

**Die Methode der Finiten Differenzen  
auf dünnen Gittern zur  
Lösung elliptischer und parabolischer  
partieller Differentialgleichungen**

**Dissertation**

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch–Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von

Thomas Schiekofer

aus

München

Bonn 1998

Angefertigt mit Genehmigung der Mathematisch–Naturwissenschaftlichen  
Fakultät der Rheinischen Friedrich–Wilhelms–Universität Bonn

1. Referent: Prof. Dr. Michael Griebel

2. Referent: Prof. Dr. Martin Rumpf

Tag der Promotion: 11. Dezember 1998

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Multilevel-Darstellung von Funktionen</b>	<b>6</b>
2.1	Multilevel-Unterraum-Splitting und Tensorproduktbasen . . .	6
2.2	Eigenschaften der Unterräume . . . . .	9
2.3	Dünne Gitter . . . . .	17
2.4	Adaptive dünne Gitter . . . . .	20
<b>3</b>	<b>Eine Algebra für Operatoren auf Funktionen</b>	<b>26</b>
3.1	Einfache algebraische Grundlagen . . . . .	26
3.2	Eine Algebra für Operatoren auf Funktionen . . . . .	27
<b>4</b>	<b>Finite Differenzen Operatoren</b>	<b>39</b>
4.1	Grundlagen . . . . .	39
4.2	Konstruktion . . . . .	43
4.3	Konsistenz . . . . .	47
4.4	Eigenschaften der FD auf dünnen Gittern . . . . .	53
4.4.1	Matrix-Typen . . . . .	53
4.4.2	Eigenwerte der Diskretisierungsmatrizen . . . . .	56
4.5	Stabilität und Konvergenz . . . . .	58
4.6	Finite Differenzen höherer Ordnung . . . . .	60
4.7	Behandlung von Neumann-Randbedingungen . . . . .	61
4.8	Behandlung allgemeiner Gebiete . . . . .	62
4.8.1	Gebietszerlegung und Finite Differenzen auf dünnen Gittern . . . . .	63
4.8.2	Behandlung krummlinig berandeter Gebiete . . . . .	64
4.8.3	Constructive Solid Geometry (CSG) . . . . .	68
<b>5</b>	<b>Anmerkungen zur Lösung von PDE's</b>	<b>75</b>
5.1	Das unidirektionale Prinzip – Ausnutzung der Tensor- produktstruktur . . . . .	75

5.2	Diskretisierung elliptischer Differentialoperatoren . . . . .	75
5.3	Iterativer Löser für die Problemstellungen: BiCGSTAB . . . . .	77
5.4	Jakobi-Vorkonditionierung . . . . .	78
5.4.1	Vorkonditionierung von Dirichlet-Problemen . . . . .	78
5.4.2	Vorkonditionierung von Neumann-Problemen . . . . .	78
5.5	Ein selbstadaptiver Algorithmus . . . . .	80
<b>6</b>	<b>Implementierung und Datenstrukturen</b>	<b>82</b>
6.1	Binäre Bäume . . . . .	82
6.2	Hash-Tabellen . . . . .	85
6.3	Verzeigerte Hash-Tabellen . . . . .	91
6.4	Kosten für die Operatoranwendung . . . . .	92
6.5	Parallelisierung . . . . .	94
6.5.1	Parallelisierungskonzept . . . . .	94
6.5.2	Parallelisierung einzelner Operationen . . . . .	95
6.5.3	Parallelisierung des Lösert . . . . .	98
<b>7</b>	<b>Numerische Resultate</b>	<b>103</b>
7.1	Diskrete Normen auf dünnen Gittern . . . . .	103
7.2	Elliptische Differentialgleichungen 2. Ordnung . . . . .	105
7.2.1	Laplace- und Poisson-Gleichung . . . . .	105
7.2.2	Helmholtz-Gleichung . . . . .	121
7.2.3	Allgemeine elliptische Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten . . . . .	125
7.2.4	Allgemeine elliptische Differentialgleichungen 2. Ordnung mit variablen Koeffizienten . . . . .	131
7.3	Behandlung elliptischer DGL auf allgemeinen Gebieten . . . . .	136
7.3.1	Gebietszerlegung . . . . .	136
7.3.2	Behandlung krummlinig berandeter Gebiete . . . . .	139
7.3.3	Gebietszerlegung bei krummlinig berandeten Gebieten . . . . .	143
7.4	Parabolische Differentialgleichungen . . . . .	147
7.4.1	Skalare parabolische Differentialgleichungen . . . . .	147

---

7.4.2	Nichtlineare Systeme parabolischer Differentialgleichungen . . . . .	151
7.4.3	Gray Scott Modell für Reaktion und Diffusion chemischer Spezies . . . . .	156
7.5	Navier Stokes Gleichungen . . . . .	164
7.5.1	Die Methode der künstlichen Kompressibilität . . . . .	165
7.5.2	Ein Projektionsverfahren zur Lösung der Navier Stokes Gleichungen auf dünnen Gittern . . . . .	172
<b>8</b>	<b>Abschließende Bemerkungen</b>	<b>183</b>
<b>A</b>	<b>Beweise</b>	<b>186</b>
A.1	Beweis zu Lemma 4.2 . . . . .	186
A.2	Beweis zu Satz 4.1 . . . . .	187
A.3	Beweis zu Satz 4.2 . . . . .	196
A.4	Beweis zu Satz 4.4 . . . . .	199
A.5	Beweis zu Satz 4.5 . . . . .	200
A.6	Beweis zu Satz 4.7 . . . . .	201
A.7	Beweis zu Satz 4.8 . . . . .	204
A.8	Beweis zu Satz 4.10 . . . . .	213
A.9	Beweis zu Satz 4.13 . . . . .	218



## 1 Einleitung

Erkenntnisse und Einsichten in naturwissenschaftliche Vorgänge kann man im wesentlichen auf drei verschiedene Arten gewinnen: durch (mechanische, physikalische, chemische, ...) Experimente, mittels theoretischer Überlegungen, die sich zu einem Großteil auf mathematische Beschreibungsmöglichkeiten wie partielle Differentialgleichungen stützen, sowie mit Hilfe numerischer Simulationen. Während die ersten beiden Vorgehensweisen seit langem von Forschern der einzelnen Fachrichtungen benutzt werden, interessiert man sich für numerische Simulationen erst durch die Entwicklung leistungsfähiger Großrechner. Numerische Simulationen werden dabei in den unterschiedlichsten Forschungsgebieten eingesetzt. Hierzu gehört etwa die Crash-Simulation bei Fahrzeugen, die alltägliche Wettervorhersage, die Optimierung von Reaktionen etwa bei Verbrennungsvorgängen oder die Optimierung der Aerodynamik von Flugzeugen zur Reduktion des Treibstoffverbrauchs. Trotz rasanter Leistungssteigerung und Erhöhung der Speicherkapazität moderner Rechner ist man heutzutage jedoch nicht in der Lage, viele interessante Problemstellungen aus den Bereichen Physik, Chemie, Astronomie oder anderen Wissenschaften auf Rechnern zu lösen (Stichwort "Grand Challenges"). Die Hauptgründe hierfür liegen zum einen im enormen Aufwand, der zur diskreten Darstellung von Lösungen der mathematischen Gleichungen aufgewendet werden muß, zum anderen in dem daraus resultierenden enormen Rechenaufwand, um diese Datenmengen überhaupt verarbeiten zu können. Aufgaben aus dem Bereich der "Grand Challenges" sind etwa die Turbulenzsimulation, das Gebiet der künstlichen Intelligenz, die Klimamodellierung, die Simulation der Verschmelzung schwarzer Löcher, die numerische Simulation der Vorgänge in einem Tokamak (thermonukleare Fusionsreaktoren) oder Simulationen von Aufgabenstellungen aus der Quantenchromodynamik.

Neben dem Einsatz leistungsfähiger Großrechner, die es erlauben, immer größere Datenmengen immer schneller zu verarbeiten, ist die stetige Weiterentwicklung numerischer Algorithmen die Grundvoraussetzung, auch die momentan scheinbar unlösbaren Probleme anzugehen. Hierbei sei an die algorithmische Entwicklung von der Eliminationsmethode nach Gauß bis zu Mehrgitterverfahren zur Lösung linearer Gleichungssysteme erinnert. Durch die Reduktion der Komplexität der Löser wurde durch Algorithmen eine enorme Geschwindigkeitssteigerung erreicht, so daß moderne numerische Lösungsverfahren nur noch einen Bruchteil des Aufwandes früherer Verfahren benötigen.

Einen wichtigen Schritt zur Reduktion der Speicherkomplexität stellen dünne Gitter dar. Sie basieren auf der stückweise linearen, hierarchischen Basis in einer Raumdimension, wobei die Erweiterung auf mehrere Dimensionen

durch den Tensorproduktansatz erfolgt. Ausgehend von einem Schema von Teilräumen erhält man die hierarchischen Basisfunktionen eines dünnen Gitters dadurch, daß man lediglich diejenigen Basisfunktionen berücksichtigt, die einen Mindestbeitrag etwa zum Interpolanten liefern. Dadurch reduziert sich die Zahl der notwendigen Gitterpunkte von  $\mathcal{O}((2^n)^d)$  im Falle eines konventionellen vollen Gitters auf  $\mathcal{O}(2^n n^{d-1})$  Gitterpunkte im Fall eines dünnen Gitters, falls die Daten genügend glatt sind ( $d$ : Dimension,  $n$ : Level der Diskretisierung). Gleichzeitig wird die Genauigkeit trotz der erheblich verminderten Anzahl an Gitterpunkten annähernd aufrecht erhalten. So verringert sich die Genauigkeit in der Maximumsnorm und der  $L_2$ -Norm lediglich von  $\mathcal{O}(h^2)$  im Fall voller Gitter auf  $\mathcal{O}(h^2 \log(h^{-1})^{d-1})$  im Fall dünner Gitter, wohingegen die Genauigkeit im Fall der Energienorm auf beiden Gittern von derselben Größenordnung  $\mathcal{O}(h)$  ist. Damit wird das Verhältnis von Aufwand zu Genauigkeit im Vergleich zu vollen Gittern erheblich verbessert. Durch ihre Konstruktion erlauben dünne Gitter zudem adaptive Verfeinerungen. Adaptivität zusammen mit den zuvor genannten Vorteilen machen dünne Gitter zu äußerst aussichtsreichen Kandidaten, die Komplexität zur Lösung oben genannter Problemstellungen weiter zu verringern. Dabei sind die theoretischen Grundlagen der dünnen Gitter nicht neu, wie Arbeiten etwa von K. I. Babenko, N. M. Korobov oder etwa S. A. Smolyak belegen. Derzeit ist auf dünnen Gittern lediglich die Methode der Finiten Elemente zur Lösung beliebiger, elliptischer partieller Differentialgleichungen 2. Ordnung verfügbar. Dabei erweist sich bereits die Implementierung der Methode der Finiten Elemente auf dünnen Gittern als teilweise schwierig. Insbesondere die Behandlung von allgemeinen Koeffizientenfunktionen beispielweise bei elliptischen Differentialgleichungen zweiter Ordnung ist hier durch die notwendige Integration nicht einfach.

Ziel der vorliegenden Arbeit ist es, mit der Methode der Finiten Differenzen auf dünnen Gittern ein Diskretisierungs-Verfahren bereitzustellen, das die Vorteile dünner Gitter ausnutzt, gleichzeitig aber einfach zu implementieren ist. Neben der notwendigen Konstruktion spezieller Finiten Differenzen Operatoren auf dünnen Gittern ist auch die theoretische und numerische Untersuchung von Konsistenz und Konvergenz Bestandteil dieser Arbeit. Dabei sind die theoretisch und numerisch zu erwartenden Konvergenzordnungen vergleichbar mit denen der Methode der Finiten Elemente auf dünnen Gittern. Es zeigt sich außerdem, daß die Konsistenzordnungen der Finiten Differenzen Operatoren identisch mit denen auf entsprechenden vollen Gittern sind. Zudem stellt sich heraus, daß etwa elliptische Differentialgleichungen zweiter Ordnung mit hinreichend glatten Koeffizientenfunktionen ohne großen programmiertechnischen Aufwand realisiert werden können. Im Hinblick auf die anfangs genannten Problemstellungen werden neben Gebietszerlegungstechniken auch Transformationen von physikalischen Gebieten auf quaderförmige Rechengebiete betrachtet. Durch die Kombination beider Vor-



gehensweisen ist man in der Lage, die Gebiete, die in realen Anwendungen von Interesse sind, auch mit Hilfe dünner Gitter handzuhaben. Zur schnellen und effizienten Lösung der resultierenden Gleichungssysteme wird eine einfache Vorkonditionierung für die Methode der Finiten Differenzen auf dünnen Gittern vorgestellt. Die Datenstruktur Hash-Tabelle ermöglicht in einer Implementierung einen Zugriff auf beliebige Daten in  $\mathcal{O}(1)$  Operationen und bietet damit die Grundlage zur Generierung effizienten Programmcodes. Zudem ist durch diese Datenstruktur die Verwaltung adaptiver Gitter problemlos und einfach möglich. Neben den bereits angegebenen Gebietszerlegungstechniken, die auf natürliche Art und Weise ein paralleles Vorgehen ermöglichen, wird eine Parallelisierung basierend auf "Open MP" für SMP-Maschinen zur Verfügung gestellt. Damit steht mit der vorgestellten Methode der Finiten Differenzen auf dünnen Gittern eine Methode zur Verfügung, die es erlaubt, auf allgemeinen Gebieten parallele, vorkonditionierte Löser für vorgegebene Problemstellungen zu konstruieren. Im folgenden sollen die einzelnen Bausteine hierzu vorgestellt werden. Die vorliegende Arbeit ist nun wie folgt gegliedert:

In Kapitel 2 werden dünne Gitter eingeführt sowie grundlegende Eigenschaften der Teilräume des zur Definition dünner Gitter hilfreichen Teilraumschemas angegeben. Zudem werden die Basistransformationen von stückweise linearer, nodaler Basis in stückweise lineare, hierarchische Basis sowie deren Inverse als Integraloperator, genauer als Wavelet-Transformierte, dargestellt. Bemerkungen zur Adaptivität bei dünnen Gittern beenden diesen Abschnitt.

Kapitel 3 befaßt sich nach einigen grundlegenden, algebraischen Eigenschaften dünner Gitter mit einer Algebra auf Operatoren für Funktionen auf dünnen Gittern. Dabei wird der Multiplikation zweier Funktionen auf dünnen Gittern besondere Aufmerksamkeit geschenkt.

In Kapitel 4 wird nach einigen grundlegenden Bemerkungen zu Finiten Differenzen die Konstruktion von Finite Differenzen Operatoren auf dünnen Gittern motiviert und durchgeführt. Theoretische Betrachtungen zur Konsistenz erster und zweiter Ableitungsoperatoren zeigen dabei dieselben Konsistenzordnungen wie auf entsprechenden vollen Gittern. Die Untersuchung der Stabilität und der Konvergenz folgen der Konstruktion der Operatoren. Diese Ergebnisse werden später im Abschnitt der Numerischen Resultate (Kapitel 7) mit zahlreichen numerischen Experimenten untermauert. Nach der Konvergenzaussage werden die Eigenschaften der Diskretisierungsmatrizen untersucht. Dabei stellt sich unter anderem heraus, daß die für theoretische Aussagen interessante  $M$ -Matrix-Eigenschaft für die Finiten Differenzen Operatoren auf dünnen Gittern nicht gegeben ist. Finite Differenzen Operatoren höherer Ordnung und die Behandlung von Neumann Randbedingungen beenden die Betrachtung der Finite Differenzen Operatoren auf

dünnen Gittern. Betrachtungen zur Lösung von Problemstellung auf allgemeinen Gebieten mit Gebietszerlegungstechniken sowie Gebietstransformationen, die nicht spezifisch für die Anwendung auf dünnen Gittern sind, sondern allgemeiner behandelt werden, runden dieses Kapitel ab.

Kapitel 5 betrachtet einige, speziell für Finite Differenzen Operatoren auf dünnen Gittern interessante Aspekte zur Lösung partieller Differentialgleichungen. In Abschnitt 5.2 wird die Diskretisierung einer elliptischen Differentialgleichung zweiter Ordnung mit Hilfe der Finiten Differenzen Operatoren auf dünnen Gittern sowie der Algebra für Operatoren auf Funktionen vorgestellt. Zudem wird ein iterativer Löser für die auftretenden linearen Gleichungssysteme mit stets nicht-symmetrischen Diskretisierungsmatrizen angegeben. Ein einfacher Vorkonditionierer und die Vorstellung eines selbst-adaptiven Algorithmus auf dünnen Gittern beenden diesen Abschnitt.

Implementiertechnische Aspekte werden in Kapitel 6 betrachtet. Dort steht neben der Diskussion einer geeigneten Datenstruktur für die Implementierung der Finiten Differenzen Operatoren, aber auch für die Implementierung dünner Gitter selbst, die Frage nach den Kosten für die Anwendung Finiten Differenzen auf dünnen Gittern sowie die Parallelisierung des vorgestellten Ansatzes mittels “Open MP” auf SMP-Maschinen im Mittelpunkt.

Das sich daran anschließende Kapitel 7 beinhaltet zahlreiche numerische Resultate für allgemeine elliptische Differentialgleichungen zweiter Ordnung, skalare parabolische Differentialgleichungen und nichtlineare, gekoppelte Systeme parabolischer Differentialgleichungen sowie der Navier Stokes Gleichungen. Dabei werden im Bereich der elliptischen Differentialgleichungen Problemstellungen der Dimensionen 2 bis 8 betrachtet.

Im Appendix befinden sich diejenigen Beweise zu Sätzen dieser Arbeit, die eine gewisse Länge überschreiten und somit den Lesefluß beeinträchtigen würden. Sämtlich der dort aufgeführten Beweise entstammen Lemmatas oder Sätzen des Kapitels 4.

An dieser Stelle möchte ich meinen Dank an diejenigen Personen aussprechen, die beim Erstellen und Gelingen dieser Arbeit hilfreich waren. Als erstes ist hier Prof. Dr. Michael Griebel zu nennen, der mein Interesse auf das Gebiet der dünnen Gitter gelenkt hat und während der Erstellung dieser Arbeit stets mit konstruktiven Vorschlägen und Ideen den Entwicklungsprozeß vorangetrieben hat. Sein Engagement und Interesse hat diese Arbeit positiv beeinflusst. Bei Professor Dr. Martin Rumpf bedanke ich mich für die Übernahme des Zweitgutachtens. Zudem gebührt mein Dank Dr. Gerhard Zumbusch, der in zahlreichen Diskussionen Hilfestellungen und Anregungen gegeben hat. Stephan Knapek und Dr. Ralf Hiptmair übernahmen das Korrekturlesen dieser Arbeit, Thomas Gerstner danke ich für zahlreiche und endlose Diskussionen, und Alexander Schweitzer gebührt mein Dank, da er

mir durch eine Einführung in OpenMP die Parallelisierung des Codes erleichterte. Zudem danke ich allen Kollegen der Abteilung Wissenschaftliches Rechnen und Numerische Simulation am Institut für Angewandte Mathematik der Universität Bonn für das gute Arbeitsklima und die ständige Bereitschaft zur Diskussion. Danken möchte ich auch dem BMBF (Bundesministerium für Bildung und Forschung) sowie dem Sonderforschungsbereich 256 der Universität Bonn für die materielle Unterstützung während der Erstellung dieser Arbeit.

Sankt Augustin, im Oktober 1998

## 2 Multilevel-Darstellung von Funktionen

### 2.1 Multilevel–Unterraum–Splitting und Tensorproduktbasen

Auf dem  $d$ -dimensionalen Einheitswürfel  $\bar{\Omega} = [0, 1]^d$  betrachten wir die Menge der Funktionen  $u \in \mathcal{X}^{2,d}(\bar{\Omega})$  mit existierenden und beschränkten gemischten Ableitungen bis maximal zweiten Grades (vgl. [Bun98] für Räume mit höheren Ableitungen), wobei

$$\mathcal{X}^{\mathbf{r},d}(\bar{\Omega}) = \left\{ u^{(d)} : \bar{\Omega} \subset \mathbb{R}^d \rightarrow \mathbb{R}, \frac{\partial^{|\mathbf{m}|} u}{\partial x_1^{m_1} \dots \partial x_d^{m_d}} \in C^0(\bar{\Omega}), \mathbf{m} \leq \mathbf{r} \right\}$$

gilt. Zudem bezeichnen wir mit  $\mathcal{X}_0^{\mathbf{r},d}(\bar{\Omega}) := \{u \in \mathcal{X}^{\mathbf{r},d}(\bar{\Omega}) : u|_{\Gamma} = 0\}$  denjenigen Unterraum von  $\mathcal{X}^{\mathbf{r},d}(\bar{\Omega})$ , dessen Funktionen am Rand  $\Gamma = \partial\bar{\Omega}$  des Gebiets verschwinden. Im weiteren Verlauf der theoretischen Betrachtungen werden wir uns der Einfachheit halber auf den Raum  $\mathcal{X}_0^{2,d}(\bar{\Omega})$  zurückziehen<sup>1</sup>.

Dabei bezeichnet  $\mathbf{r} = (r_1, \dots, r_d)$  einen Multi-Index, der den Grad der Ableitung bezüglich jeder Raumdimension angibt. Auf Multi-Indizes ( $\in \mathbb{N}_0^d$ ) führen wir die beiden Semi-Normen

$$|\mathbf{r}|_1 = \sum_{i=1}^d r_i$$

und

$$|\mathbf{r}|_{\infty} = \max_{1 \leq i \leq d} r_i$$

ein. Arithmetische Operationen wie Addition, Subtraktion, Skalarmultiplikation oder Vergleiche werden auf Multi-Indizes stets komponentenweise ausgeführt. Potenziert man eine Zahl  $a$  mit einem Multi-Index  $\mathbf{r}$ , so ist das Ergebnis der Vektor  $(a^{r_1}, \dots, a^{r_d})$ , der seinerseits wiederum als Multi-Index aufgefaßt werden kann.

