $4.739 \cdot 10^{-4}$ | 1.21 | $4.528 \cdot 10^{-3}$ | 2.99 | $1.449 \cdot 10^{-3}$ | 3.77 | $1.879 \cdot 10^{-2}$ | 1.94 | 26 | 130 |
| 5 | $4.155 \cdot 10^{-4}$ | 1.14 | $2.556 \cdot 10^{-3}$ | 1.77 | $6.414 \cdot 10^{-4}$ | 2.26 | $1.206 \cdot 10^{-2}$ | 1.56 | 33 | 274 |
| 6 | $1.930 \cdot 10^{-4}$ | 2.15 | $9.649 \cdot 10^{-4}$ | 2.65 | $2.086 \cdot 10^{-4}$ | 3.07 | $6.266 \cdot 10^{-3}$ | 1.92 | 37 | 578 |
| 7 | $7.207 \cdot 10^{-5}$ | 2.68 | $3.255 \cdot 10^{-4}$ | 2.96 | $6.068 \cdot 10^{-5}$ | 3.44 | $3.151 \cdot 10^{-3}$ | 1.99 | 49 | 1218 |
| 8 | $2.410 \cdot 10^{-5}$ | 2.99 | $9.923 \cdot 10^{-5}$ | 3.28 | $1.680 \cdot 10^{-5}$ | 3.61 | $1.576 \cdot 10^{-3}$ | 2.00 | 65 | 2562 |
| 9 | $7.561 \cdot 10^{-6}$ | 3.19 | $2.974 \cdot 10^{-5}$ | 3.34 | $4.524 \cdot 10^{-6}$ | 3.71 | $7.875 \cdot 10^{-4}$ | 2.00 | 83 | 5378 |
| 10 | $3.433 \cdot 10^{-6}$ | 2.20 | $1.415 \cdot 10^{-5}$ | 2.10 | $1.910 \cdot 10^{-6}$ | 2.37 | $4.054 \cdot 10^{-4}$ | 1.94 | 94 | 11010 |
| 11 | $1.244 \cdot 10^{-6}$ | 2.76 | $4.956 \cdot 10^{-6}$ | 2.86 | $6.107 \cdot 10^{-7}$ | 3.13 | $1.555 \cdot 10^{-4}$ | 2.61 | 118 | 22530 |
| 12 | $4.074 \cdot 10^{-7}$ | 3.05 | $1.546 \cdot 10^{-6}$ | 3.21 | $1.783 \cdot 10^{-7}$ | 3.43 | $7.202 \cdot 10^{-5}$ | 2.16 | 155 | 46082 |
| 13 | $1.259 \cdot 10^{-7}$ | 3.24 | $4.505 \cdot 10^{-7}$ | 3.43 | $4.988 \cdot 10^{-8}$ | 3.57 | $2.035 \cdot 10^{-5}$ | 3.54 | 197 | 94210 |
| 14 | $3.751 \cdot 10^{-8}$ | 3.36 | $1.373 \cdot 10^{-7}$ | 3.28 | $1.362 \cdot 10^{-8}$ | 3.66 | $7.660 \cdot 10^{-6}$ | 2.66 | 256 | 192514 |

Tabelle 20
*Beispiel 1, adaptiv erzeugte dünne Gitter (Kriterium (39)).*

| $\epsilon$ | $\hat{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\Vert\hat{e}_\epsilon^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\hat{e}_\epsilon^{(2)}\Vert_2$ | $\rho_2$ | $\Vert\hat{e}_\epsilon^{(2)}\Vert_{E_{10}}$ | $\rho_E$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-7.427 \cdot 10^{-2}$ | | $7.427 \cdot 10^{-2}$ | | $7.427 \cdot 10^{-2}$ | | $5.589 \cdot 10^{-1}$ | | 12 |
| $4^{-2}$ | $3.498 \cdot 10^{-4}$ | 212.32 | $2.603 \cdot 10^{-2}$ | 2.85 | $1.730 \cdot 10^{-2}$ | 4.29 | $2.364 \cdot 10^{-1}$ | 2.36 | 32 |
| $4^{-3}$ | $9.698 \cdot 10^{-4}$ | 0.36 | $1.213 \cdot 10^{-2}$ | 2.15 | $5.660 \cdot 10^{-3}$ | 3.06 | $1.237 \cdot 10^{-1}$ | 1.91 | 64 |
| $4^{-4}$ | $1.100 \cdot 10^{-3}$ | 0.88 | $4.221 \cdot 10^{-3}$ | 2.87 | $1.609 \cdot 10^{-3}$ | 3.52 | $6.052 \cdot 10^{-2}$ | 2.04 | 138 |
| $4^{-5}$ | $5.019 \cdot 10^{-4}$ | 2.19 | $1.324 \cdot 10^{-3}$ | 3.19 | $4.579 \cdot 10^{-4}$ | 3.51 | $3.101 \cdot 10^{-2}$ | 1.95 | 278 |
| $4^{-6}$ | $2.279 \cdot 10^{-4}$ | 2.20 | $4.144 \cdot 10^{-4}$ | 3.19 | $1.428 \cdot 10^{-4}$ | 3.21 | $1.534 \cdot 10^{-2}$ | 2.02 | 586 |
| $4^{-7}$ | $5.194 \cdot 10^{-5}$ | 4.39 | $1.194 \cdot 10^{-4}$ | 3.47 | $3.448 \cdot 10^{-5}$ | 4.14 | $7.635 \cdot 10^{-3}$ | 2.01 | 1246 |
| $4^{-8}$ | $1.223 \cdot 10^{-5}$ | 4.25 | $3.624 \cdot 10^{-5}$ | 3.29 | $1.004 \cdot 10^{-5}$ | 3.43 | $3.692 \cdot 10^{-3}$ | 2.07 | 2620 |
| $4^{-9}$ | $3.154 \cdot 10^{-6}$ | 3.88 | $9.994 \cdot 10^{-6}$ | 3.63 | $2.814 \cdot 10^{-6}$ | 3.57 | $1.584 \cdot 10^{-3}$ | 2.33 | 5524 |
| $4^{-9}$ | $8.696 \cdot 10^{-6}$ | – | $1.196 \cdot 10^{-5}$ | – | $3.356 \cdot 10^{-6}$ | – | $1.589 \cdot 10^{-3}$ | – | 5272 |

Tabelle 19 zeigt die Ergebnisse für reguläre dünne Gitter $\hat{G}_n^{(2)}$, Tabelle 20 stellt die Resultate im Falle adaptiver Verfeinerung zusammen, wobei als Kriterium für die Erzeugung weiterer Gitterpunkte jetzt (39) herangezogen wird. Die letzte Zeile von Tabelle 20 enthält dabei wieder die Werte, die sich bei schrittweiser Verkleinerung des Wertes von $\epsilon$ ergeben. Vergleicht man nun die Ergebnisse hier mit denen bei Verwendung $L_2$-normorientierter Gitter (siehe die Tabellen 6 und 7), so erkennt man deutlich die geringere Anzahl an Gitterpunkten $K_n$ beziehungsweise $K_\epsilon$. Ferner ist die Reduktion des Fehlers im Mittelpunkt sowie seiner $L_2$- und Maximumsnorm bei steigendem $n$ beziehungsweise abnehmendem $\epsilon$ jetzt langsamer, was zu kleineren Werten von $\rho_P$, $\rho_\infty$ und $\rho_2$ führt. Dieses Verhalten überrascht jedoch nicht, da wir bereits am Ende von Abschnitt 4 gesehen haben, daß auf $\hat{G}_n^{(2)}$ die $L_2$- und die Maximumsnorm des Interpolationsfehlers von der Ordnung $O(h_n^{1.6} \cdot \mathrm{ld}(h_n^{-0.8}))$ sind. Dagegen nimmt die Energienorm des Fehlers im regulären Fall nach wie vor linear ab. Zur Erläuterung der Werte von $\rho_E$ im Falle $n > 10$ sei wiederum darauf verwiesen, daß wir die Energie jeweils auf dem vollen Gitter $G_m^{(2)}$, $m = \min(n, 10)$, berechnen. Was die Konvergenzgeschwindigkeit angeht, die ja mit Hilfe der Größe $m_\delta$ beschrieben wird, so scheinen die energieorientierten dünnen Gitter leichte Vorteile zu besitzen. Wesentliche Abweichungen vom Verhalten bei Verwendung $L_2$-normorientierter dünner Gitter sind jedoch nicht auszumachen.

Für Beispiel 2 zeigt Abbildung 32 die Fehler $\tilde{e}_{10}^{(2)}$ und $\hat{e}_{10}^{(2)}$, wobei die beiden Funktionen wie zuvor von ähnlicher Gestalt sind. Auch hier sind die Fehler jedoch wieder unterschiedlich skaliert: Es gilt $\|\tilde{e}_{10}^{(2)}\|_\infty = 2.289 \cdot 10^{-2}$ und $\|\hat{e}_{10}^{(2)}\|_\infty = 2.909 \cdot 10^{-2}$.

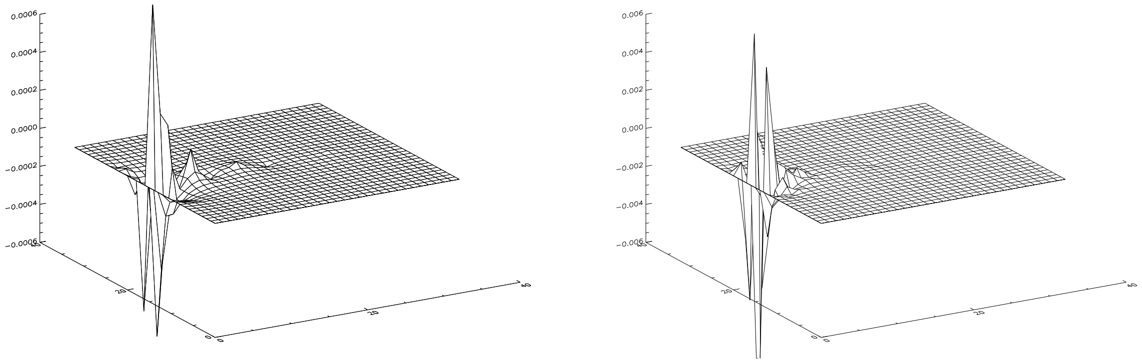


ABB. 32. *Beispiel 2: Die Fehler $\tilde{e}_{10}^{(2)}$ (links) und $\hat{e}_{10}^{(2)}$ (rechts).*

In Abbildung 33 sieht man die unterschiedliche Wirkung der beiden Kriterien (38) und (39) bei der Erzeugung neuer Gitterpunkte im Falle adaptiver Verfeinerung. Erwartungsgemäß enthält das energieorientierte Gitter weniger Gitterpunkte. Dies fällt besonders auf der Achse $y = 0$ in der Nähe der Singularität auf. An den Stellen, an denen die Lösung besonders steil verläuft (auf Strecken parallel zur $y$-Achse für kleine positive $y$, vgl. Abbildung 22), ist die Auflösung bei beiden Gittern dagegen nahezu gleich.