Mit Hilfe des Multi-Index  $\mathbf{l}$ , der in einer Multilevel-Umgebung für jede Dimension einen Level angibt, definieren wir die Familie der  $d$ -dimensionalen Gitter

$$\{\bar{\Omega}_{\mathbf{l}}\}_{\mathbf{l} \in \mathbb{N}_0^d}$$

---

<sup>1</sup>Dabei sei an dieser Stelle angemerkt, daß der in [BuGr97] definierte Raum  $\mathcal{X}^p(\bar{\Omega})$ , der statt der Zugehörigkeit der gemischten Ableitungen zu  $C^0$  die Zugehörigkeit zu  $L_p$  verlangt, an dieser Stelle nicht verwendet werden kann. Der Grund hierfür ist, daß bei der Methode der Finiten Differenzen im Gegensatz zur Methode der Finiten Elemente Punktauswertungen vorgenommen werden müssen, was bei  $L_p$ -Räumen nicht sinnvoll ist. Somit muß hier die stärkere Voraussetzung, d.h. die Verwendung von  $C^0$ , gefordert werden.

mit den Schrittweiten

$$\mathbf{h}_l = (h_1, \dots, h_d) = (2^{-l_1}, \dots, 2^{-l_d}) = 2^{-|\mathbf{l}|},$$

die bezüglich unterschiedlicher Raumdimensionen verschieden sein können, für eine feste Raumdimension jedoch ein äquidistantes Gitter beschreiben. Die mit diesen Gittern verbundenen Gitterpunkte ergeben sich zu

$$x_{\mathbf{l}, \mathbf{i}} = (x_{l_1, i_1}, \dots, x_{l_d, i_d}) = \mathbf{i} \cdot \mathbf{h}_l$$

mit  $x_{l_j, i_j} = i_j \cdot h_j$  für  $i_j \in \{0, \dots, 2^{l_j}\}$ . Dabei beschreibt der Multi-Index  $\mathbf{i}$  die Lage der Gitterpunkte auf einem gegebenen, festen Level  $l$ . Jedem Gitterpunkt  $x_{\mathbf{l}, \mathbf{i}}$  läßt sich eine stückweise,  $d$ -lineare Hutfunktion

$$\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) = \prod_{j=1}^d \varphi_{l_j, i_j}(x_j) \tag{2.1}$$

mit  $\mathbf{x} = \{x_1, \dots, x_d\}$  zuordnen, die vermöge eines Tensorprodukts eindimensionaler Funktionen aufgebaut ist. Diesen eindimensionalen Funktionen liegt die (Mutter-)Funktion

$$\varphi(x) = \begin{cases} 1 - |x| & \text{für } x \in [-1, 1] \\ 0 & \text{sonst} \end{cases}$$

zugrunde, aus der mittels Translation und Dilatation die Funktionen  $\varphi_{l_j, i_j}(x)$  vermöge

$$\varphi_{l_j, i_j}(x_j) = \varphi\left(\frac{x_j - i_j \cdot h_j}{h_j}\right) \tag{2.2}$$

erzeugt werden. Mit Hilfe der den Gitterpunkten zugeordneten Basisfunktionen definieren wir den Raum der stückweise  $d$ -linearen Funktionen

$$V_l = \text{span}\{\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) : i_j = 1, \dots, 2^{l_j} - 1, 1 \leq j \leq d, \mathbf{l} \in \mathbf{N}^d\},$$

für den die zuvor definierten Ansatzfunktionen eine Basis darstellen. Zudem definieren wir die (eindimensionalen) Räume

$$V_{l_j, j} = \text{span}\{\varphi_{l_j, i_j}(x_j) : i_j = 1, \dots, 2^{l_j} - 1, l_j \in \mathbf{N}\} \tag{2.3}$$

bezüglich Raumdimension  $j$ , womit wir den  $d$ -dimensionalen Raum  $V_l$  auch mittels

$$V_l = \bigotimes_{j=1}^d V_{l_j, j}$$

in Tensorschreibweise darstellen können. Aufbauend auf den Räumen  $V_1$  definieren wir die Differenzräume  $W_1$  gemäß

$$W_1 = V_1 \setminus \bigoplus_{j=1}^d V_{1-\mathbf{e}_j} \quad (2.4)$$

wobei der  $d$ -dimensionale Einheitsvektor  $\mathbf{e}_j$  hier als Multi-Index aufzufassen ist. Aus formalen Gründen setzen wir zudem  $V_{1-\mathbf{e}_j} = \emptyset$ , falls  $l_j = 0$ . Damit können wir den Raum

$$V = \sum_{l_1=1}^{\infty} \cdots \sum_{l_d=1}^{\infty} W_{(l_1, \dots, l_d)} = \bigoplus_{\mathbf{l} \geq 1} W_{\mathbf{l}} \quad (2.5)$$

definieren, dessen Vervollständigung  $\bar{V}$  bezüglich der  $H^1$ -Norm mit dem Sobolev-Raum  $H_0^1(\bar{\Omega})$  zusammenfällt. Hierbei steht  $\mathbf{1}$  für den Multiindex  $\mathbf{1} = (1, \dots, 1)$ . Die Differenzräume  $W_1$  bestehen nach Definition aus denjenigen Funktionen in  $V_1$ , die ungerade Indizes besitzen, d.h.

$$W_1 = \text{span}\{\varphi_{\mathbf{l};\mathbf{i}}(\mathbf{x}) \in V_1 : i_j = 1, \dots, 2^{l_j} - 1 \text{ mit } i_j \text{ ungerade}, 1 \leq j \leq d\}.$$

Damit bilden die Räume  $W_1$  eine hierarchische Basis [Fab09, Yse86, Yse92], die einen endlich-dimensionalen Unterraum des  $H_0^1(\bar{\Omega})$  aufspannt. Dies ist eine einfache Verallgemeinerung der eindimensionalen hierarchischen Basis auf den  $d$ -dimensionalen Fall, wie sie in [Fab09] zu finden ist. Zudem definieren wir die eindimensionalen Räume

$$W_{l_j,j} = V_{l_j,j} \setminus V_{l_j-1,j} \quad (2.6)$$

bezüglich Raumdimension  $j$ , die wir genauso wie die eindimensionalen Räume  $V_{l_j,j}$  zur Definition von Operatoren im Verlauf des Abschnitts benötigen.

In Abbildung 1 sind die Räume  $W_1$  für den eindimensionalen Fall zu sehen. Die Träger der Basisfunktionen im zweidimensionalen Fall zusammen mit den Mittelpunkten (gekennzeichnet durch  $\bullet$ ) der entsprechenden Basisfunktionen sind in Abbildung 2 zu finden. Dabei wird ersichtlich, daß der Schnitt der Träger zweier Basisfunktionen eines Raumes  $W_1$  mit festem  $\mathbf{l}$  leer ist.

Nach den Definitionen der einzelnen Räume können wir nun beliebige Funktionen  $u \in H_0^1(\bar{\Omega})$  mittels

$$u(\mathbf{x}) = \sum_{\mathbf{l} \geq 1} u_{\mathbf{l}}(\mathbf{x}) = \sum_{\mathbf{l} \geq 1} \sum_{\mathbf{i} \in \mathbf{I}_{\mathbf{l}}} u_{\mathbf{l};\mathbf{i}} \cdot \varphi_{\mathbf{l};\mathbf{i}}(\mathbf{x}) \quad (2.7)$$

darstellen, wobei  $u_{\mathbf{l}}(\mathbf{x}) \in W_{\mathbf{l}}$  ist und

$$u_{\mathbf{l}}(\mathbf{x}) = \sum_{\mathbf{i} \in \mathbf{I}_{\mathbf{l}}} u_{\mathbf{l};\mathbf{i}} \cdot \varphi_{\mathbf{l};\mathbf{i}}(\mathbf{x}) \quad (2.8)$$

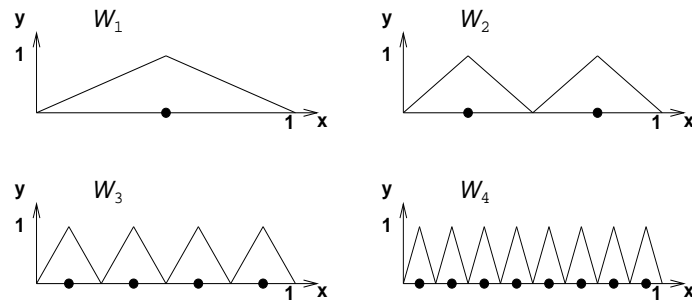


Abbildung 1: Basisfunktionen der Differenzräume  $W_l$  in einer Raumdimension für  $l \in \{1, 2, 3, 4\}$

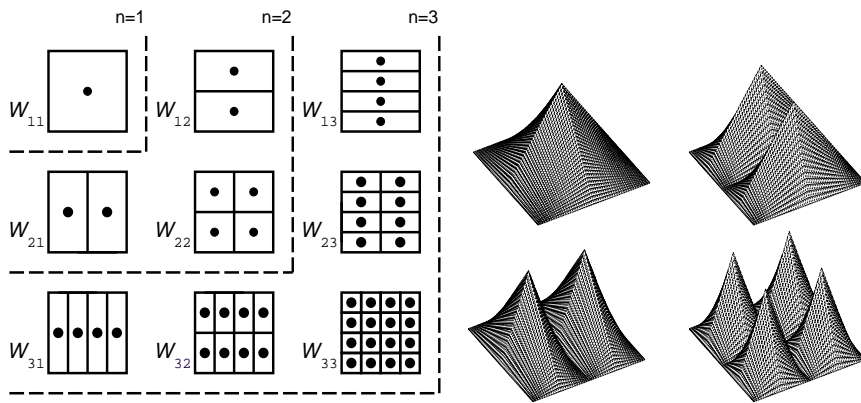


Abbildung 2: Träger der Basisfunktionen der Differenzräume in zwei Raumdimensionen (nur Basisfunktionen im Inneren von  $\bar{\Omega}$ , links) sowie die Basisfunktionen von  $W_{11}$ ,  $W_{12}$ ,  $W_{21}$  und  $W_{22}$  (rechts)

gilt. Hierbei sind die  $u_{\mathbf{i}}$  die Koeffizienten der Ansatzfunktionen (auch: hierarchische Überschüsse, vgl. [Zen90, Bun92, Bal94, Bun98]) und  $\mathbf{I}_1$  repräsentiert die Menge der Indizes

$$\mathbf{I}_1 = \{\mathbf{i} \in \mathbb{N}_0^d : i_j = 1, \dots, 2^{l_j} - 1 \text{ mit } i_j \text{ ungerade}, 1 \leq j \leq d\}. \quad (2.9)$$

## 2.2 Eigenschaften der Unterräume

In diesem Abschnitt fassen wir die Eigenschaften der Unterräume  $W_1$  sowie der sie aufbauenden Basisfunktionen zusammen. In unserer Aufzählung folgen wir eng der Darstellung in [Bun92, BuGr97, Bun98], wo auch sämtliche hier ausgelassenen Beweise zu finden sind.

Zur Formulierung der Abschätzungen im nachfolgenden Lemma 2.3 führen

wir zunächst die beiden Seminormen

$$|u|_{\mathbf{r},\infty} := \|D^{\mathbf{r}}u\|_{\infty}$$

und

$$|u|_{\mathbf{r},2} := \|D^{\mathbf{r}}u\|_2 = \left( \int_{\bar{\Omega}} |D^{\mathbf{r}}u|^2 d\mathbf{x} \right)^{1/2}$$

ein, wobei  $\mathbf{r}$  ein Multi-Index ist und  $D^{\mathbf{r}}u$  vermöge

$$D^{\mathbf{r}}u := \frac{\partial^{|\mathbf{r}|} u}{\partial x_1^{r_1} \cdots \partial x_d^{r_d}} \quad (2.10)$$

definiert ist. Zudem ist die Energienorm  $\|\cdot\|_E$  (des Laplace-Operators) definiert mittels

$$\|\cdot\|_{E,\Omega} = \left( \int_{\Omega} \sum_{j=1}^d \left( \frac{\partial u}{\partial x_j} \right)^2 d\mathbf{x} \right)^{1/2},$$

wobei wir nachfolgend  $\|\cdot\|_E \equiv \|\cdot\|_{E,\Omega}$  verwenden, wenn  $\Omega$  bekannt ist.

**Lemma 2.1** (vgl. [BuGr97], Lemma 2) Für jede stückweise  $d$ -lineare Basisfunktion  $\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$  gelten die folgenden Aussagen:

$$\begin{aligned} \|\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x})\|_{\infty} &= 1 \\ \|\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x})\|_{L^p} &= \left( \frac{2}{p+1} \right)^{d/p} \cdot 2^{-|\mathbf{l}|/p} \quad (p \geq 1, p \neq \infty) \\ \|\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x})\|_E &= \sqrt{2} \cdot \left( \frac{2}{3} \right)^{(d-1)/2} \cdot 2^{-|\mathbf{l}|/2} \left( \sum_{j=1}^d 2^{2 \cdot l_j} \right)^{1/2} \end{aligned}$$

**Lemma 2.2** (vgl. [BuGr97], Lemma 4) Sei  $u(\mathbf{x}) \in \mathcal{X}_0^{\mathbf{r},d}(\bar{\Omega})$  mit  $\mathbf{r} \geq \mathbf{2}$  bezüglich der Darstellung (2.7) gegeben. Für die hierarchischen Überschüsse  $u_{\mathbf{l},\mathbf{i}}$  gelten die Abschätzungen

$$\begin{aligned} |u_{\mathbf{l},\mathbf{i}}| &\leq 2^{-d} \cdot 2^{-2 \cdot |\mathbf{l}|_1} \cdot |u|_{\mathbf{2},\infty} \\ |u_{\mathbf{l},\mathbf{i}}| &\leq 2^{-d} \cdot \left( \frac{2}{3} \right)^{d/2} \cdot 2^{-(3/2) \cdot |\mathbf{l}|_1} \cdot |\chi_{\mathbf{l},\mathbf{i}} \cdot u|_{\mathbf{2},2} \end{aligned}$$

Dabei ist  $\chi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$  die charakteristische Funktion des Trägers der Funktion  $\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$ .

**Lemma 2.3** (vgl. [BuGr97], Lemma 1 und Lemma 5) Aus (2.4) zusammen mit (2.9) folgt

$$|W_{\mathbf{l}}| = |\mathbf{I}_{\mathbf{l}}| = 2^{|\mathbf{l}-\mathbf{1}|_1}. \quad (2.11)$$



Das Volumen des Trägers einer beliebigen Basisfunktion  $\varphi_{1,\mathbf{i}}(x, y) \in \mathcal{W}_1$  in  $d$  Dimensionen beträgt

$$\text{vol}(\text{supp}(\varphi_{1,\mathbf{i}})) = 2^{d-|\mathbf{l}_1|} .$$

Sei  $u(\mathbf{x}) \in \mathcal{X}_0^{\mathbf{r},d}(\bar{\Omega})$  mit  $\mathbf{r} \geq \mathbf{2}$  bezüglich der Darstellung (2.7) gegeben. Für  $u_1(\mathbf{x}) \in \mathcal{W}_1$  gelten die Abschätzungen

$$\begin{aligned} \|u_1(\mathbf{x})\|_\infty &\leq 2^{-d} \cdot 2^{-2 \cdot |\mathbf{l}_1|} \cdot |u|_{\mathbf{2},\infty} \\ \|u_1(\mathbf{x})\|_2 &\leq 3^{-d} \cdot 2^{-2 \cdot |\mathbf{l}_1|} \cdot |u|_{\mathbf{2},2} \\ \|u_1(\mathbf{x})\|_E &\leq \frac{1}{2 \cdot 12^{(d-1)/2}} \cdot 2^{-2 \cdot |\mathbf{l}_1|} \cdot \left( \sum_{j=1}^d 2^{2 \cdot l_j} \right)^{1/2} |u|_{\mathbf{2},\infty} \end{aligned}$$

**Beweis:** Vergleiche hierzu [Bun92, BuGr97].  $\square$

Die Berechnung der Koeffizienten  $u_{1,\mathbf{i}}$  in (2.7) bzw. (2.8) erfolgt über

$$u_{1,\mathbf{i}} = \left( \prod_{j=1}^d H_{x_{l_j, i_j}, h_j} \right) u(\mathbf{x}) =: H_{\mathbf{x}_{1,\mathbf{i}}, \mathbf{h}} u(\mathbf{x}) \quad (2.12)$$

mit

$$H_{\mathbf{x}_{1,\mathbf{i}}, \mathbf{h}} u(\mathbf{x}) : V_1 \rightarrow \bigoplus_{\mathbf{k}=1}^1 W_{\mathbf{k}} ,$$

wobei die Operatoren<sup>2</sup>  $H_{x_{l_j, i_j}, h_j} : V_{l_j} \rightarrow \bigoplus_{k_j=1}^{l_j} W_{k_j}$  in (2.12) in einer Raumdimension mittels

$$H_{x_{l_j, i_j}, h_j} = \left[ -\frac{1}{2} \quad 1 \quad -\frac{1}{2} \right]_{x_{l_j, i_j}, h_j} \quad (2.13)$$

$$H_{x_{l_j, i_j}, h_j} \circ u(x_{l_j, i_j}) = u(x_{l_j, i_j}) - \frac{1}{2} (u(x_{l_j, i_j} - h_j) + u(x_{l_j, i_j} + h_j))$$

für  $1 \leq i_j \leq 2^{l_j} - 1$  definiert sind. In  $d$  Raumdimensionen sind diese Operatoren entsprechend vermöge

$$\begin{aligned} H_{x_{l_j, i_j}, h_j} \circ u(\mathbf{x}_{1,\mathbf{i}}) &= u(x_{l_1, i_1}, \dots, x_{l_j, i_j}, \dots, x_{l_d, i_d}) - \\ &\quad \frac{1}{2} \left( u(x_{l_1, i_1}, \dots, x_{l_j, i_j} - h_j, \dots, x_{l_d, i_d}) + \right. \\ &\quad \left. u(x_{l_1, i_1}, \dots, x_{l_j, i_j} + h_j, \dots, x_{l_d, i_d}) \right) \quad (2.14) \end{aligned}$$

<sup>2</sup>In dieser Arbeit wird als Symbol für die Berechnung des hierarchischen Überschusses der Buchstabe  $H$  (Hierarchisierung) verwendet. Dieser Operator ist identisch mit dem Interpolationsoperator  $I_{x_{l_j, i_j}, l_j}$  in [Bun92, Bal94, Bun98].

definiert (vgl. [Zen90, Bun92, Bal94, Gri97], siehe auch [Fab09] für den eindimensionalen Fall). An einem Randknoten bezüglich der Raumdimension  $j$  entspricht hierbei obiger Operator der Identität. Der Term  $H_{\mathbf{x}_{1,i},\mathbf{h}}$  repräsentiert einen  $d$ -dimensionalen Stern, der den Koeffizienten  $u_{1,i}$  durch eine Linearkombination von Funktionswerten (nodalen Werten) der betrachteten Funktion  $u(\mathbf{x})$  auf  $\bar{\Omega}$  ermittelt (vgl. [Hac86, Bun92] für die Schreibweise von Operatoren, die vermöge Sternen definiert sind).  $H_{\mathbf{x}_{1,i},\mathbf{h}}$  ist dabei die Basistransformation von der stückweise  $d$ -linearen, nodalen Basis in die stückweise  $d$ -dimensionale, hierarchische Basis.

Für die Einführung der Finiten Differenzen Operatoren auf dünnen Gittern in Abschnitt 4.2 benötigen wir weitere Operatoren, die über die eindimensionalen Basistransformationen  $H_{x_{l_j,i_j},h_j}$  definiert werden. Zunächst führen wir den Operator  $\mathbf{H}^j$  für  $1 \leq j \leq d$  ein, der eine Kurzschreibweise für die Anwendung der eindimensionalen Hierarchisierung bezüglich Koordinatenrichtung  $j$  an allen Gitterpunkten darstellt und somit formal als Operator in  $d$  Dimensionen definiert ist.

**Definition 2.1 (Hierarchisierung und Enthierarchisierung):** *Unter dem Operator  $\mathbf{H}^j$  verstehen wir eine eindimensionale Hierarchisierung bezüglich der Raumdimension  $j$  mit  $1 \leq j \leq d$ , d.h. die Basistransformation von stückweise linearer, nodaler Basis in stückweise lineare, hierarchische Basis bezüglich der Raumdimension  $j$ . Dabei ist*

$$\left(\mathbf{H}^j \circ u\right)(\mathbf{x}_{1,i}) = H_{x_{l_j,i_j},h_j} \circ u(\mathbf{x}_{1,i})$$

mit  $H_{x_{l_j,i_j},h_j}$  aus (2.14) der Koeffizienten-Vektor, dessen Einträge bezüglich einer Dimension (nämlich  $j$ ) in hierarchischer und bezüglich aller übrigen Dimensionen in nodaler Basis gegeben sind. Dabei gilt

$$\begin{aligned} \mathbf{H}^j : & \left( \bigotimes_{s=1}^{j-1} Y_{l_s,s} \right) \otimes \left( V_{l_j,j} \right) \otimes \left( \bigotimes_{s=j+1}^d Y_{l_s,s} \right) \rightarrow \\ & \left( \bigotimes_{s=1}^{j-1} Y_{l_s,s} \right) \otimes \left( \bigoplus_{s_j=1}^{l_j} W_{s_j,j} \right) \otimes \left( \bigotimes_{s=j+1}^d Y_{l_s,s} \right), \end{aligned} \quad (2.15)$$

wobei  $Y_{l_s,s} \in \{V_{l_s,s}, \bigoplus_{t=1}^{l_s} W_{t,s}\}$  zu betrachten ist. Der  $d$ -dimensionale Hierarchisierungsoperator  $\mathbf{H}$ , der die Basistransformation bezüglich aller Koordinatenrichtungen durchführt und dessen Stern in (2.12) zusammen mit (2.13) eingeführt ist, ist definiert als

$$\mathbf{H} = \prod_{j=1}^d \mathbf{H}^j : \bigotimes_{s=1}^d V_{l_s,s} \rightarrow \bigotimes_{s=1}^d \left( \bigoplus_{t_s=1}^{l_s} W_{t_s} \right). \quad (2.16)$$

Zudem betrachten wir den Operator<sup>3</sup>

$$\mathbf{H}_k = \prod_{\substack{j=1 \\ j \neq k}}^d \mathbf{H}^j \quad (2.17)$$

mit

$$\begin{aligned} \mathbf{H}^j : & \left( \bigotimes_{s=1}^{j-1} V_{l_s, s} \right) \otimes \left( \bigoplus_{s_j=1}^{l_j} Y_{s_j} \right) \otimes \left( \bigotimes_{s=j+1}^d V_{l_s, s} \right) \rightarrow \\ & \left( \bigotimes_{s=1}^{j-1} \left( \bigoplus_{t=1}^{l_s} W_{t, s} \right) \right) \otimes \left( \bigoplus_{s_j=1}^{l_j} Y_{s_j} \right) \otimes \left( \bigotimes_{s=j+1}^d \left( \bigoplus_{t=1}^{l_s} W_{t, s} \right) \right) \end{aligned} \quad (2.18)$$

der die Hierarchisierung in alle Raumdimensionen mit Ausnahme der Dimension  $k$  bedeutet und führen  $\mathbf{E} = \mathbf{H}^{-1}$ ,  $\mathbf{E}^j = (\mathbf{H}^j)^{-1}$  sowie  $\mathbf{E}_k = \mathbf{H}_k^{-1}$  als Inverse der genannten Operatoren ein. Somit stellt  $\mathbf{H}_k \circ u$  einen Vektor dar, dessen Einträge bezüglich einer Dimension in nodaler (nämlich in  $k$ ) und bezüglich aller übrigen Dimensionen in hierarchischer Basis gegeben sind.

Obige Definitionen stützen sich bereits auf den diskreten Fall mit vorgegebenem Gitter. In kontinuierlicher Form kann diese Transformation durch eine Integraltransformation angegeben werden. Eine Darstellung hierzu gibt

**Satz 2.1 (kontinuierliche Darstellung der Hierarchisierung in 1D):** Sei  $u \in C^2(\bar{\Omega})$  und  $\varphi_{l,i}(x)$  wie zuvor aus  $\varphi(x)$  durch Translation und Dilatation erhalten. Der hierarchische Überschuss  $u_{l,i}$  an der Stelle  $x_{l,i}$  ist definiert als

$$u_{l,i} = \frac{h_l}{2} \int_{-\infty}^{\infty} \varphi_{l,i}(x) u''(x) dx \quad , \quad (2.19)$$

d.h. die eindimensionale Hierarchisierung ist ein Integraloperator mit dem Kern  $\varphi_{l,i}(x)$ .

**Beweis:** Wendet man partielle Integration auf das Integral in (2.19) an, so

---

<sup>3</sup>Man beachte an dieser Stelle die Indizes! Hochgestellte Indizes  $k$  bedeuten stets eine eindimensionale Hierarchisierung in Raumdimension  $k$ , wohingegen tiefgestellte Indizes  $k$  eine Hierarchisierung in alle  $d$  Raumdimensionen mit Ausnahme der angegebenen Dimension  $k$  bedeuten.

erhält man

$$\begin{aligned}
\frac{h_l}{2} \int_{-\infty}^{\infty} \varphi_{l,i}(x) u''(x) dx &= \frac{h_l}{2} \int_{x_{li}-h_l}^{x_{li}+h_l} \varphi_{l,i}(x) u''(x) dx \\
&= \underbrace{\frac{h_l}{2} [\varphi_{l,i}(x) u'(x)]_{x_{li}-h_l}^{x_{li}+h_l}}_{=0} - \frac{h_l}{2} \int_{x_{li}-h_l}^{x_{li}} \varphi_{l,i}(x)' u'(x) dx \\
&\quad - \frac{h_l}{2} \int_{x_{li}}^{x_{li}+h_l} \varphi_{l,i}(x)' u'(x) dx \\
&= \frac{1}{2} \int_{x_{li}-h_l}^{x_{li}} u'(x) dx - \frac{1}{2} \int_{x_{li}}^{x_{li}+h_l} u'(x) dx \\
&= u(x_{li}) - \frac{u(x_{li}-h_l) + u(x_{li}+h_l)}{2} \tag{2.20}
\end{aligned}$$

Ein Vergleich von (2.20) mit (2.13) liefert nun die Behauptung.  $\square$

Betrachtet man sich den Beweis des vorangegangenen Satzes 2.1 genauer, so stellt man fest, daß die beiden Teilintegrale in der zweiten Zeile der Gleichheitskette auch mit Hilfe eines Haar-Wavelets formuliert und somit zu einem Integral zusammengefaßt werden können. Der hier beobachtete Zusammenhang wird bereits in [XuSh92, LMR94] zur Konstruktion von Wavelet-Galerkin-Methoden für Randwertprobleme ausgenutzt. Ebenso interpretiert [Gri97] die Berechnung hierarchischer Koeffizienten als eine Integraltransformation, die der Wavelet-Transformation ähnlich ist. Betrachtet man die kontinuierliche Wavelet-Transformation (vgl. hierzu etwa [Dau92]), so stellt man fest, daß sich die Hierarchisierung einer gegebenen Funktion  $u$  als skalierte Wavelet-Transformation der Ableitung  $u'$  schreiben läßt. Dementsprechend läßt sich durch die Inversion der Wavelet-Transformation auch die Enthierarchisierung mittels einer Integraltransformation darstellen. Diese Aussagen sind Bestandteil des nachfolgenden Satzes.

**Satz 2.2 (Die Hierarchisierung und Enthierarchisierung als Wavelet-Transformation in 1D):** Sei  $u \in C^1(\bar{\Omega})$  und  $\psi(x)$  das Haar-Wavelet, das vermöge

$$\psi(x) = \begin{cases} 1 & : -1 \leq x < 0 \\ -1 & : 0 \leq x \leq 1 \\ 0 & : \text{sonst} \end{cases}$$

definiert ist (vgl. etwa [LMR94]). Der hierarchische Überschuß  $u_{l,i}$  an der

Stelle  $x_{l,i}$  berechnet sich aus

$$u_{l,i} = (\mathcal{H}_\psi u)(a, b) = \frac{1}{2} |a|^{1/2} \sqrt{2 \ln 2} \cdot (L_\psi u')(a, b) \quad (2.21)$$

mit den Substitutionen  $a = h_l = 2^{-l}$  und  $b = x_{l,i} = i \cdot h_l$ . In obiger Formel folgen wir der Notation in [LMR94], in der mit  $(L_\psi f)(a, b)$  die Wavelet-Transformierte

$$(L_\psi f)(a, b) = \frac{1}{\sqrt{c_\psi}} |a|^{-1/2} \int_{\mathbf{R}} f(t) \psi \left( \frac{t-b}{a} \right) dt$$

der Funktion  $f$  mit dem Wavelet  $\psi$  bezeichnet wird. Dabei gilt

$$0 < c_\psi = 2\pi \int_{\mathbf{R}} \frac{|\hat{\psi}(\omega)|}{|\omega|} d\omega < \infty ,$$

und speziell im Fall des Haar-Wavelets gilt  $c_\psi = 2 \ln 2$ . Mittels der inversen Wavelet-Transformation ergibt sich die Ausgangsfunktion  $u$  vermöge

$$\begin{aligned} u(x) &= \int u'(x) dx \\ &= \int \left( \frac{1}{\ln 2} \int_{\mathbf{R}} \int_{\mathbf{R}} |a|^{-1} \psi \left( \frac{x-b}{a} \right) (\mathcal{H}_\psi u)(a, b) \frac{da db}{a^2} \right) dx \end{aligned} \quad (2.22)$$

bis auf Integrations-Konstanten bei Integration von  $u'$ .

**Beweis:** Der Beweis für die Hierarchisierung ergibt sich sofort aus der Definition der Wavelet-Transformation sowie dem Beweis von Satz 2.1:

$$\begin{aligned} u_{l,i} &= (\mathcal{H}_\psi u)(a, b) \\ &= \frac{1}{2} |a|^{1/2} \sqrt{2 \ln 2} \cdot (L_\psi u')(a, b) \\ &= \frac{1}{2} |a|^{1/2} \sqrt{2 \ln 2} \cdot \underbrace{\left( \frac{1}{\sqrt{2 \ln 2}} |a|^{-1/2} \int_{\mathbf{R}} u'(x) \psi \left( \frac{x-b}{a} \right) dx \right)}_{\text{Haar-Wavelet-Transformierte von } u'} \\ &= \frac{1}{2} \int_{\mathbf{R}} u'(x) \psi \left( \frac{x-b}{a} \right) dx \end{aligned}$$

Ebenso ergibt sich der Beweis für die Enthierarchisierung aus der Definition

der inversen Wavelet-Transformation:

$$\begin{aligned}
u(x) &= \int u'(x) dx \\
&= \int \underbrace{\left( \frac{1}{\sqrt{2 \ln 2}} \int_{\mathbf{R}} \int_{\mathbf{R}} |a|^{-\frac{1}{2}} \psi \left( \frac{x-b}{a} \right) \cdot (L_\psi u')(a, b) \frac{da db}{a^2} \right)}_{\text{inverse Haar-Wavelet-Transformierte von } L_\psi u'} dx \\
&= \int \left( \frac{1}{\sqrt{2 \ln 2}} \int_{\mathbf{R}} \int_{\mathbf{R}} |a|^{-\frac{1}{2}} \psi \left( \frac{x-b}{a} \right) \cdot \left( 2 \cdot |a|^{-\frac{1}{2}} \frac{1}{\sqrt{2 \ln 2}} (\mathcal{H}_\psi u)(a, b) \right) \frac{da db}{a^2} \right) dx \\
&= \int \left( \frac{1}{\ln 2} \int_{\mathbf{R}} \int_{\mathbf{R}} |a|^{-1} \psi \left( \frac{x-b}{a} \right) (\mathcal{H}_\psi u)(a, b) \frac{da db}{a^2} \right) dx
\end{aligned}$$

Dabei ist jedoch zu beachten, daß bei der von uns gewählten Vorgehensweise nur die Funktion  $u'$ , die in die Wavelet-Transformation eingeht, wiedergewonnen werden kann. Daher ist eine zusätzliche Integration nötig. Somit kann  $u$  nur bis auf Konstanten rekonstruiert werden, Information geht folglich in dieser Darstellung verloren.  $\square$

Nach dem bisherigen ist klar, daß sich die  $d$ -dimensionale Hierarchisierung nun als  $d$ -dimensionales Integral schreiben läßt:

**Satz 2.3 (Integraldarstellung der Hierarchisierung in  $d$  Dimensionen):** Sei  $u \in \mathcal{X}^{\mathbf{r},d}(\bar{\Omega})$  mit  $\mathbf{r} \geq 2$  und  $\varphi_{1,\mathbf{i}}(\mathbf{x})$  definiert wie zuvor. Der hierarchische Überschuß  $u_{1,\mathbf{i}}$  an der Stelle  $\mathbf{x}$  ist definiert als

$$u_{1,\mathbf{i}} = \frac{2^{-|\mathbf{i}|_1}}{2^d} \int_{\mathbf{R}^d} \varphi_{1,\mathbf{i}}(\mathbf{x}) \frac{\partial^{2d} u(\mathbf{x})}{\partial x_1^2 \partial x_2^2 \cdots \partial x_d^2} d\mathbf{x} \quad , \quad (2.23)$$

d.h. der  $d$ -dimensionale Hierarchisierungsoperator ist ein Integraloperator mit Kern  $\varphi_{1,\mathbf{i}}(\mathbf{x})$ .

**Beweis:** vgl [BuGr97], Lemma 3  $\square$

Analog läßt sich nun die  $d$ -dimensionale Hierarchisierung als  $d$ -dimensionale Wavelet-Transformation darstellen, für die Enthierarchisierung gilt entsprechendes. Dabei wird ausgenutzt, daß sich die  $d$ -dimensionalen Ansatzfunktionen nach (2.1) als Tensorprodukt eindimensionaler Ansatzfunktionen (2.2) schreiben lassen. Führt man die Transformation (2.21) für die  $d$ -dimensionalen Ansatzfunktionen nun mit dem  $d$ -dimensionalen Analogon

der Haar-Basis aus, das ebenfalls vermöge eines Tensorprodukts gewonnen wird, so läßt sich die nötige  $d$ -dimensionale Integration sofort aufgrund der Tensorprodukt Darstellung durch  $d$  eindimensionale Integrale ausdrücken, die jeweils von der Bauart (2.21) sind. Dieselbe Vorgehensweise führt sofort auf das  $d$ -dimensionale Analogon von (2.22).

### 2.3 Dünne Gitter

Wir gehen in diesem Abschnitt von Funktionen  $u$ , die in unendlich-dimensionalen Räumen  $V$  definiert sind, über zu deren Interpolanten bzw. allgemeinen Funktionen  $\bar{u}$ , die in endlich-dimensionalen Unterräumen von  $V$  definiert sind. Für die Diskretisierung findet man in der Literatur oftmals volle Gitter (vgl. unter vielen z.B. [Hac86]), deren zugehörige endlich-dimensionale Ansatzräume mittels

$$V^{n,V,d} = \bigcup_{|\mathbf{l}|_\infty \leq n} W_{\mathbf{l}} \quad (2.24)$$

gegeben sind. Funktionen auf vollen Gittern<sup>4</sup>, also auf Gittern mit äquidistanter Schrittweite  $h$  in jede Koordinatenrichtung, können somit über

$$u^{n,V,d}(\mathbf{x}) = \mathbb{I}^{n,V,d} \circ u(\mathbf{x}) := \sum_{|\mathbf{l}|_\infty \leq n} \sum_{\mathbf{i} \in \mathbf{I}_{\mathbf{l}}} u_{\mathbf{l},\mathbf{i}} \cdot \varphi(\mathbf{x}) \quad (2.25)$$

dargestellt werden, wobei  $\mathbb{I}^{n,V,d} : V \rightarrow V^{n,V,d}$  der Interpolationsoperator des in (2.5) definierten Raums  $V$  auf den Vollgitterraum  $V^{n,V,d}$  des Levels  $n$  in  $d$  Dimensionen darstellt. In Abbildung 2 geben die gestrichelten Linien zwischen den Teilräumen an, welche Räume  $W_{\mathbf{l}}$  zu einem vollen Gitter eines vorgegebenen Levels zusammengefaßt werden. Abschätzungen für den Interpolationsfehler findet man beispielsweise in [Hac86, Tho95]. Volle Gitter besitzen das Problem des “Fluchs der Dimensionen”, d.h. die Anzahl der Punkte eines vollen Gitters nimmt exponentiell mit der Dimension zu. Hieraus resultiert die Suche nach einem Ansatzraum, der bei annähernd gleicher Genauigkeit (Fehler im Interpolanten) erheblich weniger Gitterpunkte benötigt.

Das quadratische Schema im linken Teil der Abbildung 2 zeigt in zwei Raumdimensionen ein theoretisch unendliches Tableau der Supports der stückweise bilinearen Ansatzfunktionen, die ein volles Gitter liefern. Jeder einzelne Unterraum  $W_{\mathbf{l}} = W_{l_1,l_2}$  besitzt dabei nach (2.11)  $2^{l_1+l_2-2}$  verschiedene Basisfunktionen. Aus Lemma 2.3 folgt, daß der Beitrag eines jeden Unterrums

<sup>4</sup>Hier und im folgende bedeutet ein hochgestelltes “V” stets einen Bezug zu vollen Gittern, während die Abkürzung “S” stets im Bezug zu dünnen Gittern (sparse grids) steht – eine Verwechslungsgefahr zwischen hochgestelltem V und der Bezeichnung  $V$  für Räume sollte nicht bestehen.

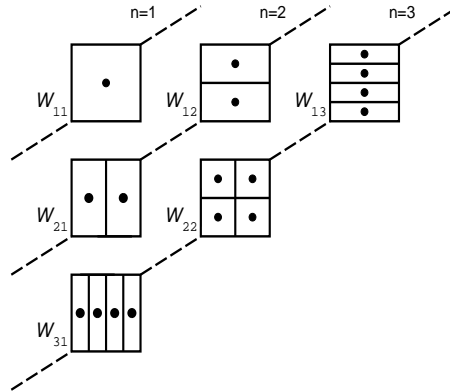


Abbildung 3: Träger der Basisfunktionen in zwei Raumdimensionen (dünnes Gitter, nur Basisfunktionen im Inneren von  $\bar{\Omega}$ )

$W_{l_1, l_2}$  mit  $l_1 + l_2 = c = \text{constant}$  zum Interpolanten von  $u$  von derselben Größenordnung  $\mathcal{O}(2^{-2c})$  bezüglich der  $L_2$ - und der  $L_\infty$ -Norm ist. Führt man diese Diskussion in mehr als zwei Dimensionen, so ergeben sich aus Lemma 2.3 sofort analoge Aussagen.

Aufgrund dieser Eigenschaften und einer Kosten-/Nutzen-Analyse (d.h. dem Vergleich des verursachten Rechenaufwands durch einen Teilraum zu seinem Beitrag zum Interpolanten) scheint es sinnvoll, statt einem quadratischem Schema ein dreieckiges Schema an Unterräumen zu verwenden, wie es in Abbildung 3 angedeutet ist. Somit werden nur diejenigen Teilräume zur Darstellung des Interpolanten verwendet, die einen bestimmten Minimalbeitrag hierzu liefern. Das "Weglassen" von Unterräumen, die über einen Großteil der Ansatzfunktionen eines vollen Gitters verfügen, jedoch nur noch einen kleinen Beitrag zum Interpolanten liefern und somit sehr teuer sind, führt auf dünne Gitter (vgl. [Zen90, Gri91a, Gri91b, Bun92, BuDo97]). Ein Beispiel für ein dünnes Gitter sieht man rechts in Abbildung 4.

Die  $d$ -linearen Ansatzräume dünner Gitter sind definiert vermöge

$$V^{n, S, d} = \bigcup_{|\mathbf{l}|_1 \leq n} W_{\mathbf{l}} . \quad (2.26)$$

Die Abbildung 4 verdeutlicht, daß dünne Gitter erheblich weniger Gitterpunkte besitzen als vergleichbare volle Gitter. Analog wie im Fall voller Gitter können Funktionen auf dünnen Gittern über

$$u^{n, S, d}(\mathbf{x}) = I^{n, S, d} \circ u(\mathbf{x}) := \sum_{|\mathbf{l}|_1 \leq n} \sum_{\mathbf{i} \in \mathbf{I}_{\mathbf{l}}} u_{\mathbf{l}, \mathbf{i}} \cdot \varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) \quad (2.27)$$

dargestellt werden, wobei  $I^{n, S, d} : V \rightarrow V^{n, S, d}$  der Interpolationsoperator des in (2.5) definierten Raums  $V$  auf den Dünngitterraum  $V^{n, S, d}$  des Le-



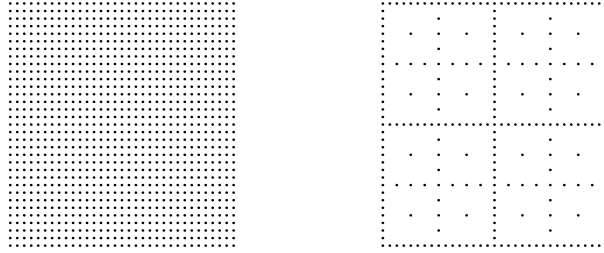


Abbildung 4: Volles und dünnes Gitter auf Level 5 in 2 Raumdimensionen

vels  $n$  in  $d$  Dimensionen darstellt. Ein Vergleich der beiden Räume (2.24) und (2.26) bzw. der Darstellung von Funktionen (2.25) und (2.27) in diesen Räumen zeigt, daß der einzige Unterschied in der Definition der Räume die unterschiedliche Wahl der Norm für die Auswahl zulässiger Unterräume ist.

Nachfolgend wollen wir einige Eigenschaften von dünnen Gittern angeben, wobei wir uns wie bereits zuvor an [Bun92, BuGr97, Bun98] orientieren.

**Lemma 2.4** Die Anzahl  $N^{n,S,d} := |V^{n,S,d}|$  der Gitterpunkte (im Inneren von  $\bar{\Omega}$ ) ergibt sich zu

$$\begin{aligned} N^{n,S,1} &= 2^n - 1 \\ N^{n,S,d} &= \sum_{i=0}^{n-1} 2^i \cdot N^{n-i,S,d-1} \quad (d \geq 2) \end{aligned}$$

**Beweis:** siehe [Bun92], Satz 3.1 □

Läßt man entgegen der bisherigen Diskussion auch Funktionen zu, die am Rand nicht verschwinden, berücksichtigt man damit folglich auch diejenigen Teilräume  $W_1$  mit  $1 \in \mathbf{N}_0^d$ , die aus Randfunktionen bestehen, so gibt Lemma 2.5 die Zahl der Punkte für diesen Fall an.

**Lemma 2.5** Die Anzahl  $M^{n,S,d}$  der Gitterpunkte eines dünnen Gitters (einschließlich Randpunkten) auf  $\bar{\Omega}$  ergibt sich zu

$$\begin{aligned} M^{n,S,1} &= 2^n + 1 \\ M^{n,S,d} &= 3 \cdot M^{n,S,d-1} + \sum_{i=1}^{n-1} 2^i \cdot M^{n-i,S,d-1} \quad (d \geq 2) \end{aligned}$$

Folglich gilt  $M^{n,S,d} = \mathcal{O}(2^n \cdot n^{d-1})$ .

**Beweis:** siehe [Bun92], Satz 3.3 □

Aus den beiden vorangegangenen Lemmata 2.4 und 2.5 geht hervor, daß dünne Gitter im Vergleich zu vollen Gittern erheblich weniger Gitterpunkte benötigen. Die Zahl der Gitterpunkte eines dünnen Gitters steigt unabhängig von der Dimension asymptotisch mit dem Faktor 2. Die Frage nach der Güte der Ansatzräume der dünnen Gitter beantwortet das folgende Lemma.

**Lemma 2.6** *Sei  $u^{n,S,d} \in V^{n,S,d}$  der Interpolant der Funktion  $u \in \mathcal{X}^{\mathbf{r},d}(\bar{\Omega})$  mit  $\mathbf{r} \geq \mathbf{2}$ . Für den Interpolationsfehler auf dünnen Gittern, gemessen in der  $L_\infty$ -Norm, der  $L_2$ -Norm sowie der Energienorm, gelten die Abschätzungen:*

$$\begin{aligned} \|u - u^{n,S,d}\|_\infty &\leq \frac{2 \cdot |u|_{\mathbf{2},\infty}}{8^d} \cdot 2^{-2n} \cdot A(d, n) = \mathcal{O}(h_n^2 \cdot n^{d-1}) \\ \|u - u^{n,S,d}\|_2 &\leq \frac{2 \cdot |u|_{\mathbf{2},2}}{12^d} \cdot 2^{-2n} \cdot A(d, n) = \mathcal{O}(h_n^2 \cdot n^{d-1}) \quad (2.28) \\ \|u - u^{n,S,d}\|_E &\leq \frac{d \cdot |u|_{\mathbf{2},\infty}}{2 \cdot 3^{(d-1)/2} \cdot 4^{d-1}} \cdot 2^{-n} = \mathcal{O}(h_n) \end{aligned}$$

Dabei ist  $h_n := 2^{-n}$  und

$$A(d, n) := \sum_{k=0}^{d-1} \binom{n+d-1}{k}$$

**Beweis:** siehe [BuGr97], Theorem 1 □

Hieraus wird ersichtlich, daß sich der Interpolationsfehler sowohl in der Maximumsnorm als auch in der  $L_2$ -Norm lediglich um einen logarithmischen Faktor verschlechtert. Dünne Gitter besitzen somit bei einer erheblich geringeren Anzahl von Gitterpunkten annähernd dieselbe Genauigkeit des Interpolanten.

Die Tabelle 1 gibt einen Überblick über die Ergebnisse der Lemmata 2.4 bis 2.6. Dabei enthalten die beiden letzten Zeilen der Tabelle einen Vergleich der Anzahl der Gitterpunkte eines dünnen und eines vollen Gitters im Fall von drei Dimensionen. In der genannten Tabelle setzen wir  $N = h^{-1}$ , d.h.  $N$  gibt die Zahl der (Rand-)Punkte an, um die Maschenweite  $h$  zu erhalten. Dabei betrachtet man bei einem regulären dünnen Gitter in  $d$  Dimensionen stets eine (eindimensionale) Kante des Gitters, um  $N$  zu erhalten, wobei im Falle eines vollen Gitters  $N$  hier die Zahl der Gitterpunkte pro Raumdimension darstellt.