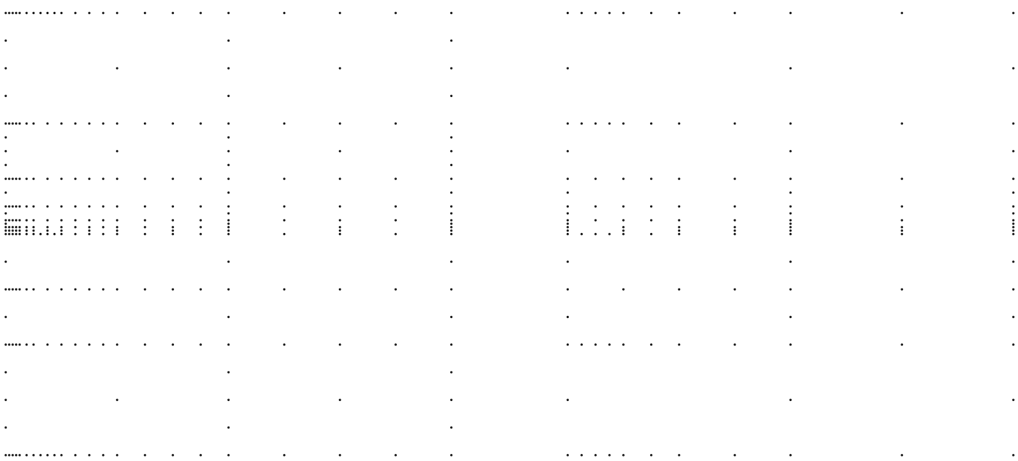


ABB. 33. *Beispiel 2: Adaptive dünne Gitter gemäß (38) ($\epsilon = 4^{-4}$, links) und (39) ($\epsilon = 4^{-4}$, rechts).*

Tabelle 21 zeigt die Resultate für Beispiel 2 im Falle regulärer dünner Gitter $\hat{G}_n^{(2)}$, Tabelle 22 enthält die Ergebnisse bei adaptiver Verfeinerung (Kriterium (39)). Der Vergleich mit den entsprechenden Resultaten bei Verwendung $L_2$-normorientierter dünner Gitter (siehe die Tabellen 8 und 9) läßt auch in diesem Beispiel wieder eine jeweils geringere Knotenanzahl bei leicht verschlechterten Werten von $\rho_2$ und $\rho_\infty$ erkennen.

Tabelle 21

*Beispiel 2, reguläre dünne Gitter $\hat{G}_n^{(2)}$.*

| $n$ | $\hat{e}_n^{(2)}(P)$ | $\rho_P$ | $\\|\hat{e}_n^{(2)}\\|_\infty$ | $\rho_\infty$ | $\\|\hat{e}_n^{(2)}\\|_2$ | $\rho_2$ | $\\|\hat{e}_n^{(2)}\\|_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $-3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $6.482 \cdot 10^{-3}$ | | 1 | 12 |
| 2 | $8.294 \cdot 10^{-3}$ | 0.48 | $3.987 \cdot 10^{-2}$ | 0.10 | $2.215 \cdot 10^{-2}$ | 0.18 | $9.846 \cdot 10^{-2}$ | 0.07 | 7 | 26 |
| 3 | $4.830 \cdot 10^{-3}$ | 1.72 | $5.668 \cdot 10^{-2}$ | 0.70 | $1.625 \cdot 10^{-2}$ | 1.36 | $1.469 \cdot 10^{-1}$ | 0.67 | 23 | 58 |
| 4 | $9.807 \cdot 10^{-4}$ | 4.93 | $5.907 \cdot 10^{-2}$ | 0.96 | $1.062 \cdot 10^{-2}$ | 1.53 | $1.772 \cdot 10^{-1}$ | 0.83 | 25 | 130 |
| 5 | $-2.222 \cdot 10^{-3}$ | 0.44 | $5.842 \cdot 10^{-2}$ | 1.01 | $9.239 \cdot 10^{-3}$ | 1.15 | $2.114 \cdot 10^{-1}$ | 0.84 | 32 | 274 |
| 6 | $-1.823 \cdot 10^{-3}$ | 1.22 | $5.046 \cdot 10^{-2}$ | 1.16 | $6.748 \cdot 10^{-3}$ | 1.37 | $2.281 \cdot 10^{-1}$ | 0.93 | 37 | 578 |
| 7 | $-1.023 \cdot 10^{-3}$ | 1.78 | $4.474 \cdot 10^{-2}$ | 1.13 | $4.635 \cdot 10^{-3}$ | 1.46 | $2.369 \cdot 10^{-1}$ | 0.96 | 45 | 1218 |
| 8 | $-5.260 \cdot 10^{-4}$ | 1.94 | $3.843 \cdot 10^{-2}$ | 1.16 | $3.116 \cdot 10^{-3}$ | 1.49 | $2.419 \cdot 10^{-1}$ | 0.98 | 57 | 2562 |
| 9 | $-2.420 \cdot 10^{-4}$ | 2.17 | $3.234 \cdot 10^{-2}$ | 1.19 | $2.100 \cdot 10^{-3}$ | 1.48 | $2.449 \cdot 10^{-1}$ | 0.99 | 79 | 5378 |
| 10 | $-1.003 \cdot 10^{-4}$ | 2.41 | $2.909 \cdot 10^{-2}$ | 1.11 | $1.806 \cdot 10^{-3}$ | 1.16 | $2.500 \cdot 10^{-1}$ | 0.98 | 89 | 11010 |
| 11 | $-4.124 \cdot 10^{-5}$ | 2.43 | $2.584 \cdot 10^{-2}$ | 1.13 | $1.358 \cdot 10^{-3}$ | 1.33 | $1.961 \cdot 10^{-1}$ | 1.27 | 110 | 22530 |
| 12 | $-1.666 \cdot 10^{-5}$ | 2.48 | $2.249 \cdot 10^{-2}$ | 1.15 | $9.658 \cdot 10^{-4}$ | 1.41 | $1.632 \cdot 10^{-1}$ | 1.20 | 148 | 46082 |

Tabelle 22

*Beispiel 2, adaptiv erzeugte dünne Gitter (Kriterium (39)).*

| $\epsilon$ | $\hat{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\\|\hat{e}_\epsilon^{(2)}\\|_\infty$ | $\rho_\infty$ | $\\|\hat{e}_\epsilon^{(2)}\\|_2$ | $\rho_2$ | $\\|\hat{e}_\epsilon^{(2)}\\|_{E_{10}}$ | $\rho_E$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $3.969 \cdot 10^{-3}$ | | $2.874 \cdot 10^{-1}$ | | 12 |
| $4^{-2}$ | $8.294 \cdot 10^{-3}$ | 0.48 | $3.987 \cdot 10^{-2}$ | 0.10 | $2.215 \cdot 10^{-2}$ | 0.18 | $2.353 \cdot 10^{-1}$ | 1.22 | 26 |
| $4^{-3}$ | $5.248 \cdot 10^{-3}$ | 1.58 | $6.045 \cdot 10^{-2}$ | 0.66 | $1.541 \cdot 10^{-2}$ | 1.44 | $2.573 \cdot 10^{-1}$ | 0.91 | 82 |
| $4^{-4}$ | $9.472 \cdot 10^{-4}$ | 5.54 | $3.754 \cdot 10^{-2}$ | 1.61 | $5.410 \cdot 10^{-3}$ | 2.85 | $2.693 \cdot 10^{-1}$ | 0.96 | 174 |
| $4^{-5}$ | $2.338 \cdot 10^{-4}$ | 4.05 | $2.691 \cdot 10^{-2}$ | 1.40 | $2.691 \cdot 10^{-3}$ | 2.01 | $1.354 \cdot 10^{-1}$ | 1.99 | 408 |
| $4^{-6}$ | $4.730 \cdot 10^{-5}$ | 4.94 | $1.706 \cdot 10^{-2}$ | 1.58 | $1.399 \cdot 10^{-3}$ | 1.92 | $6.791 \cdot 10^{-2}$ | 1.99 | 886 |
| $4^{-7}$ | $1.026 \cdot 10^{-5}$ | 4.61 | $1.141 \cdot 10^{-2}$ | 1.50 | $6.400 \cdot 10^{-4}$ | 2.19 | $3.429 \cdot 10^{-2}$ | 1.98 | 1868 |
| $4^{-8}$ | $1.590 \cdot 10^{-6}$ | 6.45 | $7.638 \cdot 10^{-3}$ | 1.49 | $3.756 \cdot 10^{-4}$ | 1.70 | $1.738 \cdot 10^{-2}$ | 1.97 | 4102 |
| $4^{-9}$ | $3.038 \cdot 10^{-7}$ | 5.23 | $4.815 \cdot 10^{-3}$ | 1.59 | $1.717 \cdot 10^{-4}$ | 2.19 | $9.040 \cdot 10^{-3}$ | 1.92 | 8982 |
| $4^{-9}$ | $-4.525 \cdot 10^{-6}$ | – | $4.814 \cdot 10^{-3}$ | – | $2.043 \cdot 10^{-4}$ | – | $9.064 \cdot 10^{-3}$ | – | 7190 |

Tabelle 23

*Beispiel 4, adaptiv erzeugte dünne Gitter (Kriterium (39)).*

| $\epsilon$ | $\hat{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\\|\hat{e}_\epsilon^{(2)}\\|_\infty$ | $\rho_\infty$ | $\\|\hat{e}_\epsilon^{(2)}\\|_2$ | $\rho_2$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-1.072 \cdot 10^{-1}$ | | $1.072 \cdot 10^{-1}$ | | $1.072 \cdot 10^{-1}$ | | 39 |
| $4^{-2}$ | $-9.864 \cdot 10^{-3}$ | 10.87 | $4.436 \cdot 10^{-2}$ | 2.42 | $1.705 \cdot 10^{-2}$ | 6.29 | 147 |
| $4^{-3}$ | $1.984 \cdot 10^{-3}$ | 4.97 | $2.153 \cdot 10^{-2}$ | 2.06 | $1.107 \cdot 10^{-2}$ | 1.54 | 283 |
| $4^{-4}$ | $1.436 \cdot 10^{-4}$ | 13.82 | $1.333 \cdot 10^{-2}$ | 1.62 | $5.214 \cdot 10^{-3}$ | 2.12 | 571 |
| $4^{-5}$ | $2.345 \cdot 10^{-4}$ | 0.61 | $5.135 \cdot 10^{-3}$ | 2.60 | $1.901 \cdot 10^{-3}$ | 2.74 | 1455 |
| $4^{-6}$ | $-1.903 \cdot 10^{-4}$ | 1.23 | $1.704 \cdot 10^{-3}$ | 3.01 | $5.821 \cdot 10^{-4}$ | 3.27 | 3287 |
| $4^{-7}$ | $-1.724 \cdot 10^{-5}$ | 11.04 | $5.793 \cdot 10^{-4}$ | 2.94 | $1.716 \cdot 10^{-4}$ | 3.39 | 7039 |
| $4^{-8}$ | $-6.138 \cdot 10^{-5}$ | 0.28 | $1.833 \cdot 10^{-4}$ | 3.16 | $5.846 \cdot 10^{-5}$ | 2.94 | 15195 |
| $4^{-9}$ | $-2.951 \cdot 10^{-5}$ | 2.08 | $5.792 \cdot 10^{-5}$ | 3.16 | $1.896 \cdot 10^{-5}$ | 3.08 | 32431 |

Auch im dreidimensionalen Fall stellen sich bei Verwendung energieorientierter dünner Gitter die erwarteten Resultate ein: Die jeweiligen Gitter enthalten weniger

Punkte als die entsprechenden $L_2$-normorientierten dünnen Gitter und führen zu schlechteren Werten von $\rho_2$ sowie $\rho_\infty$. Tabelle 23 faßt die Ergebnisse für Beispiel 4 bei adaptiver Verfeinerung mittels des energieorientierten Kriteriums (39) zusammen.