## 2.4 Adaptive dünne Gitter

Sind die durch  $\mathcal{X}^{2,d}(\bar{\Omega})$  geforderten Glattheitsvoraussetzungen nicht erfüllbar, wie es beispielsweise bei Singularitäten oder Grenzschichten der Fall

Tabelle 1: Eigenschaften dünner Gitter ( $N = h^{-1}$ )

Eigenschaft	reguläres dünnes Gitter	volles Gitter
Genauigkeit bzgl. $L_2$	$\mathcal{O}(h^2 \cdot (\log(h^{-1})^{d-1}))$	$\mathcal{O}(h^2)$
Genauigkeit bzgl. $L_\infty$	$\mathcal{O}(h^2 \cdot (\log(h^{-1})^{d-1}))$	$\mathcal{O}(h^2)$
# Gitterpunkte	$\mathcal{O}(N \cdot (\log(N)^{d-1}))$	$\mathcal{O}(N^d)$
$d = 3, N = 2^{10}$	$1.1 \cdot 10^5$	$1.1 \cdot 10^9$
$d = 3, N = 2^{20}$	$3.3 \cdot 10^8$	$1.2 \cdot 10^{18}$

ist, kann man adaptive (dünne) Gitter verwenden. Dabei ist anzumerken, daß der gewählte Ansatz über die Tensorproduktdarstellung von Funktionen bzw. der Ansatzräume Adaptivität auf einfache Weise unterstützt. Dazu geht man von der (endlichen) Darstellung (2.27) über zu der Darstellung

$$u^{\varepsilon, \|\cdot\|}(\mathbf{x}) = \sum_{\mathbf{l}, \mathbf{i}: \|u_{\mathbf{l}, \mathbf{i}} \cdot \varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})\| \geq \varepsilon} u_{\mathbf{l}, \mathbf{i}} \cdot \varphi(\mathbf{x}) \quad , \quad (2.29)$$

von der man a priori nicht weiß, ob sie endlich oder unendlich ist. Dabei werden in (2.29) nur diejenigen Basisfunktionen in Betracht gezogen, deren Beitrag in der gewählten Norm  $\|\cdot\|$  größer als eine vorgegebene Genauigkeits-schranke  $\varepsilon$  ist. Zum einen erhält man je nach Wahl der Norm verschiedene Gitter, zum anderen kann die Wahl der Norm entscheiden, ob die Summe (2.29) endlich oder unendlich ist (vgl. weiter unten). Als heuristisches Kriterium für die Auswahl einer Basisfunktion könnte hierbei

$$\|u_{\mathbf{l}, \mathbf{i}} \varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})\| = |u_{\mathbf{l}, \mathbf{i}}| \cdot \|\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})\| =: |u_{\mathbf{l}, \mathbf{i}}| \cdot \gamma_{\mathbf{l}} \quad (2.30)$$

dienen, wobei sich die  $\gamma_{\mathbf{l}}$  für die Basisfunktionen auf einem festen Level für die verschiedenen Normen zu

$$\gamma_{\mathbf{l}} = \begin{cases} 1 & , \text{ falls } \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\| = \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\|_\infty \\ 2^{-|\mathbf{l}|_1} & , \text{ falls } \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\| = \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\|_{L_1} \\ \left(\frac{2}{3}\right)^{d/2} \cdot 2^{-|\mathbf{l}|_1/2} & , \text{ falls } \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\| = \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\|_{L_2} \\ \left(2 \cdot \left(\frac{2}{3}\right)^{d-1} \cdot 2^{-|\mathbf{l}|_1} \cdot \sum_{j=1}^d 2^{2l_j}\right)^{\frac{1}{2}} & , \text{ falls } \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\| = \|\varphi_{\mathbf{l}, \mathbf{i}}(x)\|_E \end{cases} \quad (2.31)$$

ergeben. Dabei stützt sich das Kriterium (2.30) auf das Verhalten

$$|u_{\mathbf{l}, \mathbf{i}}| \leq \frac{|u|_\infty}{2^d} \cdot \prod_{j=1}^d h_j^2$$

des hierarchischen Überschusses (vgl. [Bun92]), wobei für diese Abschätzung jedoch die Bedingung  $u \in \mathcal{X}^{\mathbf{r},d}(\bar{\Omega})$  bzw.  $u \in \mathcal{X}_0^{\mathbf{r},d}(\bar{\Omega})$  (mit  $\mathbf{r} \geq \mathbf{2}$ ) vorausgesetzt wird, man sich also im Fall glatter Funktionen befindet, bei denen eine adaptive Verfeinerung nicht notwendig ist. Betrachtet man Funktionen mit inhomogenen Dirichlet-Randbedingungen, so müssen auf dem Rand gewisse Modifikationen in obige Formeln eingebaut werden, die allerdings offensichtlich sind und daher hier nicht angeführt werden. Gesucht ist an dieser Stelle diejenige Norm, die in einem selbstadaptiven Lösungsprozeß die Gitterverfeinerung derart übernimmt, daß eine “optimale” Auswahl von Gitterpunkten vorgenommen wird, wobei der Begriff “optimal” sicherlich einer (hier aber nicht geführten) Diskussion bedarf.

In der Darstellung (2.29) kann es vorkommen, daß “Löcher” in der Darstellung auftauchen und somit nicht jeder Knoten seine Vaterknoten besitzt (vgl. Abbildung 5, siehe hierzu auch [Gri97]), was insbesondere in praktischen Anwendungen (einfache Implementierungen) jedoch wichtig wäre. Durch das Vorhandensein aller Vaterknoten wird sichergestellt, daß die Basistransformationen von nodaler Basis in hierarchische Basis und umgekehrt durchgeführt werden können. Da in (2.29) Löcher auftreten können, entspricht (2.29) damit einer Kompression, bei der sämtliche hierarchische Koeffizienten, die betragsmäßig kleiner als eine vorgegebene Genauigkeitsschranke  $\varepsilon$  sind, zu 0 gesetzt sind. Somit brauchen bei diesem Vorgehen für einen Knoten, der zur Darstellung einer gegebenen Funktion benötigt wird, seine Vaterknoten nicht unbedingt ebenfalls zur Darstellung notwendig sein. Wir schränken für die weiteren Betrachtungen die adaptive Darstellung (2.29) einer Funktion  $u$  auf

$$u^{\varepsilon, \|\cdot\|}(\mathbf{x}) = \sum_{\mathbf{l}, \mathbf{i} \in \mathcal{M}(u, \varepsilon, \|\cdot\|)} u_{\mathbf{l}, \mathbf{i}} \cdot \varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) \quad (2.32)$$

ein, wobei  $\mathcal{M}(u, \varepsilon, \|\cdot\|)$  die für die Funktion  $u$  bei der Wahl der Norm  $\|\cdot\|$  als Fehlerkriterium mit dem Abbruchkriterium  $\varepsilon$  aktive Menge an Basisfunktionen sei. Die aktive Menge der Basisfunktionen ist hierbei vermöge

$$\begin{aligned} \mathcal{M}(u, \varepsilon, \|\cdot\|) = & \{(\mathbf{l}, \mathbf{i}) : \|u_{\mathbf{l}, \mathbf{i}} \cdot \varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})\| \geq \varepsilon \vee \\ & \exists(\mathbf{k}, \mathbf{j}) \text{ mit } \mathbf{k} > \mathbf{l} : \|u_{\mathbf{k}, \mathbf{j}} \cdot \varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x})\| \geq \varepsilon \wedge \\ & \text{supp}(\varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x})) \cap \text{supp}(\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})) \neq \emptyset\} \end{aligned}$$

definiert. Auf den so erhaltenen adaptiven Gittern können wie auf den regulären dünnen Gittern die Basistransformationen von stückweise  $d$ -linearer, nodaler Basis in stückweise  $d$ -lineare, hierarchische Basis sowie deren Inversen angewendet werden. Aus der Darstellung (2.32) ist auch sofort ersichtlich, daß die Identität

$$u(\mathbf{x}) = u^{0, \|\cdot\|}(\mathbf{x})$$

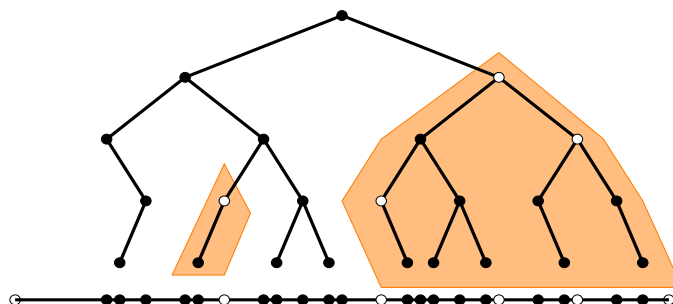


Abbildung 5: Baumdarstellung eines adaptiven Gitters in einer Raumdimension (leere Knoten stellen Knoten dar, deren hierarchischer Überschuß kleiner als  $\varepsilon$  ist)

unabhängig von der verwendeten Norm  $\|\cdot\|$  gilt.

Nachfolgend wollen wir in einer kurzen Diskussion Probleme behandeln, die bei der Wahl verschiedener Normen in (2.29) und (2.32) auftreten können.

Die einfachste Wahl für  $\|\cdot\|$  bzw. für  $\gamma_1$  in (2.30) ist sicherlich die Maximumnorm, denn hier besteht der Fehlerindikator allein aus dem hierarchischen Koeffizienten. Damit stellt sie von den vorgestellten Möglichkeiten gleichzeitig diejenige dar, bei denen der hierarchische Überschuß nicht gedämpft wird. Diese Wahl des Fehlerindikators erweist sich beispielsweise bei Sprungfunktionen als ungeeignet. Zur Verdeutlichung dient hier die eindimensionale Sprungfunktion

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

als Beispiel auf  $[-c, c]$  mit  $0 < c \in \mathbb{R}$ . Für diese Funktion gilt  $u(-c) = 0$  und  $u(c) = 1$ . Eine kurze Rechnung enthüllt, daß die hierarchischen Koeffizienten an den Stellen  $x_k = -2^{-k}$  mit  $k \geq 1$  stets  $-\frac{1}{2}$  betragen und hiermit das Schema (2.29) unendlich ist. Graphisch veranschaulicht dieses Dilemma Abbildung 6. In dieser Graphik wird die Sprungfunktion in dem Bereich  $[-c, c]$  durch die hierarchische Basis approximiert. Der Sprung der Sprungfunktion liegt bei unserer Wahl hierbei in der Intervallmitte. An derjenigen Stelle, an der jeweils der senkrechte Strich eingezeichnet ist, ist die nächste Basisfunktion hinzuzufügen. Dabei wird ersichtlich, daß der Wert, mit dem jede Basisfunktion zu gewichten ist, stets  $0 - \frac{1}{2}(0 + 1) = -\frac{1}{2}$  beträgt.

Nimmt man bei dieser Funktion beispielsweise die Gewichtung mittels der  $L_1$ -Norm oder der  $L_2$ -Norm vor, so reduziert sich der Faktor  $-\frac{1}{2}$  um die levelweise kleiner werdenden Normen der Basisfunktionen und man erhält somit eine endliche Darstellung der Funktion. Die Gewichtungen der Basisfunktionen in den Normen  $\|\cdot\|_{L_1}$  und  $\|\cdot\|_{L_2}$  können folglich als eine Art Dämpfungsfaktor des hierarchischen Überschusses, also dem entsprechen-

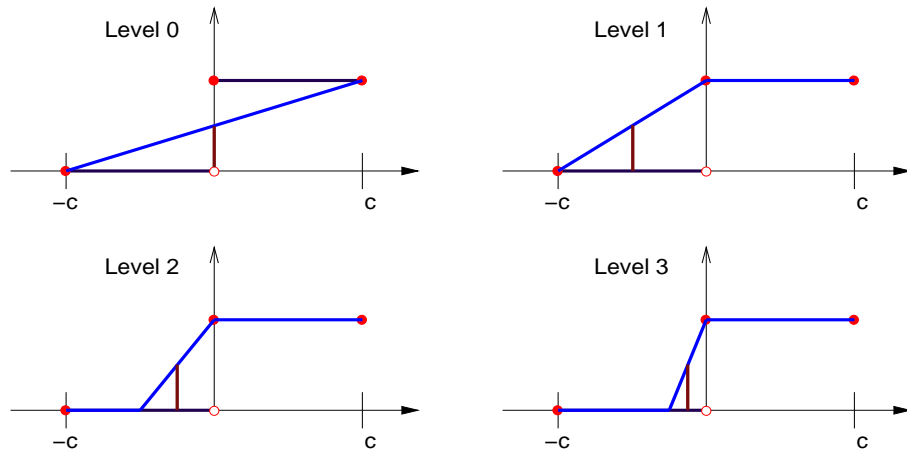


Abbildung 6: Sprungfunktion zusammen mit dem Interpolanten auf sukzessiven Levels, wobei der hierarchische Überschuß ( $\equiv$  Höhe des jeweils markierten senkrechten Strichs) der jeweils hinzuzufügenden Basisfunktion gleichbleibend ist

den Gewicht in der Maximumsnorm, interpretiert werden. Dennoch sind die  $L_1$ -Norm oder die  $L_2$ -Norm kein Allheilmittel, denn Probleme bei der Repräsentation einer gegebenen Funktion können bei einer beliebigen Wahl der Norm auftreten, wie folgendes Beispiel einer  $C^\infty(\Omega)$  Funktion zeigt: sei  $u(x) = \sin(x)$  und  $\Omega = [0, \pi]$ . Sowohl an den Intervallgrenzen als auch in der Intervallmitte ist  $u(x)$  identisch 0. Folglich berechnet sich der hierarchische Koeffizient ebenfalls zu 0 und ein adaptiver Prozeß, der eines der Kriterien in (2.31) verwendet, stoppt, unabhängig davon, welche Norm in (2.31) in Betracht gezogen wird.

Das Problem, das hier beobachtet wird, ist jenes, daß für jedes Fehlerkriterium stets eine Funktion konstruiert werden kann, für die das gewählte Fehlerkriterium versagt. Will man ein adaptives dünnes Gitter samt Interpolanten für eine gegebene Funktion konstruieren, die das Kriterium (2.32) erfüllt, so erweist sich jedoch in aller Regel die  $L_2$ -Norm als vernünftige Wahl für  $\gamma_1$ . In einem selbstadaptiven Algorithmus, der auf zuvor berechnete Ergebnisse als Entscheidungsgrundlage angewiesen ist, kann eine adaptive Verfeinerung anhand der Größen

$$\frac{2 \cdot |u|_{2,\infty}}{8^d} \cdot 2^{-2n} \cdot A(d, n) \quad \text{und} \quad \frac{2 \cdot |u|_{2,2}}{12^d} \cdot 2^{-2n} \cdot A(d, n)$$

durchgeführt werden. Ist dabei der hierarchische Überschuß an einem Knoten kleiner als eines der beiden Kriterien (dabei wird eines der beiden Kriterien für alle Gitterpunkte fest ausgewählt) kleiner als eine vorgegebene Schranke  $\eta$ , so wird an dem Knoten verfeinert. Obige Terme kommen bereits bei

---

den Abschätzungen des Interpolationsfehlers des Dünngitter-Interpolanten in Lemma 2.6 vor. Numerische Ergebnisse, die vermöge eines selbstadaptiven Algorithmus gewonnen werden können, der auf einer geringfügigen Modifikation der soeben genannten Größen aufbaut, befinden sich in dem Abschnitt 7 (“Numerische Resultate”). Die Beschreibung des angesprochenen Algorithmus zusammen mit dem verwendeten Verfeinerungskriterium befindet sich in dem Abschnitt 5 (“Anmerkungen zur Lösung partieller Differentialgleichungen”).

### 3 Eine Algebra für Operatoren auf Funktionen

#### 3.1 Einfache algebraische Grundlagen

In einem Lösungsprozeß, der auf adaptiven Gittern arbeiten kann, treten mitunter folgende Probleme auf: will man zwei Funktionen miteinander addieren oder multiplizieren, so werden diese Operationen im diskreten Fall im allgemeinen punktweise durchgeführt. Werden die beteiligten Funktionen jedoch mit unterschiedlichen adaptiven Gittern dargestellt, so muß man im allgemeinen für die Ergebnisfunktion diejenigen Gitterpunkten, die nur zur Darstellung einer der beteiligten Funktionen nötig sind, zur Darstellung berücksichtigen, wobei man nach Beendigung der Operation gegebenenfalls eine Kompression durchführen und somit wieder einige Gitterpunkte bzw. Basisfunktionen vernachlässigen kann. Geht man dabei wie eben beschrieben vor, so stellen die Punkte, die zur Darstellung der Ergebnisfunktion nötig sind, wiederum ein dünnes Gitter dar. Dieses entsteht aus der Vereinigung der Gitterpunkte der einzelnen dünnen Gitter, die zur Darstellung der Operanden nötig sind.

Es ist intuitiv klar, daß aus der Vereinigung zweier dünner Gitter (über demselben Gebiet  $\bar{\Omega}$ ) wiederum ein dünnes Gitter entsteht. Der Grund hierfür liegt in dem Unterraumschema begründet, über dem dünne Gitter definiert werden. Wird ein Unterraum  $W_1$  zur Darstellung eines dünnen Gitters benötigt, so sind ebenfalls alle Unterräume  $W_k$  mit  $k \leq 1$  in der Definition des betrachteten dünnen Gitters vertreten. Analoges gilt im Fall adaptiver dünner Gitter: wird eine Basisfunktion  $\varphi_{1,i}$  zur adaptiven Darstellung einer Funktion verwendet, so sind ebenso alle Basisfunktionen  $\varphi_{k,j}$  mit  $k \leq 1$  in der Definition des Dünngitterraums vertreten, für die  $\text{supp}(\varphi_{1,i}) \subset \text{supp}(\varphi_{k,j})$  gilt. Dies folgt sofort aus der Konstruktion der dünnen Gitter des vorangehenden Abschnitts 2.3. Die mengentheoretische Vereinigung zweier Dünngitterräume führt damit wieder auf einen Dünngitterraum. Dies ist gleichzeitig die theoretische Grundlage für die in Abschnitt 3.2 besprochene Algebra auf Funktionen und für die Operatoren des nachfolgenden Kapitels 4.

Die zuvor gemachte mengentheoretische Aussage läßt sich mit der Halbgruppentheorie beschreiben, die wir nachfolgend kurz skizzieren wollen. Dazu führen wir zunächst einige Definitionen aus [Mey80] an.

**Definition 3.1** [Mey80] *Ein Paar  $(G, \circ)$ , bestehend aus einer nichtleeren Menge  $G$  und einer assoziativen inneren Abbildung  $\circ$  auf  $G$ , heißt eine Halbgruppe.*

**Definition 3.2** [Mey80] *Eine Halbgruppe  $(G, \circ)$  heißt abelsch oder kommutativ, wenn  $a \circ b = b \circ a$  gilt für alle  $a, b \in G$ .*



**Definition 3.3** Sei  $G^S(\bar{\Omega})$  die Menge der dünnen Gitter auf dem Grundgebiet  $\bar{\Omega}$  und seien  $G_1, G_2 \in G^S(\bar{\Omega})$ . Die innere Verknüpfung  $\circ^S$  sei definiert über

$$\circ^S : \begin{cases} G^S(\bar{\Omega}) \times G^S(\bar{\Omega}) & \rightarrow G^S(\bar{\Omega}) \\ G_1 \circ^S G_2 & \mapsto G_3 \end{cases} \quad (3.1)$$

mit  $x_{1,i} \in G_3 \Leftrightarrow (x_{1,i} \in G_1 \vee x_{1,i} \in G_2)$ . Die Operation  $\circ^S$  heißt Dünngitter-Verknüpfung.

**Lemma 3.1**  $d$ -dimensionale dünne Gitter  $G \in G^S(\bar{\Omega})$  mit der punktweisen Dünngitter-Verknüpfung  $\circ^S$  von Gitterpunkten als innere Verknüpfung bilden eine abelsche Halbgruppe ( $d \in \mathbb{N}$ ).

**Beweis:** Der Beweis ist offensichtlich und wird daher ausgelassen.  $\square$

**Lemma 3.2** [Mey80] In einer abelschen Halbgruppe  $(G, \circ)$  ist eine Verknüpfung von je  $n$  Elementen  $a_1, \dots, a_n \in G$  unabhängig von der Reihenfolge der Elemente.

**Definition 3.4** [Mey80] Eine nichtleere Teilmenge  $U \subseteq G$  einer Halbgruppe  $(G, \circ)$  heißt Unterhalbgruppe (von  $G$ ), wenn für alle  $u, v \in U$  auch  $u \circ v$  in  $U$  liegt. Die Restriktion von  $\circ$  auf  $U \times U$  liefert dann eine innere Komposition auf  $U$ , diese ist assoziativ (denn  $\circ$  ist bereits assoziativ auf  $G$ ), d.h.  $(U, \circ)$  ist Halbgruppe.

### Bemerkung 3.1

Wendet man Lemma 3.4 auf dünne Gitter an, so folgt beispielsweise, daß für beliebiges, aber festes  $d$ , die Menge  $U$  der dünnen Gitter in  $d$  Dimensionen mit einem maximalen Level  $L$  in jede Koordinatenrichtung eine Unterhalbgruppe bezüglich  $\circ^S$  bilden.

## 3.2 Eine Algebra für Operatoren auf Funktionen

Die Idee einer Algebra für Operatoren auf Funktionen besteht darin, einen Satz von einfachen, grundlegenden Operatoren zusammenzustellen, mit deren Hilfe kompliziertere Operatoren zusammengesetzt werden können. Zu solchen einfachen Operatoren zählen wir beispielsweise die skalare Multiplikation, die Addition und Subtraktion, die Multiplikation und Division und ähnliche grundlegenden Operatoren, die vornehmlich aus dem Bereich der linearen Algebra entnommen sind und wie sie unter anderem in wohlbekannteren iterativen Lösern benötigt werden.

Ein weiterer Aspekt für die Entwicklung der Algebra ist der zugrundeliegende Tensorproduktansatz. Ihn kann man ausnützen, um verschiedene  $d$ -dimensionale Operatoren durch die  $d$ -malige Anwendung eindimensionaler

Operatoren zu formulieren. Folglich müssen so operatorteknisch lediglich eindimensionale Operatoren betrachtet und untersucht werden, um einen  $d$ -dimensionalen Operator zu studieren. Wie wir bei der Einführung der Finiten Differenzen auf dünnen Gittern später in Kapitel 4 sehen werden, ist dieser Ansatz essentiell für eine einfache Konstruktion (und auch Notation) speziell dieser Operatoren.

Bei der Anwendung einer der oben genannten Operatoren stellt sich die Frage, welches  $\varepsilon$  in (2.32) die Genauigkeit ist, mit der das Ergebnis des jeweiligen Operators dargestellt wird. Genauer: gegeben seien zwei Normen  $\|\cdot\|$  und  $|||\cdot|||$ , zwei Funktionen  $u_1(\mathbf{x})$  und  $u_2(\mathbf{x})$  sowie ihre Dünngitterinterpolanten  $u_1^{\varepsilon_1, |||\cdot|||}(\mathbf{x})$  und  $u_2^{\varepsilon_2, |||\cdot|||}(\mathbf{x})$  (vgl. 2.32). Ferner gelte

$$|||u_1(\mathbf{x}) - u_1^{\varepsilon_1, |||\cdot|||}(\mathbf{x})||| < \eta_1 \quad \wedge \quad |||u_2(\mathbf{x}) - u_2^{\varepsilon_2, |||\cdot|||}(\mathbf{x})||| < \eta_2$$

Die Frage lautet nun: für welches  $\eta_3$  gilt  $|||u_3(\mathbf{x}) - u_3^{\varepsilon_3, |||\cdot|||}(\mathbf{x})||| < \eta_3$ , wenn  $u_3(\mathbf{x})$  das Ergebnis des betrachteten Operators angewandt auf  $u_1(\mathbf{x})$  und  $u_2(\mathbf{x})$  ist und  $u_3^{\varepsilon_3, |||\cdot|||}(\mathbf{x})$  den entsprechenden Dünngitterinterpolanten darstellt. Zudem stellt sich die Frage, für welches  $\varepsilon_3$  die gemachte Aussage gilt<sup>5</sup>. Dabei können die beiden Normen  $\|\cdot\|$  und  $|||\cdot|||$  unabhängig voneinander gewählt werden und somit verschieden sein. Hierbei ist  $\|\cdot\|$  diejenige Norm, bezüglich der die Darstellung in (2.32) gewonnen wird, und die Norm  $|||\cdot|||$  stellt nachfolgend diejenige Norm dar, bezüglich der der jeweilige Fehler gemessen werden soll. Somit sind die Aussagen der nachfolgenden Lemmata insbesondere für die Maximumsnorm, die  $L_2$ -Norm, die  $L_1$ -Norm sowie die Energienorm gültig, wobei die Kombination der Normen für  $\|\cdot\|$  und  $|||\cdot|||$  beliebig ist. Es ist jedoch offensichtlich, daß bei festem  $\varepsilon$  bzw.  $\varepsilon_i$  für  $i = 1, 2$  in der Darstellung (2.32) durch die Wahl der jeweiligen Norm die in den Lemmata nachfolgend verwendeten Schranken  $\eta$  bzw.  $\eta_i$  unterschiedlich ausfallen und mit den verwendeten Normen gekoppelt sind.