Abbildung 34 zeigt für Beispiel 5 die unterschiedliche Wirkung der beiden Adaptionsstrategien gemäß (38) und (39). Obwohl für das energieorientierte Gitter ein kleinerer Wert für den Adaptionsparameter $\epsilon$ gewählt wurde, enthält das resultierende dreidimensionale dünne Gitter auf der Mittelebene $z = \frac{1}{2}$ deutlich weniger Punkte als das entsprechende $L_2$-normorientierte Gitter.

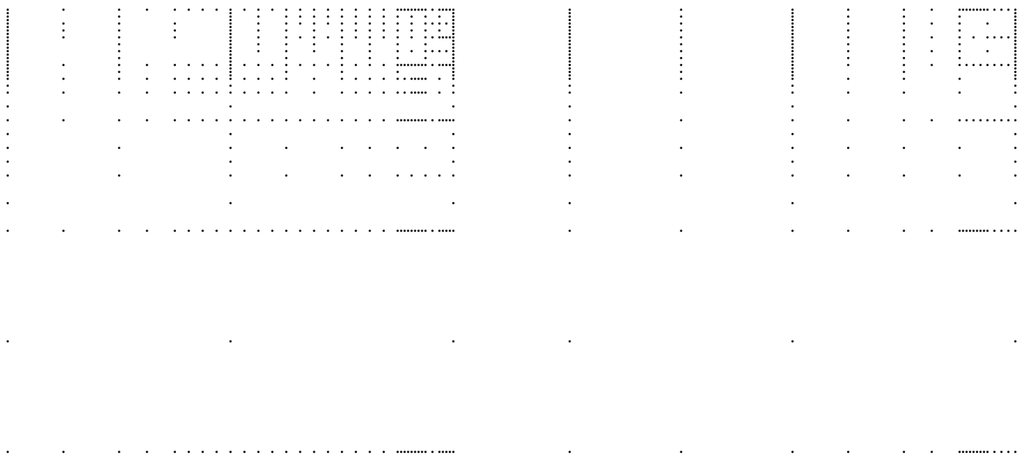


ABB. 34. *Beispiel 5, $z = \frac{1}{2}$: Adaption gemäß (38) ($\epsilon = 4^{-5}$, links) und (39) ($\epsilon = 4^{-6}$, rechts).*

Tabelle 24 enthält die Ergebnisse zu Beispiel 5 bei Verwendung regulärer dünner Gitter $\hat{G}_n^{(3)}$, Tabelle 25 zeigt die Resultate im Falle adaptiver Verfeinerung gemäß (39).

TABELLE 24
*Beispiel 5, reguläre dünne Gitter $\hat{G}_n^{(3)}$.*

| $n$ | $\hat{e}_n^{(2)}(P)$ | $\rho_P$ | $\lVert\hat{e}_n^{(2)}\rVert_\infty$ | $\rho_\infty$ | $\lVert\hat{e}_n^{(2)}\rVert_2$ | $\rho_2$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | 2 | 39 |
| 2 | $-3.657 \cdot 10^{-2}$ | 1.72 | $6.744 \cdot 10^{-2}$ | 0.93 | $3.984 \cdot 10^{-2}$ | 1.57 | 9 | 107 |
| 3 | $6.846 \cdot 10^{-3}$ | 5.34 | $7.094 \cdot 10^{-2}$ | 0.95 | $2.322 \cdot 10^{-2}$ | 1.72 | 27 | 283 |
| 4 | $-6.907 \cdot 10^{-4}$ | 9.91 | $6.524 \cdot 10^{-2}$ | 1.09 | $1.078 \cdot 10^{-2}$ | 2.15 | 35 | 715 |
| 5 | $-1.752 \cdot 10^{-3}$ | 0.39 | $6.147 \cdot 10^{-2}$ | 1.06 | $8.610 \cdot 10^{-3}$ | 1.25 | 51 | 1587 |
| 6 | $-1.252 \cdot 10^{-4}$ | 13.99 | $4.345 \cdot 10^{-2}$ | 1.41 | $3.449 \cdot 10^{-3}$ | 2.50 | 66 | 3523 |
| 7 | $-3.401 \cdot 10^{-5}$ | 3.68 | $2.410 \cdot 10^{-2}$ | 1.80 | $1.115 \cdot 10^{-3}$ | 3.09 | 99 | 7811 |
| 8 | $-8.841 \cdot 10^{-6}$ | 3.85 | $1.227 \cdot 10^{-2}$ | 1.96 | $3.606 \cdot 10^{-4}$ | 3.09 | 144 | 17219 |
| 9 | $-2.500 \cdot 10^{-6}$ | 3.54 | $5.004 \cdot 10^{-3}$ | 2.45 | $1.114 \cdot 10^{-4}$ | 3.24 | 186 | 37379 |
| 10 | $-9.268 \cdot 10^{-7}$ | 2.70 | $2.989 \cdot 10^{-3}$ | 1.67 | $5.952 \cdot 10^{-5}$ | 1.87 | 251 | 78339 |
| 11 | $-2.954 \cdot 10^{-7}$ | 3.14 | $1.485 \cdot 10^{-3}$ | 2.01 | $1.963 \cdot 10^{-5}$ | 3.03 | 342 | 164355 |

Faßt man nun die numerischen Resultate zu den gemäß Abschnitt 4 auf der Grundlage der Energienorm definierten dünnen Gittern zusammen, so zeigen sich keine wesentlichen Vorteile gegenüber den $L_2$-normorientierten dünnen Gittern. Bei gleichem $n$

TABELLE 25
*Beispiel 5, adaptiv erzeugte dünne Gitter (Kriterium (39)).*

| $\epsilon$ | $\hat{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\Vert\hat{e}_\epsilon^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\hat{e}_\epsilon^{(2)}\Vert_2$ | $\rho_2$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | 39 |
| | | 1.00 | | 1.00 | | 1.00 | |
| $4^{-2}$ | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | $6.273 \cdot 10^{-2}$ | | 39 |
| | | 37.12 | | 0.76 | | 1.95 | |
| $4^{-3}$ | $-1.690 \cdot 10^{-3}$ | | $8.257 \cdot 10^{-2}$ | | $3.217 \cdot 10^{-2}$ | | 203 |
| | | 1.96 | | 2.83 | | 4.19 | |
| $4^{-4}$ | $8.620 \cdot 10^{-4}$ | | $2.915 \cdot 10^{-2}$ | | $7.671 \cdot 10^{-3}$ | | 281 |
| | | 7.70 | | 1.89 | | 2.24 | |
| $4^{-5}$ | $-1.119 \cdot 10^{-4}$ | | $1.544 \cdot 10^{-2}$ | | $3.422 \cdot 10^{-3}$ | | 479 |
| | | 15.72 | | 2.82 | | 2.72 | |
| $4^{-6}$ | $-7.118 \cdot 10^{-6}$ | | $5.470 \cdot 10^{-3}$ | | $1.258 \cdot 10^{-3}$ | | 809 |
| | | 0.09 | | 2.38 | | 2.66 | |
| $4^{-7}$ | $-7.905 \cdot 10^{-5}$ | | $2.303 \cdot 10^{-3}$ | | $4.735 \cdot 10^{-4}$ | | 1631 |
| | | 16.57 | | 3.45 | | 2.63 | |
| $4^{-8}$ | $4.770 \cdot 10^{-6}$ | | $6.672 \cdot 10^{-4}$ | | $1.799 \cdot 10^{-4}$ | | 3521 |
| | | 1.43 | | 1.90 | | 2.95 | |
| $4^{-9}$ | $3.335 \cdot 10^{-6}$ | | $3.515 \cdot 10^{-4}$ | | $6.107 \cdot 10^{-5}$ | | 7211 |
| | | 7.39 | | 3.71 | | 2.91 | |
| $4^{-10}$ | $4.510 \cdot 10^{-7}$ | | $9.484 \cdot 10^{-5}$ | | $2.100 \cdot 10^{-5}$ | | 14927 |

beziehungsweise $\epsilon$ entstehen zwar Gitter mit erheblich weniger Knoten, die Genauigkeit der erzielten Näherungslösungen jedoch verschlechtert sich. Vergleicht man dagegen Gitter der beiden Typen mit etwa gleicher Knotenzahl, so sind die jeweils gewonnenen Lösungen auch von annähernd gleicher Qualität. Der entscheidende Vorteil des Energiekonzepts liegt also eher in den dadurch möglichen Aussagen zur Genauigkeit der auf dünnen Gittern berechneten Näherungslösungen (siehe Satz 6.2 sowie Satz 7.1).

**9.4. Numerische Resultate zu den Verfahren höherer Ordnung.** Wir betrachten jetzt wieder auf der Grundlage der $L_2$-Norm definierte dünne Gitter gemäß Abschnitt 3 und die entsprechende Adaptionsstrategie (38). Die hierfür behandelten Beispiele haben gezeigt, daß sich die $L_2$- und die Maximumsnorm des Diskretisierungsfehlers unserer Näherungslösung wie die jeweiligen Normen des Interpolationsfehlers verhalten. In Abschnitt 8.2 haben wir verschiedene Ansätze diskutiert, diese Fehlerordnung $O(h_n^2 \cdot \mathrm{ld}(h_n^{-1}))$ zu erhöhen. Zunächst wird der Baum in Gedanken zu einem Baum unendlicher Tiefe erweitert. Die neuen (fiktiven) Knoten erben dabei als hierarchischen Überschuß jeweils ein Viertel des hierarchischen Überschusses ihres Vaters im Baum. Diese Vorgehensweise bezeichnen wir im folgenden als Variante I. Unter Variante II soll derjenige Ansatz verstanden werden, der zusätzlich noch den Einsatz quadratischer Polynome bei der Interpolation simuliert und somit zu einem Verfahren der Ordnung $O(h_n^3)$ beziehungsweise $O(h_n^3 \cdot \mathrm{ld}(h_n^{-1}))$ führt. Variante III schließlich stellt das Verfahren der Ordnung $O(h_n^4)$ beziehungsweise $O(h_n^4 \cdot \mathrm{ld}(h_n^{-1}))$ dar, das stattdessen kubische Interpolation simuliert.

Als Modellproblem verwenden wir Beispiel 1 aus Abschnitt 9.1. Abbildung 35 zeigt das bei sukzessiver Verfeinerung bis $\epsilon = 4^{-5}$ entstehende dünne Gitter, falls mit Variante I gearbeitet wird. Zum Vergleich ist nochmals das entsprechende Gitter abgebildet, welches man bei Verwendung des Standardverfahrens aus Abschnitt 6 erhält. In Abbildung 36 sehen wir die beiden adaptiv verfeinerten dünnen Gitter, die sich aus dem Einsatz von Variante II beziehungsweise Varinate III ergeben. Wie man sieht, nimmt die Anzahl benötigter Gitterpunkte mit zunehmender Genauigkeit des verwendeten Lösungsverfahrens ab. So besteht der jeweilige Binärbaum im Standardfall aus 506 Knoten, bei Variante I aus 468 und bei Variante II aus 432 Knoten. Mit Variante III schließlich ergibt sich ein Baum, der nur 418 Knoten enthält.

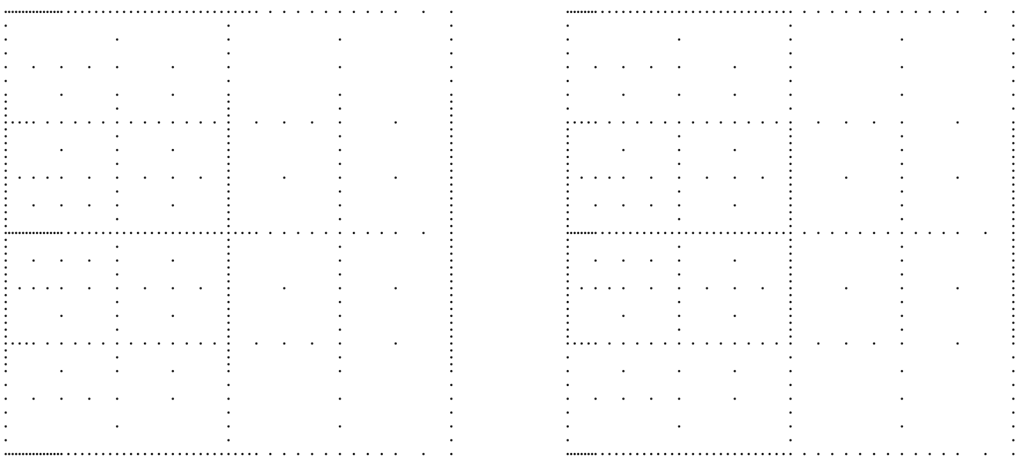


ABB. 35. *Beispiel 1: Adaption ($\epsilon = 4^{-5}$) im Standardfall (links) und bei Variante I (rechts).*

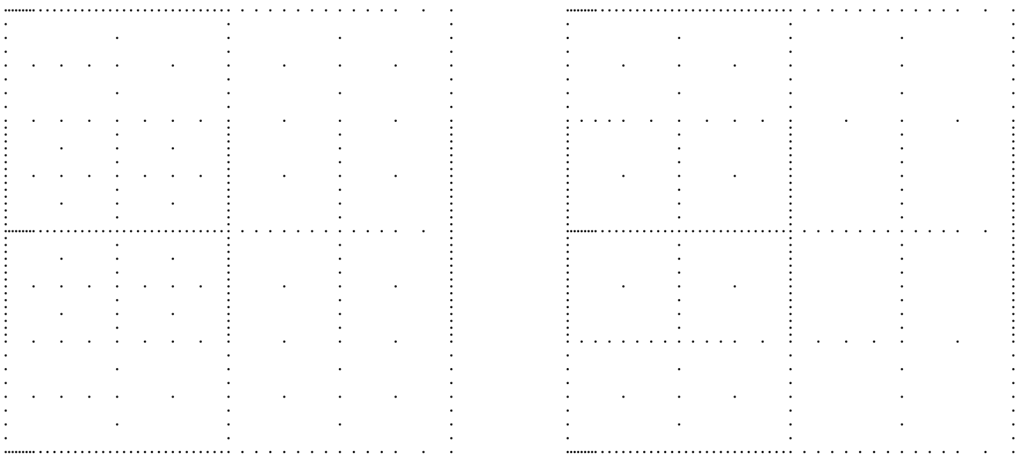


ABB. 36. *Beispiel 1: Adaption ($\epsilon = 4^{-5}$) bei Variante II (links) und Variante III (rechts).*

Tabelle 26 zeigt die Ergebnisse für Variante I im regulären Fall. Vergleicht man diese mit den Resultaten aus Tabelle 6, so erkennt man die verbesserten Werte von $\rho_P$, $\rho_\infty$ sowie $\rho_2$. Ab $n = 8$ liegen diese Werte alle sehr nahe bei 4. Es scheint also, als sei der Einfluß des logarithmischen Faktors zumindest für größere $n$ verschwunden. In Abbildung 37 ist das Verhalten des durch $h_n^2$ dividierten Fehlers im Mittelpunkt dargestellt. Der ab $n = 9$ nahezu konstante Verlauf bestätigt die Vermutung, daß für den punktweisen Fehler eine Reihenentwicklung exisitiert mit $h_n^2$ als führendem Glied. Auf jeden Fall ist das mittels Variante I erzielte Resultat von größerer Genauigkeit als dasjenige im Standardfall. So ist jetzt der maximale Fehler auf $\tilde{G}_{14}^{(2)}$ nur $2.3 \cdot 10^{-9}$, wohingegen er ohne diese Modifikation des Verfahrens noch $1.6 \cdot 10^{-8}$ betrug. Der Fehler im Mittelpunkt geht von $4.2 \cdot 10^{-9}$ auf $-2.9 \cdot 10^{-11}$ zurück. Was die Konvergenzgeschwindigkeit betrifft, so weichen die Werte von $m_\delta$ bei Variante I und beim nicht modifizierten Verfahren nur geringfügig voneinander ab (561 statt 554 für $n = 14$).

TABELLE 26
*Beispiel 1, reguläre dünne Gitter $\tilde{G}_n^{(2)}$, Variante I.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_n^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_n^{(2)}\|_2$ | $\rho_2$ | $\|\tilde{e}_n^{(2)}\|_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $7.316 \cdot 10^{-4}$ | | $7.316 \cdot 10^{-4}$ | | $7.316 \cdot 10^{-4}$ | | $1.195 \cdot 10^{-3}$ | | 10 | 12 |
| 2 | $-7.393 \cdot 10^{-4}$ | 0.99 | $2.753 \cdot 10^{-3}$ | 0.27 | $1.528 \cdot 10^{-3}$ | 0.48 | $7.330 \cdot 10^{-2}$ | 0.02 | 10 | 26 |
| 3 | $1.971 \cdot 10^{-4}$ | 3.75 | $1.438 \cdot 10^{-3}$ | 1.91 | $4.221 \cdot 10^{-4}$ | 3.62 | $5.122 \cdot 10^{-2}$ | 1.43 | 19 | 58 |
| 4 | $6.356 \cdot 10^{-6}$ | 31.01 | $8.480 \cdot 10^{-4}$ | 1.70 | $1.651 \cdot 10^{-4}$ | 2.56 | $2.844 \cdot 10^{-2}$ | 1.80 | 22 | 130 |
| 5 | $-3.068 \cdot 10^{-6}$ | 2.07 | $3.333 \cdot 10^{-4}$ | 2.54 | $4.617 \cdot 10^{-5}$ | 3.58 | $1.471 \cdot 10^{-2}$ | 1.93 | 36 | 290 |
| 6 | $-1.469 \cdot 10^{-6}$ | 2.09 | $1.066 \cdot 10^{-4}$ | 3.13 | $1.165 \cdot 10^{-5}$ | 3.96 | $7.435 \cdot 10^{-3}$ | 1.98 | 48 | 642 |
| 7 | $-4.354 \cdot 10^{-7}$ | 3.37 | $3.055 \cdot 10^{-5}$ | 3.49 | $2.834 \cdot 10^{-6}$ | 4.11 | $3.730 \cdot 10^{-3}$ | 1.99 | 51 | 1410 |
| 8 | $-1.152 \cdot 10^{-7}$ | 3.78 | $8.239 \cdot 10^{-6}$ | 3.71 | $6.821 \cdot 10^{-7}$ | 4.15 | $1.867 \cdot 10^{-3}$ | 2.00 | 89 | 3074 |
| 9 | $-2.937 \cdot 10^{-8}$ | 3.92 | $2.148 \cdot 10^{-6}$ | 3.84 | $1.643 \cdot 10^{-7}$ | 4.15 | $9.337 \cdot 10^{-4}$ | 2.00 | 125 | 6658 |
| 10 | $-7.390 \cdot 10^{-9}$ | 3.97 | $5.495 \cdot 10^{-7}$ | 3.91 | $3.971 \cdot 10^{-8}$ | 4.14 | $4.669 \cdot 10^{-4}$ | 2.00 | 161 | 14338 |
| 11 | $-1.852 \cdot 10^{-9}$ | 3.99 | $1.414 \cdot 10^{-7}$ | 3.89 | $9.641 \cdot 10^{-9}$ | 4.12 | $1.164 \cdot 10^{-4}$ | 4.01 | 228 | 30722 |
| 12 | $-4.632 \cdot 10^{-10}$ | 4.00 | $3.604 \cdot 10^{-8}$ | 3.92 | $2.350 \cdot 10^{-9}$ | 4.10 | $3.043 \cdot 10^{-5}$ | 3.83 | 299 | 65538 |
| 13 | $-1.159 \cdot 10^{-10}$ | 4.00 | $9.102 \cdot 10^{-9}$ | 3.96 | $5.746 \cdot 10^{-10}$ | 4.09 | $7.701 \cdot 10^{-6}$ | 3.95 | 407 | 139266 |
| 14 | $-2.896 \cdot 10^{-11}$ | 4.00 | $2.291 \cdot 10^{-9}$ | 3.97 | $1.409 \cdot 10^{-10}$ | 4.08 | $1.926 \cdot 10^{-6}$ | 4.00 | 561 | 294914 |