Nachfolgend finden sich entsprechende Lemmata, die für spezielle Operatoren die aufgeworfenen Fragestellungen untersuchen. Dabei wird stets davon ausgegangen, daß die betrachteten Funktionen die Glattheitsvoraussetzungen der dünnen Gitter erfüllen (vgl. Abschnitt 2.1 sowie [Bun92]). Zudem wird nachfolgend vorausgesetzt, daß die Interpolanten  $u^{\varepsilon, |||\cdot|||}(\mathbf{x})$  nach (2.32) konstruiert sind.

**Lemma 3.3** *Sei  $u_c(\mathbf{x}) = c \cdot u(\mathbf{x})$  auf  $\Omega$  mit  $c \in \mathbb{R}$  und es gelte*

$$|||u(\mathbf{x}) - u^{\varepsilon, |||\cdot|||}(\mathbf{x})||| < \eta \quad .$$

*Dann folgt:*

$$|||u_c(\mathbf{x}) - u_c^{\varepsilon, |||\cdot|||}(\mathbf{x})||| < \eta_c := |c| \cdot \eta$$

<sup>5</sup>Die aufgeworfenen Fragen sind für Operatoren gestellt, die zwei Funktionen verknüpfen. Analoge Fragestellungen sind jedoch auch für Operatoren zu stellen, die nur auf einer oder mehr als zwei Funktionen operieren.

**Beweis:**

$$\begin{aligned} \| |u_c(\mathbf{x}) - u_c^{c \cdot \varepsilon, \|\cdot\|}(\mathbf{x})| \| &= \| |c \cdot u(\mathbf{x}) - c \cdot u^{\varepsilon, \|\cdot\|}(\mathbf{x})| \| \\ &= |c| \cdot \| |u(\mathbf{x}) - u^{\varepsilon, \|\cdot\|}(\mathbf{x})| \| \\ &\leq |c| \cdot \eta =: \eta_c \end{aligned}$$

Dabei gilt das erste Gleichheitszeichen, da sich der Parameter  $c$ , mit dem die Funktionswerte (nodale Werte) multipliziert werden, direkt auf den hierarchischen Überschuss überträgt. In einer Raumdimension berechnet man den hierarchischen Überschuss von  $u_c(x)$  an der Stelle  $x_{l,i}$  mittels

$$\begin{aligned} (u_c)_{l,i} &= c \cdot u(x_{l,i}) - \frac{c \cdot u(x_{l,i} - h_l) + c \cdot u(x_{l,i} + h_l)}{2} \\ &= c \cdot \left( u(x_{l,i}) - \frac{u(x_{l,i} - h_l) + u(x_{l,i} + h_l)}{2} \right) \\ &= c \cdot u_{l,i} \end{aligned}$$

als das  $c$ -fache des hierarchischen Überschusses  $u_{l,i}$  der Funktion  $u(x)$  an der angegebenen Stelle. Folglich besitzt die Funktion  $u_c(\mathbf{x})$  dieselben Gitterpunkte bzw. Basisfunktionen für ihre Repräsentation in der hierarchischen Basis wie die Funktion  $u(x)$ , wenn zum einen  $c \cdot \varepsilon$  und zum anderen  $\varepsilon$  als Kriterium in der Formel (2.32) für die verwendete Norm  $\|\cdot\|$  benutzt wird.  $\square$

**Lemma 3.4** *Seien die Funktionen  $u_1(\mathbf{x})$  und  $u_2(\mathbf{x})$  gegeben sowie  $u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})$  und  $u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})$  ihre Dünngitterinterpolanten, für die die Abschätzungen*

$$\| |u_1(\mathbf{x}) - u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})| \| < \eta_1 \quad \wedge \quad \| |u_2(\mathbf{x}) - u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})| \| < \eta_2$$

*gelten. Dann folgt für  $u_3(\mathbf{x}) = u_1(\mathbf{x}) \pm u_2(\mathbf{x})$  die Abschätzung*

$$\| |u_3(\mathbf{x}) - u_3^{\varepsilon_3, \|\cdot\|}(\mathbf{x})| \| < \eta_3$$

*mit  $\eta_3 := \eta_1 + \eta_2$ . Dabei ist  $u_3^{\varepsilon_3, \|\cdot\|}(\mathbf{x}) = u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \pm u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})$  und  $\varepsilon_3 = \varepsilon_1 + \varepsilon_2$ .*

**Beweis:**

$$\begin{aligned} \| |u_3(\mathbf{x}) - u_3^{\varepsilon_3, \|\cdot\|}(\mathbf{x})| \| &= \| |(u_1(\mathbf{x}) \pm u_2(\mathbf{x})) - (u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \pm u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}))| \| \\ &= \| |(u_1(\mathbf{x}) - u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})) \pm (u_2(\mathbf{x}) - u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}))| \| \\ &\leq \| |u_1(\mathbf{x}) - u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})| \| + \| |u_2(\mathbf{x}) - u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})| \| \\ &\leq \eta_1 + \eta_2 := \eta_3 \end{aligned}$$

Zudem haben wir

$$\begin{aligned} \| |u_{3,1,i} \cdot \varphi_{1,i}| \| &= \| |(u_{1,1,i} \pm u_{2,1,i}) \cdot \varphi_{1,i}| \| \\ &\leq \| |u_{1,1,i} \cdot \varphi_{1,i}| \| + \| |u_{2,1,i} \cdot \varphi_{1,i}| \| \\ &\leq \varepsilon_1 + \varepsilon_2 =: \varepsilon_3 \end{aligned}$$

$\square$

**Bemerkung 3.2**

Die Addition (Subtraktion) zweier Funktionen kann genauso wie die Skalarmultiplikation sowohl auf hierarchischen als auch auf nodalen Werten durchgeführt werden. Bei der Addition (Subtraktion) von Funktionen ist an Gitterpunkten, die nur für eine der beteiligten Funktionen zur Darstellung notwendig sind, die jeweils andere Funktion zu interpolieren.

Die Abschätzung in Lemma 3.4 mit  $\varepsilon_3$  ist scharf. Zum Nachweis hierzu kann man beispielsweise die beiden Funktionen  $u_1(\mathbf{x}) = u_2(\mathbf{x}) = x_1 \cdot (1 - x_1)$  auf  $\bar{\Omega}^d = [0, 1]^d$  betrachten. Nachrechnen ergibt, daß der hierarchische Koeffizient der Funktion  $u_3(\mathbf{x}) = u_1(\mathbf{x}) + u_2(\mathbf{x})$  an einem beliebigen Knoten des dünnen Gitters die Summe der hierarchischen Koeffizienten der Einzelsummanden ist.

**Lemma 3.5** *Seien die Voraussetzungen wie in Lemma 3.4. Dann folgt für  $u_3(\mathbf{x}) = u_1(\mathbf{x}) \cdot u_2(\mathbf{x})$  die Abschätzung*

$$\| \| u_3(\mathbf{x}) - u_3^{\varepsilon_3, \|\cdot\|}(\mathbf{x}) \| \| < \eta_3$$

mit

$$\eta_3 = \min \left\{ \| \| u_1(\mathbf{x}) \| \| \cdot \eta_2 + \| \| u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \| \| \cdot \eta_1, \right. \\ \left. \| \| u_2(\mathbf{x}) \| \| \cdot \eta_1 + \| \| u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \| \| \cdot \eta_2 \right\} \quad (3.2)$$

*Eine symmetrische, aber schwächere Abschätzung liefert*

$$\eta_3 = \max \left\{ \| \| u_2(\mathbf{x}) \| \|, \| \| u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \| \| \right\} \cdot \eta_1 \\ + \max \left\{ \| \| u_1(\mathbf{x}) \| \|, \| \| u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \| \| \right\} \cdot \eta_2 \quad (3.3)$$

*Dabei ist  $u_3^{\varepsilon_3, \|\cdot\|}(\mathbf{x}) = u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})$  auf der Indezmenge*

$$\mathcal{M}_3 := \mathcal{M}_1(u_1, \varepsilon_1, \|\cdot\|) \cup \mathcal{M}_2(u_2, \varepsilon_2, \|\cdot\|)$$

**Beweis:** Wir führen zwei verschiedene Abschätzungen für die Differenz  $u_3(\mathbf{x}) - u_3^{\varepsilon_3, \|\cdot\|}(\mathbf{x})$  durch, die jeweils durch Addition einer 0 gewonnen werden. Die beiden angesprochenen Abschätzungen, deren Ergebnisse in den Formeln (3.4) und (3.5) zu finden sind, liefern dabei die beiden Terme, deren Minimum in (3.2) zu nehmen ist. Die gewichtete Kombination der ersten beiden Abschätzungen führt dann auf die symmetrische Abschätzung in (3.3).

Für die erste Abschätzung addieren wir

$$0 = u_1(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) - u_1(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})$$

zur Differenz der exakten und der genäherten Lösung und erhalten

$$\begin{aligned}
& \left\| \left( u_1(\mathbf{x}) \cdot u_2(\mathbf{x}) \right) - \left( u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \right) \right\| = \\
& = \left\| \left( u_1(\mathbf{x}) \cdot u_2(\mathbf{x}) \right) - \left( u_1(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \right) + \right. \\
& \quad \left. \left( u_1(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \right) - \left( u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \right) \right\| \\
& \leq \|u_1(\mathbf{x})\| \cdot \|u_2(\mathbf{x}) - u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})\| + \|u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})\| \cdot \|u_1(\mathbf{x}) - u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})\| \\
& \leq \|u_1(\mathbf{x})\| \cdot \eta_2 + \|u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})\| \cdot \eta_1 \tag{3.4}
\end{aligned}$$

Analog erhalten wir

$$\begin{aligned}
& \left\| \left( u_1(\mathbf{x}) \cdot u_2(\mathbf{x}) \right) - \left( u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \right) \right\| = \\
& = \left\| \left( u_1(\mathbf{x}) \cdot u_2(\mathbf{x}) \right) - \left( u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2(\mathbf{x}) \right) + \right. \\
& \quad \left. \left( u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2(\mathbf{x}) \right) - \left( u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \right) \right\| \\
& \leq \|u_2(\mathbf{x})\| \cdot \|u_1(\mathbf{x}) - u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})\| + \|u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})\| \cdot \|u_2(\mathbf{x}) - u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})\| \\
& \leq \|u_2(\mathbf{x})\| \cdot \eta_1 + \|u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})\| \cdot \eta_2 \tag{3.5}
\end{aligned}$$

nach Addition von

$$0 = u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2(\mathbf{x}) - u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2(\mathbf{x}) \quad .$$

Nimmt man nun das Minimum der rechten Seiten der beiden Abschätzungen (3.4) und (3.5), so erhält man die Aussage (3.2). Addiert man beide Abschätzungen (3.4) und (3.5) und dividiert das Ergebnis anschließend durch 2, so erhält man die Abschätzung

$$\begin{aligned}
& \left\| \left( u_1(\mathbf{x}) \cdot u_2(\mathbf{x}) \right) - \left( u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) \right) \right\| \leq \\
& \leq \frac{1}{2} \left( \left( \|u_2(\mathbf{x})\| + \|u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})\| \right) \eta_1 + \left( \|u_1(\mathbf{x})\| + \|u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})\| \right) \eta_2 \right) \\
& \leq \max \left( \|u_2(\mathbf{x})\|, \|u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})\| \right) \eta_1 + \max \left( \|u_1(\mathbf{x})\|, \|u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})\| \right) \eta_2 ,
\end{aligned}$$

die die schwächere Aussage des Lemmas beweist.  $\square$

### Bemerkung 3.3

Da die Division  $g^{-1} \circ f$  mit  $g \neq 0$  auf  $\Omega$  einer Multiplikation von  $f$  mit der Funktion  $h = g^{-1}$  entspricht, beweist das Lemma 3.5 gleichzeitig die Aussage für die Operation "Division".

Die beiden Dünngitter-Interpolanten  $u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})$  und  $u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})$  besitzen die Darstellung

$$u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) = \sum_{\mathbf{l}, \mathbf{i} \in \mathcal{M}(u_1, \varepsilon_1, \|\cdot\|)} u_{\mathbf{l}, \mathbf{i}}^1 \cdot \varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})$$

sowie

$$u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) = \sum_{\mathbf{k}, \mathbf{j} \in \mathcal{M}(u_2, \varepsilon_2, \|\cdot\|)} u_{\mathbf{k}, \mathbf{j}}^2 \cdot \varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x}) .$$

Die Multiplikation

$$u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) = \sum_{\mathbf{l}, \mathbf{i} \in \mathcal{M}(u_1, \varepsilon_1, \|\cdot\|)} \sum_{\mathbf{k}, \mathbf{j} \in \mathcal{M}(u_2, \varepsilon_2, \|\cdot\|)} (u_{\mathbf{l}, \mathbf{i}}^1 \cdot u_{\mathbf{k}, \mathbf{j}}^2) \cdot \varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x}) \quad (3.6)$$

zweier stückweise  $d$ -linearer Funktionen führt nun zu stückweise quadratischen Funktionen. Folglich besitzt das Ergebnis keine Darstellung der Form (2.27). Das Produkt zweier Basisfunktionen  $\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x})$  kann man vermöge

$$\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x}) = \sum_{\mathbf{s}, \mathbf{t}} q_{\mathbf{s}, \mathbf{t}}^{\mathbf{l}, \mathbf{i}, \mathbf{k}, \mathbf{j}} \cdot \varphi_{\mathbf{s}, \mathbf{t}}(\mathbf{x}) \quad (3.7)$$

darstellen. Dabei gilt  $\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x}) = 0$  für  $\text{supp}(\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})) \cap \text{supp}(\varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x})) = \emptyset$ . Folglich braucht man nur Basisfunktionen zu untersuchen, bei denen der eindimensionale Anteil  $\varphi_{k_m, j_m}(x_m)$  bezüglich Raumdimension  $1 \leq m \leq d$  der Basisfunktion  $\varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x})$  ein Sohnknoten (bzw. ein Enkel oder Urenkel oder ...) des eindimensionalen Anteils  $\varphi_{l_m, i_m}(x_m)$  der Basisfunktion  $\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})$  bzw. umgekehrt ist. Der Support des Produkts der beiden Basisfunktionen ist dabei identisch dem Schnitt des Support von  $\varphi_{\mathbf{k}, \mathbf{j}}(\mathbf{x})$  und  $\varphi_{\mathbf{l}, \mathbf{i}}(\mathbf{x})$ . Auf diesem Support stellt das Produkt eine stückweise quadratische Funktion dar, die am Rand des Supports verschwindet.

Um zu untersuchen, wie die resultierenden stückweise definierten quadratischen Funktionen aufgebaut sind, ziehen wir uns an dieser Stelle kurz auf den eindimensionalen Fall zurück. In den Abbildungen 7 bis 9 sind die 3 verschiedenen Möglichkeiten aufgezeigt, die nachfolgend in der Berechnung der  $q_{\mathbf{s}, \mathbf{t}}^{\mathbf{l}, \mathbf{i}, \mathbf{k}, \mathbf{j}}$  eingehen werden.

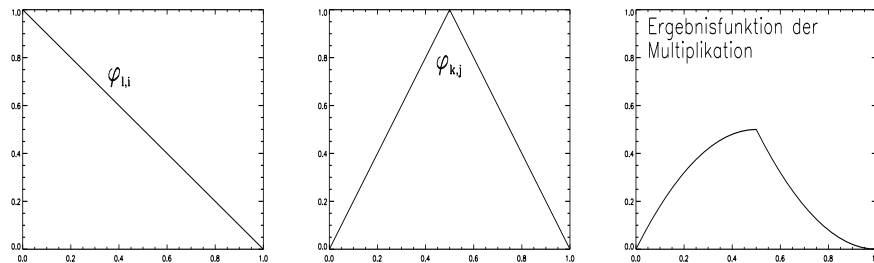


Abbildung 7: qualitatives Aussehen der Ergebnisfunktion (rechts) bei der Multiplikation zweier eindimensionaler Basisfunktionen  $\varphi_{l,i}$  (links) und  $\varphi_{k,j}$  (mitte) mit  $l < k$  und  $i \cdot 2^{k-l} < j < (i+1) \cdot 2^{k-l}$

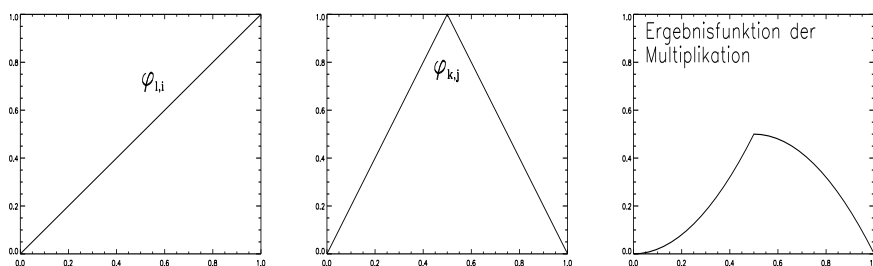


Abbildung 8: qualitatives Aussehen der Ergebnisfunktion (rechts) bei der Multiplikation zweier eindimensionaler Basisfunktionen  $\varphi_{l,i}$  (links) und  $\varphi_{k,j}$  (mitte) mit  $l < k$  und  $(i-1) \cdot 2^{k-l} < j < i \cdot 2^{k-l}$

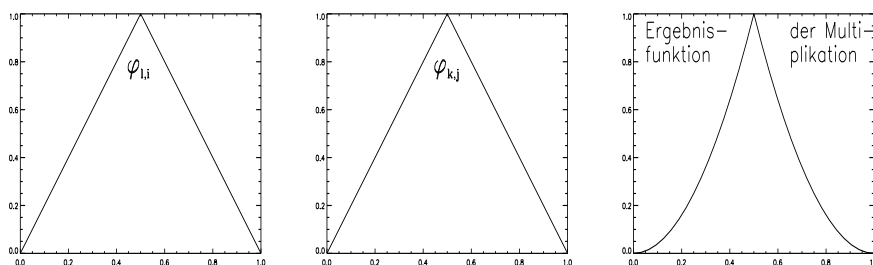


Abbildung 9: qualitatives Aussehen der Ergebnisfunktion (rechts) bei der Multiplikation zweier eindimensionaler Basisfunktionen  $\varphi_{l,i}$  (links) und  $\varphi_{k,j}$  (mitte) mit  $l = k$  und  $i = j$

Approximiert man die Ergebnisfunktionen der Abbildungen 7 bis 9 (jeweils rechte Graphik) mittels hierarchischer Basisfunktionen  $\varphi_{l,i}$ , so zeigt sich, daß in allen drei Fällen die Basisfunktion, die auf dem jeweils größten Level verwendet werden muß, mit einem positiven Gewicht versehen ist. Zur Bestimmung des Vorzeichens von zur Darstellung der Ergebnisfunktion notwendigen Basisfunktionen auf feineren Leveln dient nachfolgende Fallunterscheidung. Dabei wird vorausgesetzt, daß  $\xi < k$  und  $|\nu - j \cdot 2^{\xi-k}| < 2^{\xi-k}$  gilt, wobei die letzte Bedingung lediglich besagt, daß die Basisfunktion  $\varphi_{\xi,\nu}$  tatsächlich zur (exakten) Darstellung der Ergebnisfunktion benötigt wird.

- Abbildung 7:  $\text{sgn}(u_{\xi,\nu}) = 1$ , falls  $\nu < 2^{\xi-k} \cdot j$  und  $\text{sgn}(u_{\xi,\nu}) = -1$ , falls  $\nu > j \cdot 2^{\xi-k}$
- Abbildung 8:  $\text{sgn}(u_{\xi,\nu}) = 1$ , falls  $\nu > 2^{l-k} \cdot j$  und  $\text{sgn}(u_{\xi,\nu}) = -1$ , falls  $\nu < j \cdot 2^{\xi-k}$
- Abbildung 9: alle weiteren hierarchischen Koeffizienten haben negatives Vorzeichen

Für die nachfolgende Diskussion, die auch den  $d$ -dimensionalen Fall ab-

decken soll, definieren wir  $\mathbf{p}$  und  $\mathbf{r}$  vermöge

$$p_m = \max\{l_m, k_m\} \quad \text{für } 1 \leq m \leq d$$

und

$$r_m = \begin{cases} i_m & \text{falls } p_m \equiv l_m \text{ gilt} \\ j_m & \text{falls } p_m \equiv k_m \text{ gilt} \end{cases} ,$$

bleiben aber zunächst weiterhin bei der Betrachtung im Fall  $d = 1$ . Der hierarchische Überschuß fällt bei quadratischen Funktionen exakt mit dem Faktor  $\frac{1}{4}$  ab, so daß die Reihenentwicklung von  $\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k},\mathbf{j}}(\mathbf{x})$  vermöge des Funktionswertes von  $\varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k},\mathbf{j}}(\mathbf{x})$  im Mittelpunkt des Supports von  $\varphi_{\mathbf{p},\mathbf{r}}(\mathbf{x})$  angegeben werden kann, der sich zu

$$A_{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j}}(x_{\mathbf{l},\mathbf{i}}, x_{\mathbf{k},\mathbf{j}}) := \prod_{m=1}^d \left( 1 - \left| \frac{x_{l_m, i_m} - x_{k_m, j_m}}{h_{p_m}} \right| \right)$$

berechnet. Berücksichtigt man die zuvor durchgeführte Fallunterscheidung, so kann man die Entwicklung (3.7) in einer Dimension vermöge

$$\begin{aligned} \varphi_{l,i}(x) \cdot \varphi_{k,j}(x) &= \tag{3.8} \\ & A_{l,i,k,j}(x_{l,i}, x_{k,j}) \cdot \left\{ \varphi_{p,r}(x) + \operatorname{sgn}(p-l) \cdot \operatorname{sgn}(i \cdot 2^{k-l} - j) \cdot \left( \sum_{s=k+1}^{\infty} \sum_{t=1}^{2^{s-k-1}} \right. \right. \\ & \left. \left. \left( \frac{1}{4} \right)^{s-k} \cdot [\varphi_{s,(i-1) \cdot 2^{s-l} + 2t-1}(x) + \varphi_{s,(i-1) \cdot 2^{s-l} + 2t-1 + 2^{s-k}}(x)] \right) \right. \\ & + \operatorname{sgn}(p-k) \cdot \operatorname{sgn}(j \cdot 2^{l-k} - i) \cdot \left( \sum_{s=l+1}^{\infty} \sum_{t=1}^{2^{s-l-1}} \left( \frac{1}{4} \right)^{s-l} \cdot \right. \\ & \left. \cdot [\varphi_{s,(i-1) \cdot 2^{s-l} + 2t-1}(x) + \varphi_{s,(i-1) \cdot 2^{s-l} + 2t-1 + 2^{s-k}}(x)] \right) - \\ & \left. (1 + \operatorname{sgn}(l-k) \cdot \operatorname{sgn}(k-l)) \cdot \sum_{s=l+1}^{\infty} \sum_{t=1}^{2^{s-l}} \left( \frac{1}{4} \right)^{s-l} \varphi_{s,(i-1) \cdot 2^{s-l} + 2t-1}(x) \right\} \end{aligned}$$

darstellen, wobei  $l, i, k, j$  sowie  $p$  jeweils Level und Indizes in einer Raumdimension beschreiben. Dabei dienen die Faktoren  $\operatorname{sgn}(\cdot)$  dazu, alle drei Fälle der Fallunterscheidung mit einer einzigen Funktion für die Multiplikation zweier Basisfunktionen abzuhandeln. Sie steuern im wesentlichen nichts anderes als die richtige Vorzeichenwahl, dienen folglich nur zur Bestimmung von Koeffizienten  $+1$  oder  $-1$  für die jeweiligen Fälle. In Kurzschreibweise kann man (3.8) vermöge

$$\begin{aligned} \varphi_{l,i}(x) \cdot \varphi_{k,j}(x) &= \\ & A_{l,i,k,j}(x_{l,i}, x_{k,j}) \cdot \sum_{p,r} F_{l,i,k,j,p,r} \left( \frac{1}{4} \right)^{s-p} \varphi_{p,r}(x) \end{aligned}$$



darstellen, wobei die richtige Vorzeichenwahl des Koeffizienten der jeweiligen Basisfunktion durch  $F_{l,i,k,j,p,r}$  bestimmt wird (Definition von  $F_{l,i,k,j,p,r}$ ).