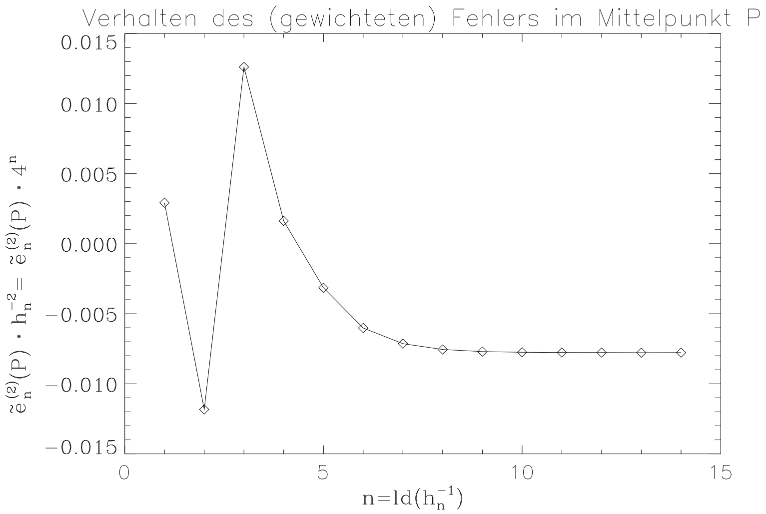



ABB. 37. *Variante I: Mit $h_n^{-2}$ multiplizierter Fehler $\tilde{e}_n^{(2)}(P)$ im Mittelpunkt $P = (\frac{1}{2}, \frac{1}{2})$.*

In Tabelle 27 sind die Resultate bei adaptiver Verfeinerung zusammengestellt. Der Vergleich mit den entsprechenden Werten aus Tabelle 7 zeigt erneut, daß Variante I bessere Näherungslösungen produziert. Im Falle sukzessiver Reduzierung von $\epsilon$ bis $\epsilon = 4^{-9}$ (letzte Zeile in den beiden Tabellen) erhalten wir bei Variante I mit 10120 Knoten einen maximalen Fehler von $7.7 \cdot 10^{-7}$. Im Mittelpunkt beträgt der Fehler $1.3 \cdot 10^{-7}$. Das Standardverfahren produziert demgegenüber eine Näherungslösung, deren Fehler im Mittelpunkt $1.2 \cdot 10^{-6}$ und deren maximaler Fehler $1.5 \cdot 10^{-6}$ beträgt. Der zugehörige Baum besteht dabei aus 10794 Knoten. Lediglich die Energienorm des Fehlers ist ohne Modifikation des Verfahrens kleiner als bei Variante I, nämlich $5.2 \cdot 10^{-4}$ gegenüber $7.6 \cdot 10^{-4}$.

Tabelle 27

*Beispiel 1, adaptiv erzeugte dünne Gitter (Kriterium (38)), Variante I.*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\Vert\tilde{e}_\epsilon^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\tilde{e}_\epsilon^{(2)}\Vert_2$ | $\rho_2$ | $\Vert\tilde{e}_\epsilon^{(2)}\Vert_{E10}$ | $\rho_E$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-7.393 \cdot 10^{-4}$ | | $2.753 \cdot 10^{-3}$ | | $1.528 \cdot 10^{-3}$ | | $2.976 \cdot 10^{-1}$ | | 26 |
| $4^{-2}$ | $-3.897 \cdot 10^{-4}$ | 1.90 | $1.569 \cdot 10^{-3}$ | 1.75 | $5.190 \cdot 10^{-4}$ | 2.94 | $1.542 \cdot 10^{-1}$ | 1.93 | 52 |
| $4^{-3}$ | $-7.680 \cdot 10^{-5}$ | 5.07 | $1.411 \cdot 10^{-3}$ | 1.11 | $4.914 \cdot 10^{-4}$ | 1.06 | $8.073 \cdot 10^{-2}$ | 1.91 | 100 |
| $4^{-4}$ | $8.666 \cdot 10^{-5}$ | 0.89 | $2.833 \cdot 10^{-4}$ | 4.98 | $8.016 \cdot 10^{-5}$ | 6.13 | $4.119 \cdot 10^{-2}$ | 1.96 | 222 |
| $4^{-5}$ | $-1.917 \cdot 10^{-5}$ | 4.52 | $9.673 \cdot 10^{-5}$ | 2.93 | $2.404 \cdot 10^{-5}$ | 3.33 | $2.057 \cdot 10^{-2}$ | 2.00 | 516 |
| $4^{-6}$ | $-2.117 \cdot 10^{-6}$ | 9.06 | $2.862 \cdot 10^{-5}$ | 3.38 | $6.270 \cdot 10^{-6}$ | 3.83 | $9.978 \cdot 10^{-3}$ | 2.06 | 1188 |
| $4^{-7}$ | $2.414 \cdot 10^{-7}$ | 8.77 | $7.706 \cdot 10^{-6}$ | 3.71 | $1.762 \cdot 10^{-6}$ | 3.56 | $4.849 \cdot 10^{-3}$ | 2.06 | 2748 |
| $4^{-8}$ | $-2.409 \cdot 10^{-8}$ | 10.02 | $2.111 \cdot 10^{-6}$ | 3.65 | $4.576 \cdot 10^{-7}$ | 3.85 | $2.235 \cdot 10^{-3}$ | 2.17 | 6488 |
| $4^{-9}$ | $6.411 \cdot 10^{-9}$ | 3.76 | $4.978 \cdot 10^{-7}$ | 4.24 | $1.115 \cdot 10^{-7}$ | 4.10 | $7.124 \cdot 10^{-4}$ | 3.14 | 15804 |
| $4^{-9}$ | $1.274 \cdot 10^{-7}$ | – | $7.731 \cdot 10^{-7}$ | – | $1.473 \cdot 10^{-7}$ | – | $7.608 \cdot 10^{-4}$ | – | 10120 |

Wir betrachten nun die mittels Variante II erzielten Resultate. Tabelle 28 beschreibt den regulären Fall. Der Fehler im Mittelpunkt verhält sich dabei für $n > 8$ bereits wie $O(h_n^4 \cdot \mathrm{ld}(h_n^{-1}))$. Hier scheinen Symmetrieeffekte zur Eliminierung der Fehlerterme ungerader Ordnung zu führen. Die Werte von $\rho_\infty$ zeigen demgegenüber $O(h_n^3 \cdot \mathrm{ld}(h_n^{-1}))$-Charakteristik. Auch die Werte von $\rho_2$ liegen bei 8 und sprechen so für ein Verfahren dritter Ordnung. Vergleicht man das auf $\tilde{G}_{14}^{(2)}$ mit Hilfe von Variante II erzielte Resultat mit den Ergebnissen im Standardfall und bei Variante I (siehe die Tabellen 6 sowie 26), so ist die Genauigkeit der jetzt gewonnenen Näherungslösung um Größenordnungen besser. Der Fehler im Mittelpunkt liegt bei $-5.8 \cdot 10^{-16}$ gegenüber $4.2 \cdot 10^{-9}$ (nicht modifiziertes Verfahren) beziehungsweise $-2.9 \cdot 10^{-11}$ (Variante I), der maximale Fehler geht von $1.6 \cdot 10^{-8}$ im Standardfall beziehungsweise $2.3 \cdot 10^{-9}$ bei Variante I auf nunmehr $1.3 \cdot 10^{-12}$ zurück. Die Werte von $m_\delta$ sind im Vergleich zum Standardfall jetzt zwar größer (z.B. bei $n = 14$: 694 statt 554), sie nehmen jedoch beim Übergang von $n$ zu $n+1$ wie zuvor etwa um einen Faktor $\sqrt{2}$ zu. Größenordnungsmäßig verschlechtert sich die Konvergenzgeschwindigkeit also nicht.

Tabelle 28

*Beispiel 1, reguläre dünne Gitter $\tilde{G}_n^{(2)}$, Variante II.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\Vert\tilde{e}_n^{(2)}\Vert_\infty$ | $\rho_\infty$ | $\Vert\tilde{e}_n^{(2)}\Vert_2$ | $\rho_2$ | $\Vert\tilde{e}_n^{(2)}\Vert_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $7.316 \cdot 10^{-4}$ | | $7.316 \cdot 10^{-4}$ | | $7.316 \cdot 10^{-4}$ | | $1.195 \cdot 10^{-3}$ | | 10 | 12 |
| 2 | $-4.669 \cdot 10^{-4}$ | 1.57 | $1.874 \cdot 10^{-3}$ | 0.39 | $9.662 \cdot 10^{-4}$ | 0.76 | $7.260 \cdot 10^{-2}$ | 0.02 | 10 | 26 |
| 3 | $2.500 \cdot 10^{-4}$ | 1.87 | $7.469 \cdot 10^{-4}$ | 2.51 | $3.146 \cdot 10^{-4}$ | 3.07 | $4.973 \cdot 10^{-2}$ | 1.46 | 17 | 58 |
| 4 | $-1.047 \cdot 10^{-4}$ | 2.39 | $2.576 \cdot 10^{-4}$ | 2.90 | $8.020 \cdot 10^{-5}$ | 3.92 | $2.786 \cdot 10^{-2}$ | 1.78 | 21 | 130 |
| 5 | $-1.146 \cdot 10^{-5}$ | 9.14 | $6.597 \cdot 10^{-5}$ | 3.90 | $1.339 \cdot 10^{-5}$ | 5.99 | $1.446 \cdot 10^{-2}$ | 1.93 | 30 | 290 |
| 6 | $-9.916 \cdot 10^{-7}$ | 11.56 | $1.227 \cdot 10^{-5}$ | 5.38 | $1.852 \cdot 10^{-6}$ | 7.23 | $7.322 \cdot 10^{-3}$ | 1.97 | 43 | 642 |
| 7 | $-7.490 \cdot 10^{-8}$ | 13.24 | $1.939 \cdot 10^{-6}$ | 6.33 | $2.426 \cdot 10^{-7}$ | 7.63 | $3.676 \cdot 10^{-3}$ | 1.99 | 60 | 1410 |
| 8 | $-5.465 \cdot 10^{-9}$ | 13.71 | $2.813 \cdot 10^{-7}$ | 6.89 | $3.041 \cdot 10^{-8}$ | 7.98 | $1.840 \cdot 10^{-3}$ | 2.00 | 86 | 3074 |
| 9 | $-3.839 \cdot 10^{-10}$ | 14.24 | $3.834 \cdot 10^{-8}$ | 7.34 | $3.736 \cdot 10^{-9}$ | 8.14 | $9.203 \cdot 10^{-4}$ | 2.00 | 128 | 6658 |
| 10 | $-2.685 \cdot 10^{-11}$ | 14.30 | $5.027 \cdot 10^{-9}$ | 7.63 | $4.532 \cdot 10^{-10}$ | 8.24 | $4.602 \cdot 10^{-4}$ | 2.00 | 181 | 14338 |
| 11 | $-1.838 \cdot 10^{-12}$ | 14.61 | $6.450 \cdot 10^{-10}$ | 7.79 | $5.468 \cdot 10^{-11}$ | 8.29 | $1.214 \cdot 10^{-4}$ | 3.79 | 258 | 30722 |
| 12 | $-1.254 \cdot 10^{-13}$ | 14.66 | $8.174 \cdot 10^{-11}$ | 7.89 | $6.581 \cdot 10^{-12}$ | 8.31 | $3.078 \cdot 10^{-5}$ | 3.94 | 357 | 65538 |
| 13 | $-8.465 \cdot 10^{-15}$ | 14.81 | $1.029 \cdot 10^{-11}$ | 7.94 | $7.921 \cdot 10^{-13}$ | 8.31 | $7.718 \cdot 10^{-6}$ | 3.99 | 498 | 139266 |
| 14 | $-5.829 \cdot 10^{-16}$ | 14.52 | $1.291 \cdot 10^{-12}$ | 7.97 | $9.542 \cdot 10^{-14}$ | 8.30 | $1.926 \cdot 10^{-6}$ | 4.01 | 694 | 294914 |

In Tabelle 29 sind die Resultate bei adaptiver Verfeinerung zusammengefaßt. Die letzte Zeile zeigt hierbei für den Fall sukzessiver Verfeinerung bis $\epsilon = 4^{-9}$, daß mit 8168 Knoten ein maximaler Fehler von $1.9 \cdot 10^{-7}$ erreicht wird. Mit unserem Standardverfahren ergibt sich in derselben Situation ein Baum mit 10794 Knoten, wobei der maximale Fehler $1.5 \cdot 10^{-6}$ beträgt.