Mit den Ergebnissen der Diskussion des eindimensionalen Falls können wir nun das entsprechende Ergebnis für (3.7) in  $d$  Dimensionen mittels

$$\begin{aligned} \varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k},\mathbf{j}}(\mathbf{x}) &= \sum_{\mathbf{s},\mathbf{t}} q_{\mathbf{s},\mathbf{t}}^{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j}} \cdot \varphi_{\mathbf{s},\mathbf{t}}(\mathbf{x}) \\ &= A_{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j}}(x_{\mathbf{l},\mathbf{i}}, x_{\mathbf{k},\mathbf{j}}) \cdot F_{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j},\mathbf{p},\mathbf{r}} \cdot \\ &\quad \sum_{\mathbf{s},\mathbf{t}=\mathbf{p},\mathbf{r}} \left(\frac{1}{4}\right)^{\mathbf{s}-\mathbf{p}} \cdot \varphi_{\mathbf{s},\mathbf{t}}(\mathbf{x}) \end{aligned} \quad (3.9)$$

angeben, wobei hier wie auch nachfolgend die Summe bezüglich  $\mathbf{s}$  in (3.9) eine unendliche Summe darstellt (vgl. hierzu (3.8)). Dabei steuert auch hier  $F_{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j},\mathbf{p},\mathbf{r}}$  das richtige Vorzeichen des Koeffizienten in der Darstellung bezüglich der hierarchischen Basis. An dieser Stelle seien die abkürzenden Schreibweisen  $\mathbf{LI}$  für  $\mathbf{l}, \mathbf{i} \in \mathcal{M}(u_1, \varepsilon_1, \|\cdot\|)$  sowie  $\mathbf{KJ}$  für  $\mathbf{k}, \mathbf{j} \in \mathcal{M}(u_2, \varepsilon_2, \|\cdot\|)$  vereinbart. Hiermit und unter den gemachten Annahmen ergibt sich somit die Multiplikation (3.6) der Interpolanten zu

$$\begin{aligned} u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x}) \cdot u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x}) &= \sum_{\mathbf{LI}} \sum_{\mathbf{KJ}} (u_{\mathbf{l},\mathbf{i}}^1 \cdot u_{\mathbf{k},\mathbf{j}}^2) \cdot \varphi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \cdot \varphi_{\mathbf{k},\mathbf{j}}(\mathbf{x}) \\ &\stackrel{(3.7)}{=} \sum_{\mathbf{LI}} \sum_{\mathbf{KJ}} u_{\mathbf{l},\mathbf{i}}^1 \cdot u_{\mathbf{k},\mathbf{j}}^2 \sum_{\mathbf{s},\mathbf{t}=\mathbf{p},\mathbf{r}} q_{\mathbf{s},\mathbf{t}}^{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j}} \cdot \varphi_{\mathbf{s},\mathbf{t}}(\mathbf{x}) \\ &= \sum_{\mathbf{LI}} \sum_{\mathbf{KJ}} \sum_{\mathbf{s},\mathbf{t}=\mathbf{p},\mathbf{r}} \left( u_{\mathbf{l},\mathbf{i}}^1 \cdot u_{\mathbf{k},\mathbf{j}}^2 \cdot q_{\mathbf{s},\mathbf{t}}^{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j}} \right) \cdot \varphi_{\mathbf{s},\mathbf{t}}(\mathbf{x}) \\ &= \sum_{\mathbf{LI}} \sum_{\mathbf{KJ}} \sum_{\mathbf{s},\mathbf{t}=\mathbf{p},\mathbf{r}} \left( u_{\mathbf{l},\mathbf{i}}^1 \cdot u_{\mathbf{k},\mathbf{j}}^2 \cdot A_{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j}}(x_{\mathbf{l},\mathbf{i}}, x_{\mathbf{k},\mathbf{j}}) \cdot \right. \\ &\quad \left. \cdot F_{\mathbf{l},\mathbf{i},\mathbf{k},\mathbf{j},\mathbf{p},\mathbf{r}} \cdot \left(\frac{1}{4}\right)^{\mathbf{s}-\mathbf{p}} \right) \cdot \varphi_{\mathbf{s},\mathbf{t}}(\mathbf{x}) \\ &=: \sum_{\mathbf{s},\mathbf{t}} u_{\mathbf{s},\mathbf{t}} \cdot \varphi_{\mathbf{s},\mathbf{t}}(\mathbf{x}) \ , \end{aligned} \quad (3.10)$$

wobei die Koeffizienten  $u_{\mathbf{s},\mathbf{t}}$  noch zu bestimmen sind. Der Einfachheit halber wollen wir uns wie zu Anfang des Kapitels 2 auf Funktionen beschränken, die am Rand  $\Gamma$  von  $\bar{\Omega}$  verschwinden. Damit sind ebenfalls die hierarchischen Koeffizienten von Rand-Basisfunktionen identisch 0. Somit gilt sowohl  $\mathbf{l} \geq \mathbf{1}$  als auch  $\mathbf{k} \geq \mathbf{1}$  in Formel (3.7) bzw. (3.6). Aus obiger Diskussion ergibt sich, daß eine Basisfunktion  $\varphi_{\mathbf{p},\mathbf{r}}(\mathbf{x})$  dann und nur dann zur Darstellung eines Produkts (3.7) benötigt wird, wenn beide am Produkt beteiligten Funktionen entweder mit  $\varphi_{\mathbf{p},\mathbf{r}}(\mathbf{x})$  identisch sind oder eine Vater-Funktion von  $\varphi_{\mathbf{p},\mathbf{r}}(\mathbf{x})$  sind. Bei der Bestimmung der gesuchten Koeffizienten handelt es

sich folglich auch um ein kombinatorisches Problem. Um nun den Koeffizienten  $u_{\mathbf{s},\mathbf{t}}$  in (3.10) mittels einer Formel angeben zu können, benötigen wir noch weitere Definitionen, die nachfolgend für eindimensionale  $p$  und  $r$ , also für  $p, r \in \mathbb{R}$ , angegeben werden.

Sei  $\mathbf{w}(p, r) \in \mathbb{R}^p$  für  $p, r \in \mathbb{R}$  derjenige Vektor, der den Index  $r$  in bitweiser Darstellung repräsentiert und eine Gesamtlänge von  $p$  Bits besitzt (Beispiel:  $\mathbf{w}(2, 3) = [1, 1]^T$ ,  $\mathbf{w}(4, 9) = [1001]^T$ ). Sei ferner  $\mathbf{Q}_{p,r} \in \mathbb{R}^p$  derjenige Vektor, der aus den Indizes der eindimensionalen Funktion  $\varphi_{p,r}(x)$  und deren Vaterknoten auf allen höheren Leveln aufgebaut ist, für den folglich  $\mathbf{Q}_{p,r}(r) = r$  gilt und die übrigen Einträge rekursiv vermöge

$$\mathbf{Q}_{p,r}(j) = \frac{1}{2} \left( \mathbf{Q}_{p,r}(j+1) + (1 - 2 \cdot \mathbf{w}(p, r)_{j+1}) \right)$$

für  $2 \leq j < r$  definiert sind. Dabei ist  $\mathbf{w}(p, r)_{j+1}$  das  $(j+1)$ -te Bit des Vektors  $\mathbf{w}(p, r)$  und zudem sei  $\mathbf{Q}_{p,r}(1) = 1$ .

Mit diesen Definitionen führen wir (3.10) fort und  $u_{\mathbf{s},\mathbf{t}}$  ergibt sich für den Fall einer Raumdimension zu

$$u_{\mathbf{s},\mathbf{t}} = \sum_{l=1}^s \sum_{k=1}^s A_{l, \mathbf{Q}_{s,t}(l), k, \mathbf{Q}_{s,t}(k)}(x_{l, \mathbf{Q}_{s,t}(l)}, x_{k, \mathbf{Q}_{s,t}(k)}) \cdot \left( \frac{1}{4} \right)^{s - \max\{l, k\}} \cdot u_{l, \mathbf{Q}_{s,t}(l)}^1 \cdot u_{k, \mathbf{Q}_{s,t}(k)}^2 \quad (3.11)$$

Im Falle von  $d$  Dimensionen ergeben sich die Koeffizienten  $u_{\mathbf{s},\mathbf{t}}$  zu der  $d$ -fachen Doppelsumme

$$u_{\mathbf{s},\mathbf{t}} = \sum_{l_1=1}^{s_1} \cdots \sum_{l_d=1}^{s_d} \sum_{k_1=1}^{s_1} \cdots \sum_{k_d=1}^{s_d} A_{l_1, \mathbf{Q}_1, k_1, \mathbf{Q}_k}(x_{l_1, \mathbf{Q}_1}, x_{k_1, \mathbf{Q}_k}) \cdot \left( \frac{1}{4} \right)^{|\mathbf{s}| - \sum_{i=1}^d \max\{l_i, k_i\}} \cdot u_{l_1, \mathbf{Q}_1}^1 \cdot u_{k_1, \mathbf{Q}_k}^2, \quad (3.12)$$

wobei

$$\mathbf{Q}_1 = \left( \mathbf{Q}_{s_1, t_1}(l_1), \dots, \mathbf{Q}_{s_d, t_d}(l_d) \right) \in \mathbb{R}^d$$

der Vektor ist, der die Indizes der entsprechenden eindimensionalen Basisfunktionen  $\varphi_{s_i, t_i}(x)$  für  $1 \leq i \leq d$  enthält. Die Berechnung der Koeffizienten  $u_{\mathbf{s},\mathbf{t}}$  erfordert somit eine  $2d$ -fache Summe.

Die Summe in (3.10) ist eine unendliche Summe, die die Multiplikation (3.7) exakt darstellt. Folglich darf der Ausdruck (3.10) – will man keinen zusätzlichen Fehler in dem Ergebnis der Multiplikation verursachen – nicht vermöge eines Kriteriums für den hierarchischen Überschuß abgeschnitten werden. Dabei ist zu beachten, daß  $u_{\mathbf{s},\mathbf{t}}$  im Grenzfall  $\mathbf{s} \rightarrow \infty$  beliebig klein wird.

Begnügt man sich für die Darstellung von  $u_3^{\varepsilon_3, \|\cdot\|}(\mathbf{x})$  wie für die Darstellung von  $u_1^{\varepsilon_1, \|\cdot\|}(\mathbf{x})$  und  $u_2^{\varepsilon_2, \|\cdot\|}(\mathbf{x})$  damit, daß man alle Basisfunktionen betrachtet, deren hierarchischer Überschuß größer als eine vorgegebene Schranke  $\varepsilon_3$  ist, so hat man in (3.10) unter Beachtung von (2.32) diejenigen Basisfunktionen zu betrachten, für die  $\|u_{\mathbf{s}, \mathbf{t}} \cdot \varphi_{\mathbf{s}, \mathbf{t}}\| \geq \varepsilon_3$  gilt, wobei sich die Koeffizienten  $u_{\mathbf{s}, \mathbf{t}}$  nach (3.12) berechnen.

Um die Schwierigkeiten in der Berechnung der Multiplikation zweier Dünngitter-Funktionen in der vorangegangenen Diskussion in der Praxis zu umgehen, kann man wie folgt bei der Multiplikation vorgehen: ausgehend von den nodalen Werten an den Punkten des dünnen Gitters multipliziert man die Funktionen punktweise, interpretiert die so erhaltenen Werte als neue nodale Werte auf dem dünnen Gitter und berechnet aus diesen die hierarchischen Koeffizienten der stückweise  $d$ -linearen hierarchischen Basis. Wie in [Dor97] angegeben, gilt

$$\mathbb{I}^{n, S, d} \circ (fg) = \mathbb{I}^{n, S, d} \circ \left( \left( \mathbb{I}^{n, S, d} \circ f \right) \cdot \left( \mathbb{I}^{n, S, d} \circ g \right) \right) .$$

Dies besagt, daß der Interpolant des Produkts  $fg$  (vgl. hierzu (2.27)) auf einem dünnen Gitter vom Level  $n$  in  $d$  Dimensionen auch dadurch gewonnen werden kann, indem man auf den Punkten des dünnen Gitters die Funktionswerte der Funktionen  $f$  und  $g$  miteinander multipliziert und diese Ergebnisse dann zur Berechnung des Interpolanten auf dem dünnen Gitter heranzieht. Bei der Multiplikation von Funktionswerten von  $f$  und  $g$  auf gemeinsamen Punkten  $\mathbf{x}_{\mathbf{1}, \mathbf{i}}$  des dünnen Gitters ist das Ergebnis (bis auf Maschinengenauigkeit, vgl. IEEE-Standard) genau. Werden  $f$  und  $g$  hingegen auf unterschiedlichen Gittern dargestellt, so müssen wie zuvor Werte an Gitterpunkten, die bei der Darstellung von  $f$ , nicht jedoch zur Darstellung von  $g$  benötigt werden (oder umgekehrt), interpoliert werden.

Seien  $G_f$  und  $G_g$  die dünnen Gitter, die zur Darstellung des Interpolanten der Funktionen  $f$  und  $g$  verwendet werden, und sei  $G_{fg} = G_f \circ^S G_g$ . Sei weiterhin angenommen, daß punktweise in jedem Dünngitter-Punkt des Gitters  $G_{fg}$  die Abschätzungen

$$\begin{aligned} |(f - f^{\varepsilon_f, \|\cdot\|})(\mathbf{x}_{\mathbf{1}, \mathbf{i}})| &< \eta_f \\ |(g - g^{\varepsilon_g, \|\cdot\|})(\mathbf{x}_{\mathbf{1}, \mathbf{i}})| &< \eta_g \end{aligned}$$

gelten, wobei  $\mathbf{x}_{\mathbf{1}, \mathbf{i}}$  ein Gitterpunkt von  $G_{fg}$  sei. Dann sind  $\eta_f$  und  $\eta_g$  in der Größenordnung des Interpolationsfehlers auf dem jeweiligen Gitter. Ist  $\mathbf{x}_{\mathbf{1}, \mathbf{i}}$  kein Punkt des Gitters  $G_f$ , so folgt

$$\begin{aligned} |(fg - f^{\varepsilon_f, \|\cdot\|} g^{\varepsilon_g, \|\cdot\|})(\mathbf{x}_{\mathbf{1}, \mathbf{i}})| &< |g(\mathbf{x}_{\mathbf{1}, \mathbf{i}})| \cdot |(f - f^{\varepsilon_f, \|\cdot\|})(\mathbf{x}_{\mathbf{1}, \mathbf{i}})| \\ &< |g(\mathbf{x}_{\mathbf{1}, \mathbf{i}})| \cdot \eta_f . \end{aligned}$$

Analog erhält man

$$\begin{aligned} |(fg - f^{\varepsilon_f, \|\cdot\|} g^{\varepsilon_g, \|\cdot\|})(\mathbf{x}_{1,i})| &< |f(\mathbf{x}_{1,i})| \cdot |(g - g^{\varepsilon_g, \|\cdot\|})(\mathbf{x}_{1,i})| \\ &< |f(\mathbf{x}_{1,i})| \cdot \eta_g \ , \end{aligned}$$

falls der Punkt  $\mathbf{x}_{1,i}$  nicht in  $G_g$  enthalten ist. Somit gilt punktweise in jedem Punkt des dünnen Gitters  $G_{fg}$  die Abschätzung

$$\begin{aligned} |(fg - f^{\varepsilon_f, \|\cdot\|} g^{\varepsilon_g, \|\cdot\|})(\mathbf{x}_{1,i})| &< \max\{|g(\mathbf{x}_{1,i})| \cdot \eta_f, |f(\mathbf{x}_{1,i})| \cdot \eta_g\} \\ &=: \eta_{fg} \ . \end{aligned} \tag{3.13}$$

Da die Funktionen  $f, g \in \mathcal{X}^{\mathbf{r}, d}(\bar{\Omega}^d)$  mit  $\mathbf{r} \geq \mathbf{2}$  beschränkt sind, ist die “Störung”  $\eta_{fg}$ , die an bestimmten Gitterpunkten den Fehler in (3.13) angibt, ebenfalls von der Größenordnung des Interpolationsfehlers. Folglich ist auch der Interpolationsfehler des Produkts  $f^{\varepsilon_f, \|\cdot\|} g^{\varepsilon_g, \|\cdot\|}$  zweier Interpolanten von der Ordnung des Diskretisierungsfehlers, da sich der Interpolationsfehler bei obiger Vorgehensweise durch den Interpolationsfehler, der bei der Interpolation des Produkts der Funktionen  $f$  und  $g$  anfällt, und dem Interpolationsfehler der Störung abschätzen läßt. Demnach kann in einer Implementierung die Multiplikation von Dünngitter-Funktionen punktweise auf den Gitterpunkten des gemeinsamen Gitters  $G_{fg}$  durchgeführt werden, was im Vergleich zur zuerst diskutierten Vorgehensweise in  $\mathcal{O}(N \log(N)^{d-1})$  Operationen in  $d$  Dimensionen gemacht werden kann.

## 4 $d$ -dimensionale Finite Differenzen Operatoren auf dünnen Gittern

### 4.1 Grundlagen

In diesem Abschnitt führen wir die Finiten Differenzen Operatoren auf dünnen Gittern ein, die insbesondere für elliptische Differentialoperatoren zweiter Ordnung

$$L = \sum_{i=1}^d \sum_{j=1}^d a_{ij}(\mathbf{x}) \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} + \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial}{\partial x_i} + c(\mathbf{x}) \quad (4.1)$$

in  $d$  Dimensionen benötigt werden. Dies sind hier Finite Differenzen Operatoren für erste und zweite Ableitungen sowie gemischte Ableitungen. Im vorigen Abschnitt 3.2 wurden die Addition und Subtraktion sowie die Multiplikation und Division von zwei Dünngitter-Funktionen bereits eingeführt, die auch im diskreten Differentialoperator angewandt werden müssen. Im Falle konstanter Koeffizientenfunktionen in  $L$  reduziert sich dabei die Multiplikation bzw. Division von Funktionen auf eine skalare Multiplikation mit einer Konstanten. Folglich ist an dieser Stelle nur noch die Diskretisierung von Ableitungen zu untersuchen. Von besonderem Interesse sind hierbei neben der Bauart der Operatoren Fragen aus dem Bereich der Konsistenz, Stabilität und Konvergenz.

Es ist hiermit jedoch klar, daß durch das Bereitstellen dieser Operatoren grundsätzlich der Grundstein für die Diskretisierung anderer Typen von Differentialgleichungen (z.B. parabolischer DGL, Navier Stokes Gleichungen etc.) gelegt ist. Dabei geschieht die Diskretisierung stets durch die Funktionalalgebra und den Finiten Differenzen Operatoren auf dünnen Gittern. Das Kapitel 7 der numerischen Beispiele beschäftigt sich neben elliptischen Differentialgleichung unter anderem auch mit einigen der soeben genannten Typen von Gleichungen.

Die in einer Raumdimension bekannten Diskretisierungssterne für Finite Differenzen Operatoren auf Gittern mit konstanter Maschenweite pro Raumdimension (vgl. u.a. [Hac86, For88, Str89, GrRo92, Tho95, For97]) sind die *zentrale Differenz*

$$\begin{aligned} \delta_{l,i}^0 \circ u(x) &= \frac{1}{2 \cdot 2^{-l}} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}_{l,i} \circ u(x) \\ &= \frac{u(x_i + 2^{-l}) - u(x_i - 2^{-l})}{2 \cdot 2^{-l}} \end{aligned} \quad (4.2)$$

deren Konsistenzordnung zweite Ordnung  $\mathcal{O}((2^{-l})^2)$  ist, sowie die *Rück-*

wärts-Differenz

$$\begin{aligned}\delta_{l,i}^- \circ u(x) &= \frac{1}{2^{-l}} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix}_{l,i} \circ u(x) \\ &= \frac{u(x_i) - u(x_i - 2^{-l})}{2^{-l}}\end{aligned}\quad (4.3)$$

und die Vorwärts-Differenz

$$\begin{aligned}\delta_{l,i}^+ \circ u(x) &= \frac{1}{2^{-l}} \begin{bmatrix} 0 & -1 & 1 \end{bmatrix}_{l,i} \circ u(x) \\ &= \frac{u(x_i + 2^{-l}) - u(x_i)}{2^{-l}},\end{aligned}\quad (4.4)$$

deren Konsistenzordnung jeweils erste Ordnung  $\mathcal{O}(2^{-l})$  ist. Für die Approximation zweiter Ableitungen nehmen wir den bekannten Stern

$$\begin{aligned}\delta_{l,i} \circ u(x) &= \frac{1}{2^{-2l}} \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}_{l,i} \circ u(x) \\ &= \frac{u(x_i + 2^{-l}) - 2 \cdot u(x_i) + u(x_i - 2^{-l})}{(2^{-l})^2},\end{aligned}\quad (4.5)$$

der wie die zentrale Differenz für erste Ableitungen von zweiter Diskretisierungsordnung ist. Dabei setzen wir für alle der eben genannten Sterne ein äquidistantes Gitter der Maschenweite  $2^{-l}$  in einer Raumdimension voraus. In mehreren Dimensionen werden die Sterne lokal in der gewählten Raumdimension betrachtet, wobei lokal ein volles Gitter zugrundegelegt ist<sup>6</sup>. Auf vollen Gittern in  $d$  Dimensionen ist dabei klar, daß (lokal) in jede Raumdimension äquidistante Gitter auftreten. Dies trifft auch auf dünne Gitter in  $d$  Raumdimensionen zu, für die lokal ebenfalls für eine beliebige, aber feste Raumdimension äquidistante eindimensionale Gitter (allerdings mit von den Gitterpunkten abhängigen Schrittweiten) vorliegen. Für eine einfache Notation der noch zu definierenden Finiten Differenzen Operatoren auf dünnen Gittern führen wir die Bezeichnungen

$$\mathbf{D}_i^{\pm,0} \quad \text{und} \quad \mathbf{D}_{ii} \quad (4.6)$$

ein, die die Anwendung einer ersten bzw. zweiten Ableitung an jedem Gitterpunkt in Koordinatenrichtung  $i$  bedeuten. Dabei ist stets als (lokale) Maschenweite der Abstand zum nächsten Nachbarknoten in (positive oder negative) Koordinatenrichtung  $i$  zu wählen. Die hochgestellten Bezeichner  $\pm$  und  $0$  bei  $\mathbf{D}_i^{\pm,0}$  deuten dabei die verschiedenen Wahlmöglichkeiten der Finiten Differenzensterne  $\delta_{l,i}^{\pm}$  und  $\delta_{l,i}^0$  (vgl. Sterne (4.2) bis (4.4)) an.

Die Erweiterung der eindimensionalen Finiten Differenzen Sterne auf Sterne in mehreren Dimensionen ist dabei klar (vgl. [Hac86, For88, Str89, GrRo92,

---

<sup>6</sup>Dies bedeutet, daß die beiden Schrittweiten 'links' und 'rechts' bzgl. der betrachteten Raumdimension gleich groß sind.