Tabelle 29
*Beispiel 1, adaptiv erzeugte dünne Gitter (Kriterium (38)), Variante II.*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_\epsilon^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_\epsilon^{(2)}\|_2$ | $\rho_2$ | $\|\tilde{e}_\epsilon^{(2)}\|_{E10}$ | $\rho_E$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-4.669 \cdot 10^{-4}$ | | $1.874 \cdot 10^{-3}$ | | $9.662 \cdot 10^{-4}$ | | $2.973 \cdot 10^{-1}$ | | 26 |
| $4^{-2}$ | $-1.844 \cdot 10^{-3}$ | 0.25 | $2.389 \cdot 10^{-3}$ | 0.78 | $1.265 \cdot 10^{-3}$ | 0.76 | $2.476 \cdot 10^{-1}$ | 1.20 | 32 |
| $4^{-3}$ | $-1.037 \cdot 10^{-5}$ | 177.82 | $5.864 \cdot 10^{-4}$ | 4.07 | $2.746 \cdot 10^{-4}$ | 4.61 | $8.908 \cdot 10^{-2}$ | 2.78 | 96 |
| $4^{-4}$ | $5.586 \cdot 10^{-5}$ | 0.19 | $2.458 \cdot 10^{-4}$ | 2.39 | $6.316 \cdot 10^{-5}$ | 4.35 | $4.582 \cdot 10^{-2}$ | 1.94 | 198 |
| $4^{-5}$ | $-1.198 \cdot 10^{-5}$ | 4.66 | $2.546 \cdot 10^{-5}$ | 9.65 | $8.075 \cdot 10^{-6}$ | 7.82 | $2.310 \cdot 10^{-2}$ | 1.98 | 432 |
| $4^{-6}$ | $3.470 \cdot 10^{-7}$ | 34.52 | $1.481 \cdot 10^{-5}$ | 1.72 | $3.026 \cdot 10^{-6}$ | 2.67 | $1.165 \cdot 10^{-2}$ | 1.98 | 904 |
| $4^{-7}$ | $5.178 \cdot 10^{-7}$ | 0.67 | $2.865 \cdot 10^{-6}$ | 5.17 | $6.028 \cdot 10^{-7}$ | 5.02 | $5.666 \cdot 10^{-3}$ | 2.06 | 2008 |
| $4^{-8}$ | $7.538 \cdot 10^{-8}$ | 6.87 | $1.085 \cdot 10^{-6}$ | 2.64 | $1.696 \cdot 10^{-7}$ | 3.55 | $2.682 \cdot 10^{-3}$ | 2.11 | 4220 |
| $4^{-9}$ | $1.157 \cdot 10^{-8}$ | 6.52 | $2.260 \cdot 10^{-7}$ | 4.80 | $2.701 \cdot 10^{-8}$ | 6.28 | $1.015 \cdot 10^{-3}$ | 2.64 | 9248 |
| $4^{-9}$ | $1.471 \cdot 10^{-8}$ | – | $1.909 \cdot 10^{-7}$ | – | $2.585 \cdot 10^{-8}$ | – | $1.070 \cdot 10^{-3}$ | – | 8168 |

Abschließend soll noch die dritte Variante unseres Verfahrens diskutiert werden. Abbildung 38 zeigt zunächst auf $G_5^{(2)}$ die auf dem regulären dünnen Gitter $\tilde{G}_{10}^{(2)}$ berechnete Näherungslösung $\tilde{u}_{10}^{(2)}$ sowie ihren Fehler $\tilde{e}_{10}^{(2)}$. Dieser weist wiederum das typische oszillierende Verhalten auf.

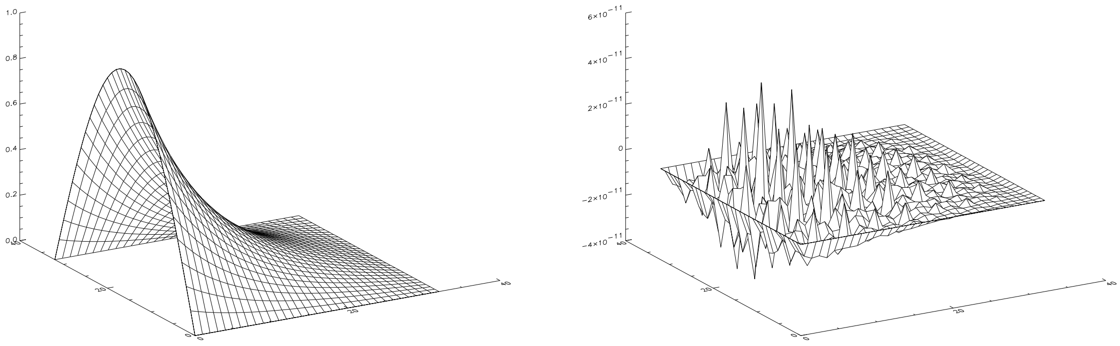


Abb. 38. *Beispiel 1: Die Näherungslösung $\tilde{u}_{10}^{(2)}$ und ihr Fehler $\tilde{e}_{10}^{(2)}$ bei Variante III.*

Tabelle 30 enthält nun die Ergebnisse für Variante III im regulären Fall, die im Vergleich zu den mittels Variante II erzielten Resultaten nochmals eine spürbare Verbesserung bedeuten. Da sich im Falle $n = 14$ bereits Rundungsfehler bemerkbar machen, betrachten wir die Situation bei $n = 13$. Hier liegt für Variante III der maximale Fehler nur noch bei $3.1 \cdot 10^{-14}$, während er in Tabelle 28 noch um etwa einen Faktor 328 größer war. Ebenso ist die $L_2$-Norm des Fehlers mit $2.5 \cdot 10^{-15}$ jetzt erheblich kleiner als bei Variante II ($7.9 \cdot 10^{-13}$). Abbildung 39 veranschaulicht das ab $n = 8$ sichtbare $O(h_n^4 \cdot \mathrm{ld}(h_n^{-1}))$-Verhalten des Fehlers im Mittelpunkt. Der Verlauf der Kurve legt dabei erneut die Existenz einer Reihenentwicklung des Fehlers mit $h_n^4 \cdot \mathrm{ld}(h_n^{-1})$ als führendem

Glied nahe. Die Werte $m_\delta$ steigen im Vergleich zu Variante II nochmals geringfügig an (bei $n = 14$ jetzt 715 statt 694), was sich jedoch größenordnungsmäßig nicht auf die Konvergenzgeschwindigkeit auswirkt.

TABELLE 30
*Beispiel 1, reguläre dünne Gitter $\tilde{G}_n^{(2)}$, Variante III.*

| $n$ | $\tilde{e}_n^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_n^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_n^{(2)}\|_2$ | $\rho_2$ | $\|\tilde{e}_n^{(2)}\|_{E_m}$ | $\rho_E$ | $m_\delta$ | $K_n$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $7.316 \cdot 10^{-4}$ | | $7.316 \cdot 10^{-4}$ | | $7.316 \cdot 10^{-4}$ | | $1.195 \cdot 10^{-3}$ | | 10 | 12 |
| 2 | $-4.669 \cdot 10^{-4}$ | 1.57 | $1.364 \cdot 10^{-3}$ | 0.54 | $7.749 \cdot 10^{-4}$ | 0.94 | $7.179 \cdot 10^{-2}$ | 0.02 | 10 | 26 |
| 3 | $2.205 \cdot 10^{-4}$ | 2.12 | $3.629 \cdot 10^{-4}$ | 3.76 | $1.914 \cdot 10^{-4}$ | 4.05 | $4.968 \cdot 10^{-2}$ | 1.45 | 17 | 58 |
| 4 | $-6.935 \cdot 10^{-5}$ | 3.18 | $1.064 \cdot 10^{-4}$ | 3.41 | $3.506 \cdot 10^{-5}$ | 5.46 | $2.787 \cdot 10^{-2}$ | 1.78 | 20 | 130 |
| 5 | $1.063 \cdot 10^{-6}$ | 65.24 | $1.983 \cdot 10^{-5}$ | 5.37 | $3.830 \cdot 10^{-6}$ | 9.15 | $1.447 \cdot 10^{-2}$ | 1.93 | 30 | 290 |
| 6 | $6.940 \cdot 10^{-7}$ | 1.53 | $2.484 \cdot 10^{-6}$ | 7.98 | $3.339 \cdot 10^{-7}$ | 11.47 | $7.323 \cdot 10^{-3}$ | 1.98 | 45 | 642 |
| 7 | $6.714 \cdot 10^{-8}$ | 10.34 | $2.359 \cdot 10^{-7}$ | 10.53 | $2.520 \cdot 10^{-8}$ | 13.25 | $3.676 \cdot 10^{-3}$ | 1.99 | 66 | 1410 |
| 8 | $5.027 \cdot 10^{-9}$ | 13.36 | $1.938 \cdot 10^{-8}$ | 12.17 | $1.793 \cdot 10^{-9}$ | 14.05 | $1.840 \cdot 10^{-3}$ | 2.00 | 93 | 3074 |
| 9 | $3.403 \cdot 10^{-10}$ | 14.77 | $1.481 \cdot 10^{-9}$ | 13.09 | $1.231 \cdot 10^{-10}$ | 14.57 | $9.203 \cdot 10^{-4}$ | 2.00 | 135 | 6658 |
| 10 | $2.252 \cdot 10^{-11}$ | 15.11 | $1.057 \cdot 10^{-10}$ | 14.01 | $8.360 \cdot 10^{-12}$ | 14.72 | $4.602 \cdot 10^{-4}$ | 2.00 | 192 | 14338 |
| 11 | $1.462 \cdot 10^{-12}$ | 15.40 | $7.219 \cdot 10^{-12}$ | 14.64 | $5.619 \cdot 10^{-13}$ | 14.88 | $1.214 \cdot 10^{-4}$ | 3.79 | 274 | 30722 |
| 12 | $9.528 \cdot 10^{-14}$ | 15.34 | $4.793 \cdot 10^{-13}$ | 15.06 | $3.769 \cdot 10^{-14}$ | 14.91 | $3.078 \cdot 10^{-5}$ | 3.94 | 367 | 65538 |
| 13 | $6.134 \cdot 10^{-15}$ | 15.53 | $3.131 \cdot 10^{-14}$ | 15.31 | $2.517 \cdot 10^{-15}$ | 14.97 | $7.718 \cdot 10^{-6}$ | 3.99 | 526 | 139266 |
| 14 | $3.608 \cdot 10^{-16}$ | 17.00 | $2.331 \cdot 10^{-15}$ | 13.43 | $2.061 \cdot 10^{-16}$ | 12.21 | $1.926 \cdot 10^{-6}$ | 4.01 | 715 | 294914 |

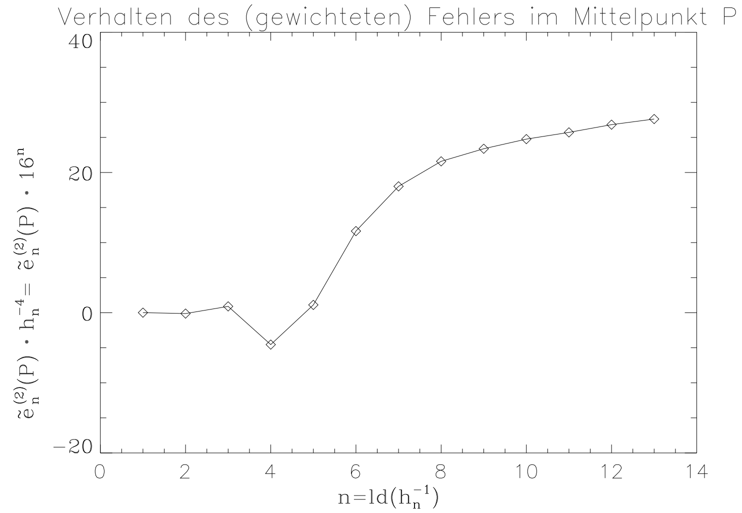



ABB. 39. *Variante III: Mit $h_n^{-4}$ multiplizierter Fehler $\tilde{e}_n^{(2)}(P)$ im Mittelpunkt $P = (\frac{1}{2}, \frac{1}{2})$.*

Tabelle 31 schließlich stellt die Ergebnisse für den Fall adaptiver Verfeinerung bereit. Der Vergleich mit Tabelle 29 zeigt eine weitere Verbesserung in der $L_2$- und der Maximumsnorm des Fehlers, die jedoch nicht so deutlich ausfällt wie im regulären Fall. Der Fehler im Mittelpunkt ist sogar wieder größer.

Zusammenfassend halten wir fest, daß die in Abschnitt 8.2 vorgestellten Techniken zur Erhöhung der Ordnung des Fehlers sehr gute Ergebnisse liefern. Dies gilt insbesondere für den Fall regulärer dünner Gitter, wo die Ordnung $O(h_n^4 \cdot \mathrm{ld} h_n^{-1}))$ sowohl für den punktweisen Fehler als auch für die $L_2$- und die Maximumsnorm des Fehlers erreicht

TABELLE 31
*Beispiel 1, adaptiv erzeugte dünne Gitter (Kriterium (38)), Variante III.*

| $\epsilon$ | $\tilde{e}_\epsilon^{(2)}(P)$ | $\rho_P$ | $\|\tilde{e}_\epsilon^{(2)}\|_\infty$ | $\rho_\infty$ | $\|\tilde{e}_\epsilon^{(2)}\|_2$ | $\rho_2$ | $\|\tilde{e}_\epsilon^{(2)}\|_{E_{10}}$ | $\rho_E$ | $K_\epsilon$ |
|---|---|---|---|---|---|---|---|---|---|
| $4^{-1}$ | $-4.669 \cdot 10^{-4}$ | | $1.364 \cdot 10^{-3}$ | | $7.749 \cdot 10^{-4}$ | | $2.969 \cdot 10^{-1}$ | | 26 |
| | | 0.63 | | 0.33 | | 0.40 | | 1.45 | |
| $4^{-2}$ | $-7.452 \cdot 10^{-4}$ | | $4.138 \cdot 10^{-3}$ | | $1.942 \cdot 10^{-3}$ | | $2.047 \cdot 10^{-1}$ | | 42 |
| | | 1.76 | | 4.27 | | 5.24 | | 2.42 | |
| $4^{-3}$ | $-4.241 \cdot 10^{-4}$ | | $9.684 \cdot 10^{-4}$ | | $3.706 \cdot 10^{-4}$ | | $8.456 \cdot 10^{-2}$ | | 98 |
| | | 3.89 | | 2.46 | | 4.37 | | 1.89 | |
| $4^{-4}$ | $-1.090 \cdot 10^{-4}$ | | $3.937 \cdot 10^{-4}$ | | $8.485 \cdot 10^{-5}$ | | $4.470 \cdot 10^{-2}$ | | 196 |
| | | 8.77 | | 7.01 | | 5.41 | | 1.97 | |
| $4^{-5}$ | $1.243 \cdot 10^{-5}$ | | $5.619 \cdot 10^{-5}$ | | $1.567 \cdot 10^{-5}$ | | $2.268 \cdot 10^{-2}$ | | 420 |
| | | 3.10 | | 3.42 | | 4.19 | | 2.02 | |
| $4^{-6}$ | $4.014 \cdot 10^{-6}$ | | $1.642 \cdot 10^{-5}$ | | $3.742 \cdot 10^{-6}$ | | $1.125 \cdot 10^{-2}$ | | 872 |
| | | 16.21 | | 5.04 | | 5.20 | | 2.06 | |
| $4^{-7}$ | $2.477 \cdot 10^{-7}$ | | $3.257 \cdot 10^{-6}$ | | $7.193 \cdot 10^{-7}$ | | $5.466 \cdot 10^{-3}$ | | 1846 |
| | | 60.03 | | 6.27 | | 7.35 | | 2.12 | |
| $4^{-8}$ | $4.126 \cdot 10^{-9}$ | | $5.194 \cdot 10^{-7}$ | | $9.793 \cdot 10^{-8}$ | | $2.583 \cdot 10^{-3}$ | | 3896 |
| | | 0.29 | | 3.42 | | 4.28 | | 2.66 | |
| $4^{-9}$ | $-1.430 \cdot 10^{-8}$ | | $1.518 \cdot 10^{-7}$ | | $2.288 \cdot 10^{-8}$ | | $9.706 \cdot 10^{-4}$ | | 8146 |
| $4^{-9}$ | $-2.583 \cdot 10^{-8}$ | – | $1.133 \cdot 10^{-7}$ | – | $2.032 \cdot 10^{-8}$ | – | $1.032 \cdot 10^{-3}$ | – | 7648 |

werden kann. Die Energienorm fällt dagegen erwartungsgemäß nach wie vor linear mit $O(h_n)$ ab. Betrachtet man die mit Hilfe der verschiedenen Verfahren auf dem regulären dünnen Gitter $\tilde{G}_{14}^{(2)}$ berechneten Näherungslösungen beziehungsweise deren Fehler (siehe Tabelle 32), so sind der Fehler im Mittelpunkt sowie die $L_2$- und die Maximumsnorm des Fehlers bei Variante III am kleinsten. Nur die Energienorm des Fehlers ist im Falle des nicht modifizierten Algorithmus' kleiner als bei allen ordnungserhöhenden Varianten. Dies war zu erwarten, da unser Standardverfahren Näherungslösungen mit minimalem Fehler in der Energienorm produziert (siehe Satz 6.2 und Satz 7.1).

TABELLE 32
*Die verschiedenen Techniken auf $\tilde{G}_{14}^{(2)}$ im Vergleich.*

| Verfahren | $\|\tilde{e}_{14}^{(2)}(P)\|$ | $\|\tilde{e}_{14}^{(2)}\|_\infty$ | $\|\tilde{e}_{14}^{(2)}\|_2$ | $\|\tilde{e}_{14}^{(2)}\|_{E_{10}}$ |
|---|---|---|---|---|
| Standard | $4.164 \cdot 10^{-9}$ | $1.643 \cdot 10^{-8}$ | $1.718 \cdot 10^{-9}$ | $8.372 \cdot 10^{-7}$ |
| Variante I | $2.896 \cdot 10^{-11}$ | $2.291 \cdot 10^{-9}$ | $1.409 \cdot 10^{-10}$ | $1.926 \cdot 10^{-6}$ |
| Variante II | $5.829 \cdot 10^{-16}$ | $1.291 \cdot 10^{-12}$ | $9.542 \cdot 10^{-14}$ | $1.926 \cdot 10^{-6}$ |
| Variante III | $3.608 \cdot 10^{-16}$ | $2.331 \cdot 10^{-15}$ | $2.061 \cdot 10^{-16}$ | $1.926 \cdot 10^{-6}$ |

**10. Schlußbemerkungen.** In dieser Arbeit haben wir zwei unterschiedliche Konzepte dünner Gitter im $d$-dimensionalen Raum vorgestellt und diskutiert. Dabei wurde zum einen die Anzahl der Punkte in den verschiedenen Gittern bestimmt, zum anderen wurden Aussagen zu den Approximationseigenschaften der jeweiligen auf stückweise $d$-linearen hierarchischen Basisfunktionen aufbauenden Dünn-Gitter-Funktionenräume getroffen. Als erste Anwendung dünner Gitter, die über die Repräsentation glatter Funktionen hinausgeht, haben wir ein Verfahren zur Lösung von Poisson-Problemen mit Dirichlet-Nebenbedingungen präsentiert, wobei auch Maßnahmen zur Erhöhung der Ordnung zur Sprache gekommen sind. Als Ergebnis halten wir fest, daß dünne Gitter sowohl bei der Repräsentation glatter Funktionen von $d$ Veränderlichen als auch bei der Lösung der Poisson-Gleichung im $d$-dimensionalen Raum bei annähernd gleicher Genauigkeit der jeweiligen Resultate mit signifikant weniger Gitterpunkten auskommen.

Eine weitere Methode zur Gewinnung von Lösungen für Randwertprobleme auf dünnen Gittern stellt die in [18] eingeführte Kombinationstechnik dar. Im zweidimensionalen Fall wird hierbei das betrachtete Problem zunächst auf $2n-1$ verschiedenen, in $\tilde{G}_n^{(2)}$ enthaltenen vollen Gittern mit jeweils $O(2^{n+1})$ beziehungsweise $O(2^n)$ Gitterpunkten unabhängig voneinander gelöst. Die sich daraus ergebenden $2n-1$ Näherungslösungen werden anschließend zu einer einzigen Näherungslösung auf dem regulären dünnen Gitter $\tilde{G}_n^{(2)}$ linear kombiniert. Während das in der vorliegenden Arbeit beschriebene Verfahren also direkt auf dem dünnen Gitter mittels der Methode der Finiten Elemente eine Lösung bestimmt, erfordert die Kombinationstechnik kein explizites Lösen des gegebenen Problems auf dem dünnen Gitter. Allerdings ist ihre Anwendbarkeit bislang auf reguläre dünne Gitter beschränkt. Für Details zur Kombinationstechnik sei auf [18] und [15] verwiesen.