Tho95, For97]). So erhält man beispielsweise aus (4.5) den bekannten 5-Punkte-Stern

$$\begin{aligned} \Delta_{\mathbf{l},\mathbf{i}} \circ u(\mathbf{x}) &= \frac{1}{2^{-2l}} \begin{bmatrix} & & 1 & & \\ & 1 & -4 & 1 & \\ & & & & \\ & & & & 1 \\ & & & & & \end{bmatrix}_{\mathbf{x}_{\mathbf{l},\mathbf{i}}, \mathbf{h}_{\mathbf{l}}} \circ u(x) \\ &= \frac{1}{2^{-2l}} \left( u(x_{l_1, i_1} - 2^{-l}, x_{l_2, i_2}) + u(x_{l_1, i_1} + 2^{-l}, x_{l_2, i_2}) + \right. \\ &\quad \left. + u(x_{l_1, i_1}, x_{l_2, i_2} - 2^{-l}) + u(x_{l_1, i_1}, x_{l_2, i_2} + 2^{-l}) - \right. \\ &\quad \left. - 4 \cdot u(x_{l_1, i_1}, x_{l_2, i_2}) \right) , \end{aligned} \quad (4.7)$$

für die Diskretisierung des Laplace-Operators auf einem äquidistanten Gitter der Maschenweite  $2^{-l} \equiv 2^{-l_1} \equiv 2^{-l_2}$  in jede Raumdimension, der ebenfalls wie (4.5) konsistent von 2. Ordnung ist. Der Beweis dieser Aussage benötigt die Operatoren

$$\mathbf{R}_{\mathbf{l},\mathbf{i}}^d : f \longrightarrow f_{\mathbf{l}} \quad (4.8)$$

und

$$\tilde{\mathbf{R}}_{\mathbf{l},\mathbf{i}}^d : f \longrightarrow \tilde{f}_{\mathbf{l}} , \quad (4.9)$$

die die Restriktion von kontinuierlichen Funktionen  $f$  auf die Gitterpunkte  $\Omega_{\mathbf{l}}^d = \bar{\Omega}_{\mathbf{l}}^d \setminus \Gamma_{\mathbf{l}}^d$  bzw. auf die Gitterpunkte  $\bar{\Omega}_{\mathbf{l}}^d$  bedeuten. Ein Nachweis der Konsistenzordnung 2 für den Stern (4.7) unter Verwendung der Restriktionsoperatoren findet sich beispielsweise in [Hac86]. Es ist jedoch nicht möglich, den Stern (4.7) als Diskretisierung des Laplace-Operators auf dünnen Gittern anzuwenden, wie uns das noch folgende (negative) Lemma 4.2 zeigt.

Bevor wir dieses Lemma formulieren, wollen wir nachweisen, daß sich obige Finiten Differenzen Operatoren wie schon die Hierarchisierung (also die Basis transformationen von stückweiser linearer, hierarchischer Basis in stückweise lineare, nodale Basis) auch als Integraloperator formulieren lassen.

**Lemma 4.1** *Sei  $u \in C^2(\mathbb{R})$  und die Hutfunktion  $\varphi_{l,i}(x)$  wie in (2.2) mittels*

$$\begin{aligned} \varphi_{l,i}(x) &: \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto \varphi_{l,i}(x) &= \begin{cases} \varphi_{l,i}^1(x) := \frac{x - (x_i - 2^{-l})}{2^{-l}} , & x_i - 2^{-l} \leq x \leq x_i \\ \varphi_{l,i}^2(x) := \frac{(x_i + 2^{-l}) - x}{2^{-l}} , & x_i < x \leq x_i + 2^{-l} \\ 0 , & x \notin [x_i - 2^{-l}, x_i + 2^{-l}] \end{cases} \end{aligned}$$

definiert. Dann gewinnt man die diskreten Ableitungen  $\delta_{l,i}^{\pm}$  und  $\delta_{l,i}^0$  über fol-

gende Integral-Identitäten:

$$\begin{aligned}\delta_{l,i}^- &: 0 = \int_{-\infty}^{\infty} \varphi_{l,i}^1(x) u''(x) dx \\ \delta_{l,i}^+ &: 0 = \int_{-\infty}^{\infty} \varphi_{l,i}^2(x) u''(x) dx \\ \delta_{l,i}^0 &: 0 = \int_{-\infty}^{x_i} \varphi_{l,i}^1(x) u''(x) dx - \int_{x_i}^{\infty} \varphi_{l,i}^2(x) u''(x) dx\end{aligned}$$

Da die eindimensionale Hierarchisierung (2.13) einer gewichteten, diskreten zweiten Ableitung entspricht, kann ebenso  $\delta_{l,i}$  als Integraloperator formuliert werden (vgl. Satz 2.1).

**Beweis für  $\delta_{l,i}^-$ :**

$$\begin{aligned}0 &= \int_{-\infty}^{\infty} \varphi_{l,i}^1(x) u''(x) dx \\ &= \int_{x_i - 2^{-l}}^{x_i} \varphi_{l,i}^1(x) u''(x) dx \\ &= \left[ \varphi_{l,i}^1(x) u'(x) \right]_{x_i - 2^{-l}}^{x_i} - \int_{x_i - 2^{-l}}^{x_i} (\varphi_{l,i}^1)'(x) u'(x) dx \\ &= u'(x_i) - \frac{u(x_i) - u(x_i - 2^{-l})}{2^{-l}} \\ \Rightarrow u'(x_i) &= \frac{u(x_i) - u(x_i - 2^{-l})}{2^{-l}}\end{aligned}$$

**Beweis für  $\delta_{l,i}^+$ :**

$$\begin{aligned}0 &= \int_{-\infty}^{\infty} \varphi_{l,i}^2(x) u''(x) dx \\ &= \int_{x_i}^{x_i + 2^{-l}} \varphi_{l,i}^2(x) u''(x) dx \\ &= \left[ \varphi_{l,i}^2(x) u'(x) \right]_{x_i}^{x_i + 2^{-l}} - \int_{x_i}^{x_i + 2^{-l}} (\varphi_{l,i}^2)'(x) u'(x) dx\end{aligned}$$



$$\begin{aligned}
&= -u'(x_i) - \frac{u(x_i + 2^{-l}) - u(x_i)}{-2^{-l}} \\
\Rightarrow u'(x_i) &= \frac{u(x_i + 2^{-l}) - u(x_i)}{2^{-l}}
\end{aligned}$$

**Beweis für  $\delta_{l,i}^0$ :**

$$\begin{aligned}
0 &= \int_{-\infty}^{x_i} \varphi_{l,i}^1(x) u''(x) dx - \int_{x_i}^{\infty} \varphi_{l,i}^2(x) u''(x) dx \\
&= \left[ u'(x_i) - \frac{u(x_i) - u(x_i - 2^{-l})}{2^{-l}} \right] + \left[ u'(x_i) - \frac{u(x_i + 2^{-l}) - u(x_i)}{2^{-l}} \right] \\
&= 2u'(x_i) - \frac{u(x_i + 2) - u(x_i - 2^{-l})}{2^{-l}} \\
\Rightarrow u'(x_i) &= \frac{u(x_i + 2) - u(x_i - 2^{-l})}{2 \cdot 2^{-l}}
\end{aligned}$$

**Beweis für  $\delta_{l,i}$ :** Ein Vergleich der Operatoren zeigt, daß lokal die Identität

$$\delta_{l_j, i_j} = \frac{2}{h_j^2} \cdot H_{x_{l_j, i_j}, h_j}$$

mit  $h_j = 2^{-l_j}$  gilt. Damit erhält man die Integraldarstellung von  $\delta_{l_j, i_j}$  bzw.  $\delta_{l,i}$  aus Satz 2.1.  $\square$

## 4.2 Konstruktion

Bevor wir die Finiten Differenzen Operatoren auf dünnen Gittern einführen, formulieren wir nun ein Lemma, das – wie schon vor Lemma 4.1 angemerkt – besagt, daß sich die Diskretisierung eines Gitters mit äquidistanter Maschenweite in jeder Koordinatenrichtung nicht auf ein dünnes Gitter übertragen läßt.

**Lemma 4.2** *Die Diskretisierung eines elliptischen partiellen Differentialoperators  $L$ , definiert wie in (4.1), mittels der Sterne (4.2), (4.3), (4.4) und (4.5) – jeweils angewandt mit der lokalen Maschenweite in jede Raumdimension – führt auf dünnen Gittern zu einer nicht-konsistenten Diskretisierung.*

**Beweis:** Siehe Abschnitt A.1, Seite 186  $\square$

Die negative Aussage dieses Lemma ist die Motivation, spezielle Finite Differenzen Operatoren auf dünnen Gittern zu konstruieren. Dabei erweisen sich die Operatoren im Fall der dünnen Gitter als nicht wesentlich komplizierter als im Fall äquidistanter Gitter. Mit den Transformationen “Hierarchisierung” und “Enthierarchisierung” haben wir bereits – neben den schon

eingeführten Finite Differenzen Sternen (4.2) bis (4.5) – alle Zutaten zur Definition der Finiten Differenzen Operatoren für erste und zweite Ableitungen auf dünnen Gittern.

**Definition 4.1** Um die Schreibweise zu vereinfachen, führen wir die Notation

$$u^{k,x_j}(\cdot) \equiv \frac{\partial^k u}{\partial x_j^k}(\cdot) \quad (4.10)$$

ein. Dabei bedeutet der erste hochgestellte Index  $k$  einen Ableitungsgrad, die zweite angegebene, hochgestellte Größe  $x_j$  gibt die Raumdimension an, in der die entsprechende Ableitung zu betrachten ist. Zudem stellt

$$u^{k_1,x_{j_1},k_2,x_{j_2},\dots,k_m,x_{j_m}}(\cdot) \equiv \frac{\partial^{k_1+k_2+\dots+k_m} u}{\partial x_{j_1}^{k_1} \partial x_{j_2}^{k_2} \dots \partial x_{j_m}^{k_m}}(\cdot)$$

gemischte Ableitungen dar, wobei  $0 \leq m \leq d$  zu beachten ist. Die Größen  $(k_s, x_{j_s})$  sind dabei stets als Tupel analog zu (4.10) zu betrachten.

Zur Angabe der benötigten Glattheitsbedingungen für die nachfolgenden Finiten Differenzen Operatoren auf dünnen Gittern definieren wir die Räume  $C_{mix}^{t,s}(\bar{\Omega}) = C_{mix}^{t,1+s\mathbf{e}_1}(\bar{\Omega}) \cap \dots \cap C_{mix}^{t,1+s\mathbf{e}_d}(\bar{\Omega})$  (vgl. [GrKn98] bzgl. der Schreibweise) mit  $C_{mix}^{\mathbf{r}}(\bar{\Omega}) = C^{r_1}(\bar{\Omega}) \otimes \dots \otimes C^{r_d}(\bar{\Omega})$ . Dabei seien  $t, s \geq 0$ ,  $\mathbf{1} = (1, \dots, 1)$  und  $\mathbf{e}_i$  mit  $1 \leq i \leq d$  der  $i$ -te Einheitsvektor. Die Räume  $C_{mix,L}^{t,s}(\bar{\Omega}) \subset C_{mix}^{t,s}(\bar{\Omega})$  und  $C_{mix,L}^{\mathbf{r}}(\bar{\Omega}) \subset C_{mix}^{\mathbf{r}}(\bar{\Omega})$  bezeichnen dabei diejenigen Funktionen, die bezüglich allen Raumdimensionen Lipschitz-stetige Funktionen darstellen. Die Normen  $\|\cdot\|_{C_{mix}^{t,s}(\bar{\Omega})}$  und  $\|\cdot\|_{C_{mix}^{\mathbf{r}}(\bar{\Omega})}$  sowie  $\|\cdot\|_{C_{mix,L}^{t,s}(\bar{\Omega})}$  und  $\|\cdot\|_{C_{mix,L}^{\mathbf{r}}(\bar{\Omega})}$  sind dabei analog zu [GrKn98] bzw. [Hac86] definiert.

Um die Notation in den Beweisen für die Konsistenzordnungen der einzelnen Finiten Differenzen Operatoren auf dünnen Gittern einfach zu halten, führen wir zudem Operatoren ein, die für einen gegebenen Punkt eines dünnen Gitters den nächsten Nachbarknoten bzw. den hierarchischen Vater in eine beliebige (positive/negative) Raumdimension angeben.

**Definition 4.2** Für jeden Gitterpunkt  $\mathbf{x}_{1,i}$  eines dünnen (oder besser: dyadischen) Gitters sei  $\mathcal{H}_j^+(\mathbf{x}_{1,i})$  der hierarchische Vater in positiver Raumdimension  $j$  sowie  $\mathcal{H}_j^-(\mathbf{x}_{1,i})$  der hierarchische Vater in negativer Raumdimension  $j$ . Die Knoten  $\mathcal{N}_j^+(\mathbf{x}_{1,i}) = \mathbf{x}_{1,i} + h_j \mathbf{e}_j$  und  $\mathcal{N}_j^-(\mathbf{x}_{1,i}) = \mathbf{x}_{1,i} - h_j \mathbf{e}_j$  bezeichnen den nächsten Nachbarknoten in positive und negative Raumdimension  $j$ , wobei  $h_j$  die lokale Maschenweite in Raumdimension  $j$  und  $\mathbf{e}_j$  der  $j$ -te Einheitsvektor ist.

**Bemerkung 4.1**

Die Basistransformation von stückweise  $d$ -linearer, nodaler Basis in stückweise  $d$ -lineare, hierarchische Basis kann unter Benutzung von  $\mathcal{H}_j^\pm(\cdot)$  angegeben werden. Finite Differenzen können mit Hilfe von  $\mathcal{N}_j^\pm(\cdot)$  formuliert werden. Zudem sind die Definitionen untereinander kommutativ:

$$\mathcal{H}_{j_1}^+(\mathcal{H}_{j_2}^-(\mathbf{x}_{\mathbf{l},\mathbf{i}})) = \mathcal{H}_{j_2}^-(\mathcal{H}_{j_1}^+(\mathbf{x}_{\mathbf{l},\mathbf{i}}))$$

und

$$\mathcal{N}_{j_1}^+(\mathcal{N}_{j_2}^-(\mathbf{x}_{\mathbf{l},\mathbf{i}})) = \mathcal{N}_{j_2}^-(\mathcal{N}_{j_1}^+(\mathbf{x}_{\mathbf{l},\mathbf{i}})) \ .$$

Im allgemeinen kann man aber die Reihenfolge der Bestimmung von hierarchischen Vaterknoten und nächsten Nachbarknoten nicht vertauschen: im allgemeinen gilt

$$\mathcal{H}_{j_1}^+(\mathcal{N}_{j_2}^-(\mathbf{x}_{\mathbf{l},\mathbf{i}})) \neq \mathcal{N}_{j_2}^-(\mathcal{H}_{j_1}^+(\mathbf{x}_{\mathbf{l},\mathbf{i}})) \ .$$

Wir haben nun alle notwendigen Vorarbeiten durchgeführt, um nachfolgend die Finiten Differenzen Operatoren auf dünnen Gittern einzuführen. Die anschließenden Sätze 4.1 bis 4.10 sowie Satz 4.13 und deren Beweise beschäftigen sich mit dem Aufbau der Operatoren sowie deren Konsistenzordnungen. Dabei zeigt sich, daß Finite Differenzen Operatoren auf dünnen Gittern stets nach einem festen Schema aufgebaut sind, unabhängig davon, welche Ableitung betrachtet werden soll. Zudem kann jeder Operator in einer nodalen sowie einer hierarchischen Schreibweise angegeben werden. Nachfolgend werden beide Möglichkeiten vorgestellt.

**Nodale Schreibweise:** Ausgehend von den bekannten Finiten Differenzen Operatoren, die auf nodalen Werten arbeiten, definieren wir Finite Differenzen Operatoren auf dünnen Gittern, die ebenfalls nodale Werte, also Funktionswerte, benötigen. Diese Operatoren bestehen aus folgenden drei Teiloperatoren:

- Durchführung einer Basistransformation von stückweise linearer, nodaler Basis in stückweise lineare, hierarchische Basis in alle Raumdimensionen außer derjenigen Raumdimension  $j$ , in die die Finite Differenz angewendet werden soll
- Anwendung eines Finite Differenzen Sterns in Raumdimension  $j$ , wobei hier die Sterne (4.2) bis (4.5) bzw. Sterne höherer Ordnung verwendet werden, die bereits bei vollen Gittern Anwendung finden. Als Maschenweite für die Finiten Differenzen Sterne wird dabei jeweils die lokale Schrittweite zum nächsten Nachbarn genommen.
- Durchführung einer Basistransformation von stückweise linearer, hierarchischer Basis in stückweise lineare, nodale Basis, wiederum in alle Raumdimensionen außer derjenigen Raumdimension  $j$ , in die die Finite Differenz in Schritt 2 angewendet worden ist

Ein Beispiel für dieses Schema findet sich in Abbildung 10, wo eine Finite Differenz für eine zweite Ableitung in  $x$ -Richtung in zwei Raumdimensionen dargestellt wird.

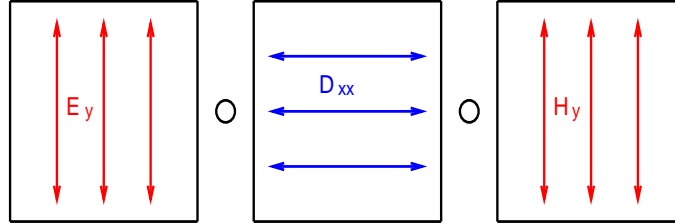


Abbildung 10: Aufbau der Finiten Differenz auf dünnen Gittern für eine zweite Ableitung in  $x$ -Richtung in zwei Dimensionen

Mit diesem Schema lassen sich nun beliebige Finite Differenzen Operatoren auf dünnen Gittern konstruieren:

- Laplace-Operator:

$$\Delta^S = \sum_{i=1}^d \mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i = \sum_{i=1}^d \mathbf{D}_{ii}^S$$

- gemischte zweite Ableitung:

$$\frac{\partial^2}{\partial x_k \partial x_m} = \mathbf{E}_k \circ \mathbf{D}_k \circ \mathbf{H}_k \circ \mathbf{E}_m \circ \mathbf{D}_m \circ \mathbf{H}_m$$

- $\nabla \cdot (A\nabla)$ -Operator:

$$\frac{\partial}{\partial x_k} \left( a_{km} \frac{\partial}{\partial x_m} \right) = \mathbf{E}_k \circ \mathbf{D}_k \circ \mathbf{H}_k \circ \left( \mathbf{a}_{km} \cdot \mathbf{E}_m \circ \mathbf{D}_m \circ \mathbf{H}_m \right)$$

- konvektiver Anteil einer elliptischen DGL:

$$\vec{\beta} \cdot \nabla = \sum_{i=1}^d \beta_i \cdot \mathbf{E}_i \circ \mathbf{D}_i \circ \mathbf{H}_i = \sum_{i=1}^d \beta_i \cdot \mathbf{D}_i^S$$

Neben der nodalen Notation ist es möglich, eine hierarchische Schreibweise der Operatoren anzugeben. Diese Operatoren arbeiten dann auf den hierarchischen Koeffizienten  $u_{l,i}$  aus (2.8) bzw. (2.23).

**hierarchische Schreibweise:** Auch in diesem Fall bestehen die Operatoren aus einem Schema, daß durch die Anwendung von drei Einzeloperatoren aufgebaut ist:

- Durchführung einer eindimensionalen Basistransformation von stückweise linearer, hierarchischer Basis in stückweise lineare, nodale Basis in diejenige Raumdimensionen  $j$ , in die die Finite Differenz angewendet werden soll
- Anwendung eines Finite Differenzen Sterns in Raumdimension  $j$ , wobei hier die Sterne (4.2) bis (4.5) bzw. Sterne höherer Ordnung verwendet werden, die bereits bei vollen Gittern Anwendung finden. Als Maschenweite für die Finiten Differenzen Sterne wird dabei jeweils die lokale Schrittweite zum nächsten Nachbarn genommen.
- Durchführung einer Basistransformation von stückweise linearer, nodaler Basis in stückweise lineare, hierarchische Basis, wiederum in diejenige Raumdimension  $j$ , in die die Finite Differenz in Schritt 2 sowie die erste Transformation in Schritt 1 angewendet worden ist

Das Ergebnis ist wiederum in hierarchischer Darstellung, die Funktionswerte können durch Anwendung einer  $d$ -dimensionalen Enthierarchisierung  $\mathbf{E}$  berechnet werden. Der Laplace-Operator lautet in dieser hierarchischen Notation etwa

$$\Delta^{S,hier} = \sum_{i=1}^d \mathbf{H}^i \circ \mathbf{D}_{ii} \circ \mathbf{E}^i \equiv \mathbf{H} \circ \Delta^S \circ \mathbf{E} .$$

Beide Darstellungsmöglichkeiten, also nodale und hierarchische Darstellung der Finiten Differenzen, sind vermöge

$$\mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i \equiv \mathbf{E} \circ (\mathbf{H}^i \circ \mathbf{D}_{ii} \circ \mathbf{E}^i) \circ \mathbf{H} \quad (4.11)$$

miteinander verbunden, wobei hier als Beispiel eine zweite Ableitung verwendet wird. Analoge Identitäten gelten für beliebige Sterne anstatt  $\mathbf{D}_{ii}^S$ .

### 4.3 Konsistenz

**Satz 4.1 Konsistenz zweite Ableitung,  $d$  Dimensionen,  $u \in C^\infty$**

*Der Operator*

$$\mathbf{D}_{ii}^S = \mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i \quad (4.12)$$

*ist eine Finite Differenzen Diskretisierung für zweite Ableitungen auf dünnen Gittern in Koordinatenrichtung  $i$  mit  $1 \leq i \leq d$ . Unter der Voraussetzung  $u \in C^\infty(\bar{\Omega})$  besitzt der Operator die Konsistenzordnung 2 bezüglich der feinsten Maschenweite  $h$  des Gitters (Randschrittweite):*

$$\|\mathbf{D}_{ii}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^2 u}{\partial x_i^2}\|_\infty \leq C \cdot h^2 \|u\|_{C^\infty(\bar{\Omega})} \quad (4.13)$$

**Beweis:** Siehe Abschnitt A.2, Seite 187 □

**Bemerkung 4.2**

Die Forderung  $u \in C^\infty(\bar{\Omega})$  ist sehr stark und im allgemeinen von der Lösung eines Laplace-Problems, bei der die Operatoren  $\mathbf{D}_{ii}^S$  für  $1 \leq i \leq d$  die entscheidende Rolle spielen, nicht zu erfüllen. Um den Fehler, der durch diese Operatoren verursacht wird, besser zu verstehen, ist es jedoch sinnvoll, zunächst diese starke Restriktion zu fordern. Wir werden nachfolgend sehen, daß man weitaus schwächere Forderungen an die Funktion  $u$  stellen kann, um dennoch die Konsistenzordnung zwei nachzuweisen. Auf der Grundlage des Beweises des Satzes 4.1 kann der Beweis für Satz 4.2 einfach gewonnen werden.