Inzwischen sind dünne Gitter auch bei allgemeineren als den hier behandelten Randwertproblemen erfolgreich eingesetzt worden. In [10] wird das hier vorgestellte Verfahren um die Möglichkeit der Behandlung zweidimensionaler Laplace-Probleme auf krummlinig berandeten Gebieten erweitert, in [42] gelangt eine modifizierte Version des Dünn-Gitter-Lösers aus dieser Arbeit bei der Lösung der Stokes-Gleichungen zum Einsatz. Dem ständig an Bedeutung gewinnenden Aspekt der Parallelisierbarkeit numerischer Methoden wurde in [25] mit der Implementierung unseres Verfahrens auf einer iPSC/2-Hypercube-Parallelrechnerarchitektur Rechnung getragen.

Über den Bereich der partiellen Differentialgleichungen hinaus gelangen dünne Gitter derzeit mit Erfolg bei der Fouriertransformation (siehe [40]) sowie bei der numerischen Quadratur (siehe [39]) zur Anwendung.

Durch die im Vergleich zu herkömmlichen, vollen Gittern drastische Reduzierung der Anzahl der erforderlichen Gitterpunkte, die mit einem nur geringen Verlust an Genauigkeit des jeweils erzielten Resultats einhergeht, stellen dünne Gitter für zahlreiche Aufgabenstellungen der numerischen Mathematik — insbesondere bei Problemen mit sehr vielen Unbekannten oder Problemen hoher Dimensionalität — ein interessantes und nützliches Hilfsmittel dar.

[1] O. Axelsson und V. A. Barker, *Finite Element Solution of Boundary Value Problems: Theory and Computation*, Academic Press, Orlando, 1984.

[2] D. Bai und A. Brandt, *Local mesh refinement multi-level techniques*, SIAM J. Sci. Stat. Comput., 8 (1987), S. 109–134.

[3] R. E. Bank, *PLTMG User's Guide, Edition 5.0*, Technical Report, Department of Mathematics, University of California at San Diego, (1988).

[4] ——, *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations*, SIAM, Philadelphia, 1990.

[5] R. E. Bank, T. Dupont und H. Yserentant, *The hierarchical basis multigrid method*, Numerische Mathematik, 52 (1988), S. 427–458.

[6] R. E. Bank und H. Yserentant, *Some remarks on the hierarchical basis multigrid method*, in Domain Decomposition Methods (Proceedings), Los Angeles, Januar 1988, T. F. Chan, R. Glowinski, J. Périaux und O. Widlund, Hrsg., SIAM, Philadelphia, 1989, S. 140–146.

[7] A. Brandt, *Multi-level adaptive solutions to boundary value problems*, Math. Comp., 31 (1977), S. 333–390.

[8] H.-J. Bungartz, *An adaptive Poisson solver using hierarchical bases and sparse grids*, in Proceedings of the IMACS International Symposium on Iterative Methods in Linear Algebra, Brüssel, 2.4.–4.4. 1991, P. de Groen und R. Beauwens, Hrsg., Elsevier, Amsterdam, 1992 (auch: Technische Universität München, SFB-Bericht Nr. 342/21/91 A).

[9] P. Deuflhard, P. Leinen und H. Yserentant, *Concepts of an adaptive hierarchical finite element code*, IMPACT of Computing in Science and Engineering, 1 (1989), S. 3–35.

[10] B. Edlinger, *Dünn besetzte Gitter zur Lösung elliptischer Differentialgleichungen bei krummlinig berandeten Gebieten*, Technische Universität München, Diplomarbeit, (1990).

[11] G. Faber, *Über stetige Funktionen*, Mathematische Annalen, 66 (1909), S. 81–94.

[12] O. Friedrich, *Spezifikation und Implementierung von Werkzeugen für die Beschreibung adaptiver Mehrgitterverfahren*, Technische Universität München, Diplomarbeit, (1990).

[13] W. J. Gordon, *Distributive lattices and the approximation of multivariate functions*, in Approximation with Special Emphasis on Spline Functions, I. J. Schoenberg, Hrsg., Academic Press, New York, 1969, S. 223–277.

[14] W. J. Gordon und C. A. Hall, *Transfinite element methods: blending-function interpolation over arbitrary curved element domains*, Numerische Mathematik, 21 (1973), S. 109–129.

[15] M. Griebel, *The combination technique for the sparse grid solution of PDE's on multiprocessor machines*, Technische Universität München, SFB-Bericht Nr. 342/14/91 A, (1991).

[16] ——, *Parallel multigrid methods on sparse grids*, in [24], (1991), S. 211–221 (auch: Technische Universität München, SFB–Bericht Nr. 342/30/90 A).

[17] ——, *A parallelizable and vectorizable multi-level algorithm on sparse grids*, in Parallel Algorithms for Partial Differential Equations: Proceedings of the Sixth GAMM-Seminar, Kiel, 19.1.–21.1. 1990, Notes on Numerical Fluid Mechanics (Vol. 31), W. Hackbusch, Hrsg., Vieweg, Braunschweig, 1991, S. 94–100 (auch: Technische Universität München, SFB–Bericht Nr. 342/20/90 A).

[18] M. Griebel, M. Schneider und C. Zenger, *A combination technique for the solution of sparse grid problems*, in Proceedings of the IMACS International Symposium on Iterative Methods in Linear Algebra, Brüssel, 2.4.–4.4. 1991, P. de Groen und R. Beauwens, Hrsg., Elsevier, Amsterdam, 1992 (auch: Technische Universität München, SFB-Bericht Nr. 342/19/90 A).

[19] W. Hackbusch, *Multigrid Methods and Applications*, Springer-Verlag, Berlin Heidelberg, 1985.

[20] ——, *Theorie und Numerik elliptischer Differentialgleichungen*, Teubner, Stuttgart, 1986.

[21] ——, *Iterative Lösung großer schwachbesetzter Gleichungssysteme*, Teubner, Stuttgart, 1991.

[22] W. Hackbusch und U. Trottenberg, Hrsg., *Multigrid Methods I*, Proceedings of the 1st European Conference on Multigrid Methods, Köln-Porz, November 1981, Lecture Notes in Mathematics 960, Springer-Verlag, Berlin Heidelberg, 1982.

[23] ——, Hrsg., *Multigrid Methods II*, Proceedings of the 2nd European Conference on Multigrid Methods, Köln, Oktober 1985, Lecture Notes in Mathematics 1228, Springer-Verlag, Berlin

Heidelberg, 1986.
[24] ———, Hrsg., *Multigrid Methods III*, Proceedings of the 3rd European Conference on Multigrid Methods, Bonn, Oktober 1990, International Series of Numerical Mathematics (Vol.98), Birkhäuser Verlag, Basel, 1991.
[25] W. Hahn, *Parallelisierung eines adaptiven hierarchischen Dünn-Gitter-Verfahrens*, Technische Universität München, Diplomarbeit, (1990).
[26] J. P. Hennart und E. H. Mund, *On the h- and p-versions of the extrapolated Gordon's projector with applications to elliptic equations*, SIAM J. Sci. Stat. Comput., 9 (1988), S. 773–791.
[27] P. Hofmann, *Asymptotic expansions of the discretization error of boundary value problems of the Laplace equation in rectangular domains*, Numerische Mathematik, 9 (1967), S. 302–322.
[28] W. Hoppenstedt, Hrsg., *Sparse Ideas on Sparse Grids*, Proceedings of the First International Conference on Applied Loriotics (LAN), Castel di Mare (Lago di Bodo), 11.3.–16.3. 1991, Müller-Lüdenscheid Publishing Company, Hammerfest, Timbuktu, Furth im Wald, 1991.
[29] T. Linner, *Konvergenz iterativer Verfahren mit dünnen Gittern*, Technische Universität München, Diplomarbeit, (1991).
[30] S. McCormick, *Fast adaptive composite grid methods: theory for the variational case*, in Defect correction methods, theory and applications (Computing, Supplementum 5), K. Böhmer und H. J. Stetter, Hrsg., Springer-Verlag, Wien New York, 1984, S. 115–121.
[31] S. McCormick und J. Thomas, *The fast adaptive composite grid method for elliptic equations*, Math. Comp., 46 (1986), S. 439–456.
[32] C. Pflaum, *Datenstrukturen für dünne Gitter*, Technische Universität München, Fortgeschrittenenpraktikum, (1991).
[33] J. Stoer, *Numerische Mathematik I*, Springer-Verlag, Berlin Heidelberg, 5. Auflage, 1989.
[34] J. Stoer und R. Bulirsch, *Numerische Mathematik II*, Springer-Verlag, Berlin Heidelberg, 3. Auflage, 1990.
[35] G. Strang und G. J. Fix, *An Analysis of the Finite Element Method*, Prentice Hall, Englewood Cliffs, New Jersey, 1973.
[36] H. Yserentant, *On the multi-level splitting of finite element spaces*, Numerische Mathematik, 49 (1986), S. 379–412.
[37] H. Yserentant, *Hierarchical bases*, in ICIAM 91: Proceedings of the Second International Conference on Industrial and Applied Mathematics, R. E. O'Malley, Jr. et al., Hrsg., SIAM, Philadelphia, 1992.
[38] C. Zenger, *Sparse grids*, in Parallel Algorithms for Partial Differential Equations: Proceedings of the Sixth GAMM-Seminar, Kiel, 19.1.–21.1. 1990, Notes on Numerical Fluid Mechanics (Vol. 31), W. Hackbusch, Hrsg., Vieweg, Braunschweig, 1991 (auch: Technische Universität München, SFB-Bericht Nr. 342/18/90 A).
[39] ———, *Funktionale Programmierung paralleler Algorithmen für numerische und technisch-wissenschaftliche Anwendungen*, in GAMM-Seminar über numerische Algorithmen auf Transputer-Systemen, Heidelberg, 31.5.–1.6. 1991, Teubner-Skripten zur Numerik, G. Bader, G. Wittum und R. Rannacher, Hrsg., Teubner, Stuttgart, 1992.
[40] C. Zenger und K. Hallatschek, *Fouriertransformation auf dünnen Gittern mit hierarchischen Basen*. Erscheint in Numerische Mathematik (auch: Technische Universität München, SFB-Bericht Nr. 342/9/91 A).
[41] O. C. Zienkiewicz, D. W. Kelly, J. Gago und I. Babuška, *Hierarchical finite element approaches, error estimates and adaptive refinement*, in The Mathematics of Finite Elements and Applications IV, J. R. Whiteman, Hrsg., Academic Press, London, 1982.
[42] S. Zimmer, *Lösung der Stokes-Gleichungen durch ein adaptives Verfahren mit hierarchischen Basisfunktionen*, Technische Universität München, Diplomarbeit, (1991).