**Satz 4.2 Konsistenz zweite Ableitung, 2 Dimensionen,  $u \in C_{mix}^{2+2 \cdot \mathbf{e}_j}$**

*Unter der Voraussetzung  $u \in C_{mix}^{2+2 \cdot \mathbf{e}_j}(\bar{\Omega})$  ist der Operator*

$$\mathbf{D}_{jj}^S = \mathbf{E}_j \circ \mathbf{D}_{jj} \circ \mathbf{H}_j \quad (4.14)$$

*eine Finite Differenzen Diskretisierung für zweite Ableitungen in zwei Dimensionen auf dünnen Gittern in Koordinatenrichtung  $j$  für  $1 \leq j \leq 2$  mit der Konsistenzordnung 2 bezüglich der feinsten Maschenweite des Gitters (Randschrittweite):*

$$\|\mathbf{D}_{jj}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^2 u}{\partial x_j^2}\|_\infty \leq C \cdot h^2 \|u\|_{C_{mix}^{2+2 \cdot \mathbf{e}_j}(\bar{\Omega})} \quad (4.15)$$

**Beweis:** Siehe Abschnitt A.3, Seite 196 □

**Satz 4.3 Konsistenz Laplace-Operator, 2 Dimensionen,  $u \in C_{mix}^{2,2}$**

*Der Dünngitter-Laplace-Operator in zwei Dimensionen*

$$\Delta^S = \sum_{j=1}^2 \mathbf{D}_{jj}^S \quad (4.16)$$

*auf regulären dünnen Gittern besitzt unter der Voraussetzung  $u \in C_{mix}^{2,2}(\bar{\Omega})$  (vgl. auch Satz 4.2) die Diskretisierungsordnung zwei bezüglich der feinsten Maschenweite des dünnen Gitters (Randmaschenweite):*

$$\|\Delta^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \Delta u\|_\infty \leq C \cdot h^2 \|u\|_{C_{mix}^{2,2}(\bar{\Omega})} \quad (4.17)$$

*Dabei ist wie im Fall der Diskretisierung auf vollen Gittern die Konstante  $C$  in der Abschätzung (4.17) abhängig von der Dimension.*

**Beweis:** Die Idee des Beweises ist dieselbe wie im Vollgitter-Fall: die Addition von Operatoren, deren Diskretisierungsordnung zweite Ordnung ist (vgl. hierzu Satz 4.2), ist wiederum ein Operator der Diskretisierungsordnung zwei. Folglich ist die Konsistenzordnung des Dünngitter-Laplace-Operator ebenfalls zwei, wobei jedoch hier die Glattheitsvoraussetzungen für jede einzelne, am Laplace-Operator beteiligte zweite Ableitung beachtet werden und daher  $u \in C_{mix}^{2,2}(\bar{\Omega})$  gefordert werden muß:

$$\begin{aligned}
\|\Delta^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \Delta u\|_\infty &= \left\| \sum_{j=1}^d \mathbf{D}_{jj}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \Delta u \right\|_\infty \\
&\leq \sum_{j=1}^d \left\| \mathbf{D}_{jj}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^2 u}{\partial x_j^2} \right\|_\infty \\
&\stackrel{(4.15)}{\leq} d \cdot C \cdot h^2 \|u\|_{C_{mix}^{2,2}(\bar{\Omega})} \\
&= C(d) \cdot h^2 \|u\|_{C_{mix}^{2,2}(\bar{\Omega})} \quad \square
\end{aligned}$$

**Satz 4.4 Konsistenz zweite Ableitung, 2 Dimensionen,  $u \in C_{mix,L}^{1+2 \cdot \mathbf{e}_j}$  sowie Konsistenz Laplace, 2 Dimensionen,  $u \in C_{mix,L}^{1,2}$**

Unter den Voraussetzungen  $u \in C_{mix,L}^{1+2 \cdot \mathbf{e}_j}(\bar{\Omega})$  ist der Operator

$$\mathbf{D}_{jj}^S = \mathbf{E}_j \circ \mathbf{D}_{jj} \circ \mathbf{H}_j \quad (4.18)$$

eine Finite Differenzen Diskretisierung für zweite Ableitungen auf dünnen Gittern in Koordinatenrichtung  $j$  für  $1 \leq j \leq 2$  in zwei Dimensionen und besitzt die Konsistenzordnung  $\mathcal{O}(h^{1+\sigma})$  mit  $\sigma \geq \frac{1}{2}$  bezüglich der feinsten Maschenweite des Gitters (Randschrittweite):

$$\left\| \mathbf{D}_{jj}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^2 u}{\partial x_j^2} \right\|_\infty \leq C \cdot h^{1+\sigma} \|u\|_{C_{mix,L}^{1+2 \cdot \mathbf{e}_j}(\bar{\Omega})} \quad (4.19)$$

Unter den Voraussetzungen  $u \in C_{mix,L}^{1,2}(\bar{\Omega})$  besitzt der Laplace-Operator ebenfalls die Konsistenz-Ordnung  $\mathcal{O}(h^{1+\sigma})$  mit obigen  $\sigma$ :

$$\|\Delta^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \Delta u\|_\infty \leq C \cdot h^{1+\sigma} \|u\|_{C_{mix,L}^{1,2}(\bar{\Omega})} \quad (4.20)$$

**Beweis:** Der Beweis für die erste Aussage von Satz 4.4 befindet sich im Abschnitt A.4, Seite 199. Der Beweis zur Aussage über den Laplace-Operator läuft analog zum Beweis von Satz 4.3.  $\square$

Die Sätze 4.2, 4.3 und 4.4 sind jeweils nur in zwei Raumdimensionen gültig. Eine Verallgemeinerung auf beliebige Dimensionen  $d$  liefern nun nachfolgend die Sätze 4.5 bis 4.7.

**Satz 4.5 Konsistenz zweite Ableitung,  $d$  Dimensionen,  $u \in C_{mix}^{2+2 \cdot \mathbf{e}_j}$**

Unter der Voraussetzung  $u \in C_{mix}^{2+2 \cdot \mathbf{e}_j}(\bar{\Omega})$  ist der Operator

$$\mathbf{D}_{jj}^S = \mathbf{E}_j \circ \mathbf{D}_{jj} \circ \mathbf{H}_j \quad (4.21)$$

eine Finite Differenzen Diskretisierung für zweite Ableitungen in  $d$  Dimensionen auf dünnen Gittern in Koordinatenrichtung  $j$  für  $1 \leq j \leq d$  mit der Konsistenzordnung 2 bezüglich der feinsten Maschenweite des Gitters (Randschrittweite):

$$\|\mathbf{D}_{jj}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^2 u}{\partial x_j^2}\|_\infty \leq C \cdot h^2 \|u\|_{C_{mix}^{2+2 \cdot \mathbf{e}_j}(\bar{\Omega})} \quad (4.22)$$

**Beweis:** Siehe Abschnitt A.5, Seite 200 □

**Satz 4.6 Konsistenz Laplace-Operator,  $d$  Dimensionen,  $u \in C_{mix}^{2,2}$**

Unter der Voraussetzung  $u \in C_{mix}^{2,2}(\bar{\Omega})$  besitzt der  $d$ -dimensionale Laplace-Operator

$$\Delta^S = \sum_{j=1}^d \mathbf{D}_{jj}^S \quad (4.23)$$

auf dünnen Gittern die Konsistenzordnung 2 bezüglich der feinsten Maschenweite des Gitters (Randschrittweite):

$$\|\Delta^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \Delta u\|_\infty \leq C \cdot h^2 \|u\|_{C_{mix}^{2,2}(\bar{\Omega})} \quad (4.24)$$

Dabei hängt die Konstante in (4.24) von der Dimension  $d$  ab.

**Beweis:** Der Beweis geht analog zum Beweis von Satz 4.3, hier unter Verwendung des Ergebnisses aus Satz 4.5. □

**Satz 4.7 Konsistenz zweite Ableitung,  $d$  Dimensionen,  $u \in C_{mix,L}^{1+2 \cdot \mathbf{e}_j}$  und Konsistenz Laplace,  $d$  Dimensionen,  $u \in C_{mix,L}^{1,2}$**

Unter den Voraussetzungen  $u \in C_{mix,L}^{1+2 \cdot \mathbf{e}_j}(\bar{\Omega})$  ist der Operator

$$\mathbf{D}_{jj}^S = \mathbf{E}_j \circ \mathbf{D}_{jj} \circ \mathbf{H}_j \quad (4.25)$$

eine Finite Differenzen Diskretisierung für zweite Ableitungen auf dünnen Gittern in Koordinatenrichtung  $j$  für  $1 \leq j \leq d$  in  $d$  Dimensionen und



besitzt die Konsistenzordnung  $\mathcal{O}(h^{1+\sigma})$  mit  $\sigma \geq \frac{1}{d}$  bezüglich der feinsten Maschenweite des Gitters (Randschrittweite):

$$\|\mathbf{D}_{jj}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^2 u}{\partial x_j^2}\|_\infty \leq C \cdot h^{1+\sigma} \|u\|_{C^{1+2 \cdot e_j}(\bar{\Omega})} \quad (4.26)$$

Unter den Voraussetzungen  $u \in C_{mix,L}^{1,2}(\bar{\Omega})$  besitzt der Laplace-Operator ebenfalls die Konsistenz-Ordnung  $\mathcal{O}(h^{1+\sigma})$  mit obigen  $\sigma$ :

$$\|\Delta^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \Delta u\|_\infty \leq C \cdot h^{1+\sigma} \|u\|_{C_{mix,L}^{1,2}(\bar{\Omega})} \quad (4.27)$$

**Beweis:** Siehe Abschnitt A.6, Seite 201 □

Nach der Betrachtung von Finiten Differenzen für zweite Ableitungen kommen wir nun zur Behandlung der Konsistenz von ersten Ableitungen.

**Satz 4.8 Konsistenz erste Ableitungen,  $d$  Dimensionen,  $u \in C^\infty$**

*Der Operator*

$$\mathbf{D}_j^S = \mathbf{E}_j \circ \mathbf{D}_j \circ \mathbf{H}_j \quad (4.28)$$

ist eine Finite Differenzen Diskretisierung für erste Ableitungen auf dünnen Gittern in Koordinatenrichtung  $j$  mit  $1 \leq j \leq d$ , wobei der Operator  $\mathbf{D}_j$  einen der in Formel (4.6) definierten Finiten Differenzen Sterne repräsentiert. Unter den Voraussetzungen wie in Satz 4.1 besitzen diese Operatoren die Konsistenzordnung zwei bzw. eins:

$$\begin{aligned} \mathbf{D}_j^{S,0} &= \mathbf{E}_j \circ \mathbf{D}_j^0 \circ \mathbf{H}_j & : & \text{Konsistenzordnung zwei} \\ \mathbf{D}_j^{S,+} &= \mathbf{E}_j \circ \mathbf{D}_j^+ \circ \mathbf{H}_j & : & \text{Konsistenzordnung eins} \\ \mathbf{D}_j^{S,-} &= \mathbf{E}_j \circ \mathbf{D}_j^- \circ \mathbf{H}_j & : & \text{Konsistenzordnung eins} \end{aligned}$$

Es gelten somit die Abschätzungen

$$\begin{aligned} \|\mathbf{D}_j^{S,0} \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial u}{\partial x_j}\|_\infty &\leq C \cdot h^2 \|u\|_{C^\infty(\bar{\Omega})} \\ &\text{und} \\ \|\mathbf{D}_j^{S,\pm} \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial u}{\partial x_j}\|_\infty &\leq C \cdot h^1 \|u\|_{C^\infty(\bar{\Omega})} \end{aligned} \quad (4.29)$$

Dabei bezieht sich die Angabe der Konsistenzordnung wiederum auf die feinste Gitterweite (Randgitterweite). Zudem ergeben sich dieselben Identitäten wie auf vollen Gittern:

$$\mathbf{D}_{jj}^S = \mathbf{D}_j^{S,-} \circ \mathbf{D}_j^{S,+} \quad (4.30)$$

$$= \mathbf{D}_j^{S,+} \circ \mathbf{D}_j^{S,-} \quad (4.31)$$

Der Operator  $\mathbf{D}_i^{S,0}$  kann zudem mittels

$$\mathbf{D}_j^{S,0} = \frac{1}{2} \left( \mathbf{D}_j^{S,+} + \mathbf{D}_j^{S,-} \right) \quad (4.32)$$

formuliert werden.

**Beweis:** Siehe Abschnitt A.7, Seite 204 □

#### Satz 4.9 Konsistenz erste Ableitungen, $d$ Dimensionen

Der Operator  $\mathbf{D}_j^{S,0}$  ist unter der Voraussetzung  $u \in C_{mix}^{2+\mathbf{e}_j}(\bar{\Omega})$  eine Finite Differenzen Diskretisierung für erste Ableitungen in Koordinatenrichtung  $j$  auf dünnen Gittern mit der Konsistenzordnung 2. Unter der Voraussetzung  $u \in C_{mix,L}^{1+\mathbf{e}_j}(\bar{\Omega})$  ist der Operator  $\mathbf{D}_j^{S,0}$  konsistent von der Ordnung  $\mathcal{O}(h^{1+1/d})$ . Zudem sind die Operatoren  $\mathbf{D}_j^{S,+}$  und  $\mathbf{D}_j^{S,-}$  (vgl. Satz 4.8) unter der Voraussetzung  $u \in C_{mix}^2(\bar{\Omega})$  konsistent von der Ordnung 1.

**Beweis:** Der Beweis der Aussagen verläuft analog zu den Beweisen der Sätze 4.4 und 4.7. □

#### Bemerkung 4.3

Ein volles Gitter auf dem  $d$ -dimensionalen Einheitsintervall  $[0, 1]^d$ , das in jede Raumdimension  $j$  die Maschenweite  $2^{-n}$  besitzt, kann ebenso wie ein dünnes Gitter als Grundlage für die Darstellung einer Funktion in der stückweise  $d$ -linearen, hierarchischen Basis genommen werden. Auf diesen Gittern können die Finiten Differenzen Operatoren für dünne Gitter ebenfalls angewendet werden. Konsistenz, Stabilität, Konvergenz sowie zahlreiche Eigenschaften der Diskretisierungsmatrizen im Fall Finiten Differenzen auf vollen Gittern sind aus der Literatur (vgl. [Hac86, Str89, GrRo92, Tho95]) hinreichend bekannt.

Ein Vergleich der Operatoren für volle Gitter (vgl. die soeben zitierten Quellen) sowie der Finiten Differenzen für dünne Gitter zeigt dabei, daß im Falle voller Gitter beide Finiten Differenzen Operatoren identisch sind und somit auch das gleiche Fehlerverhalten aufweisen. Der Grund hierfür ist die Tatsache, daß sich die Hierarchisierung und die Enthierarchisierung in Schritt 1 bzw. Schritt 3 der Konstruktion der Operatoren im Endeffekt auf vollen Gittern aufheben<sup>7</sup>, wie man dem Beweis des Satzes 4.10 entnehmen kann. Damit lassen sich für den Fall voller Gitter alle Aussagen, die sich über die bekannten Finiten Differenzen machen lassen, sofort auf die Finiten Differenzen für dünne Gitter übertragen. Im Anhang beweisen wir hierzu noch folgenden Satz:

---

<sup>7</sup>Man beachte, daß dies im Fall von dünnen Gittern gerade nicht der Fall ist und genau durch diesen Umstand eine konsistente Diskretisierung gewährleistet wird.

**Satz 4.10 Gleichheit der Operatoren auf vollen Gittern**

Die Finiten Differenzen Operatoren  $\mathbf{D}_{jj}^S$  aus Satz 4.1 und die gewöhnliche Finite Differenzen Approximation mittels des Sterns  $h_j^{-2} \cdot [1 \ -2 \ 1]$  sind auf vollen Gittern mit Maschenweiten  $h_j = 2^{-l_j}$  in Raumdimension  $j$  identisch.

**Beweis:** Siehe Abschnitt A.8, Seite 213 □

## 4.4 Eigenschaften Finiter Differenzen Operatoren auf dünnen Gitter

### 4.4.1 Matrix-Typen

Im vorigen Abschnitt 4.3 konnten wir zeigen (Satz 4.10), daß die neu eingeführten Finiten Differenzen auf dünnen Gittern und die aus der Literatur (vgl. [Hac86, For88, Str89, GrRo92, Tho95, For97]) bekannten Finiten Differenzen Operatoren auf vollen Gittern mit dyadischer Schrittweite in jede Raumdimension übereinstimmen. Folglich besitzen die Finiten Differenzen auf dünnen Gitter im Fall äquidistanter dyadischer Gitter sämtliche Eigenschaften wie die bekannten Finiten Differenzen auf vollen Gittern. Wir wollen an dieser Stelle eine Reihe dieser Eigenschaften auflisten und untersuchen, welche dieser Eigenschaften sich auch auf den Fall dünner Gitter übertragen lassen. Dabei stützen wir uns in diesem Abschnitt hauptsächlich auf Untersuchungen, die wir anhand konkret berechneter Matrizen durchgeführt haben, jedoch nicht auf theoretische Beweise für die Diskretisierungsmatrizen. Die Aussagen dieses Abschnitts basieren auf den Diskretisierungsmatrizen des Laplace-Operators in nodaler Schreibweise.

Aus der Konstruktion der Finiten Differenzen Operatoren auf dünnen Gittern ist ersichtlich, daß wir selbst für selbstadjungierte Operatoren wie etwa den Laplace-Operator nicht-symmetrische Diskretisierungsmatrizen erhalten. Dies liegt darin begründet, daß in der Konstruktion der Operatoren im dritten Schritt die inverse Operation des ersten Schrittes angewendet wird, nicht jedoch die transponierte Operation. Letztgenannte Vorgehensweise, also die Verwendung des transponierten Operators im letzten Schritt der Konstruktion, die wir hier jedoch nicht anwenden können, würde auf symmetrische Matrizen führen (vgl. hierzu die Methode der Finiten Elemente auf dünnen Gittern [Bun92], die symmetrische Matrizen bei selbstadjungierten Operatoren liefert).

**Definition 4.3 (L-, L<sub>0</sub>- und M-Matrix, vgl. [GrRo92], Definition 1.4)**

Eine Matrix  $A$  heißt *L<sub>0</sub>-Matrix*, wenn  $a_{ij} \leq 0$  gilt für  $i \neq j$ . Eine Matrix heißt *L-Matrix*, wenn  $a_{ii} > 0$  gilt und außerdem  $a_{ij} \leq 0$  für  $i \neq j$ . Eine *L<sub>0</sub>-Matrix*, für die  $A^{-1}$  existiert und  $A^{-1} \geq 0$  gilt, heißt *M-Matrix*.

Kriterien, wann eine gegebene Matrix eine  $M$ -Matrix ist, finden sich etwa in [Hac86, GrRo92].

**Satz 4.11** (vgl. [Hac86], Satz 4.4.1) Sei  $L_h$  die Matrix der Fünfpunkteformel auf einem vollen Gitter bezüglich  $\bar{\Omega} = [0, 1]^2$ . Dann hat  $L_h$  folgende Eigenschaften:

- (a)  $L_h$  ist eine  $M$ -Matrix
- (b)  $L_h$  ist positiv definit
- (c)  $\|L_h\|_\infty \leq 8h^{-2}$ ,  $\|L_h^{-1}\|_\infty \leq \frac{1}{8}$
- (d)  $\|L_h^{-1}\|_2 \leq \frac{1}{16}$

Zudem erfüllt die Diskretisierung des Laplace-Operators mittels der Fünfpunkteformel das diskrete Analogon der Mittelwerteigenschaft und das Maximum-Minimum-Prinzip. Wie zuvor bemerkt, besitzen die Finiten Differenzen für dünne Gitter dieselben Eigenschaften, wenn man sie auf dyadische, volle Gitter anwendet.

#### Lemma 4.3 L- und M-Matrix Eigenschaft

Sei  $L_\Delta^S$  die Diskretisierungsmatrix der Diskretisierung des Laplace-Operators mittels Finiten Differenzen für dünne Gitter auf dünnen Gittern. Dann ist  $L_\Delta^S$  weder eine  $L_0$ -Matrix noch eine  $L$ -Matrix. Folglich ist  $L_\Delta^S$  keine  $M$ -Matrix.

**Beweis:** Explizite Berechnung der Sterne für jeden einzelnen Gitterpunkt eines dünnen Gitters zeigt, daß neben dem positiven Zentraleintrag auch positive Nebendiagonaleinträge auftreten. Somit greift keine der Bezeichnungen in Definition 4.3.  $\square$

#### Definition 4.4 (Vergleichsmatrix, vgl. [HoJo91])

Die Vergleichsmatrix  $M(A) = (m_{i,j})_{1 \leq i,j \leq n}$  einer  $(n \times n)$ -Matrix  $A = (a_{i,j})_{1 \leq i,j \leq n}$  ist gegeben durch

$$m_{i,j} = \begin{cases} |a_{i,j}|, & \text{falls } i = j \\ -|a_{i,j}|, & \text{falls } i \neq j \end{cases}$$

Eine Matrix  $A$  heißt  $H$ -Matrix, falls die Vergleichsmatrix  $M(A)$  eine  $M$ -Matrix ist.

#### Lemma 4.4 H-Matrix Eigenschaft

Sei  $L_\Delta^S$  die Diskretisierungsmatrix der Diskretisierung des Laplace-Operators mittels Finiten Differenzen für dünne Gitter auf dünnen Gittern. Dann ist  $L_\Delta^S$  keine  $H$ -Matrix.

**Numerischer Nachweis bei kleinen Leveln** ( $l \leq 9$ ): Explizite Berechnung der Vergleichsmatrizen der Inversen der Diskretisierungsmatrizen  $L_{\Delta}^S$  mit MATLAB [Mat97] zeigt, daß die Inversen negative Einträge besitzen. Folglich ist die  $M$ -Matrix-Eigenschaft nicht gegeben.  $\square$

Betrachtet man die Matrizen  $L_{\Delta}^S$ , so stellt man zudem fest, daß keine Diagonaldominanz vorliegt. Zwar ist die Zeilensumme der Matrixeinträge stets entweder 1 (für Randknoten) oder 0 (für innere Knoten), da jedoch für innere Knoten positive Nebendiagonalelemente auftreten, kann keine Diagonaldominanz vorliegen.

**Definition 4.5 (P-Matrix, vgl. [HoJo91])**

*Eine reellwertige  $(n \times n)$ -Matrix  $A$  heißt P-Matrix, falls alle  $(k \times k)$ -Hauptminoren von  $A$  positiv sind für  $1 \leq k \leq n$ . Dies ist gleichbedeutend mit der Aussage, daß jeder reellwertige Eigenwert jeder  $(k \times k)$ -Hauptminore positiv ist.*

**Satz 4.12 P-Matrix Eigenschaft**

*Sei  $L_{\Delta}^S$  die Diskretisierungsmatrix der Diskretisierung des Laplace-Operators mittels Finiter Differenzen für dünne Gitter auf dünnen Gittern. Dann ist  $L_{\Delta}^S$  eine P-Matrix.*

**Numerischer Nachweis bei kleinen Leveln** ( $l \leq 9$ ): Explizite Berechnung der Eigenwerte der Hauptminoren mit MATLAB [Mat97] ergibt, daß die Eigenwerte stets positiv sind. Zur Verdeutlichung dient hier die Abbildung 11, bei der die Eigenwerte der  $(k \times k)$ -Hauptminoren von  $L_{\Delta}^S$  auf dem Level 5 angetragen sind (linkes Dreieck der quadratischen Ausgabe). Diese Graphik ergibt sich qualitativ für alle untersuchten Diskretisierungsmatrizen. Zum einfacheren Verständnis wird die Graphik symmetrisiert dargestellt.  $\square$

Als Konsequenz aus Satz 4.12 ergibt sich die Trägheit [HoJo91] der Diskretisierungsmatrizen  $L_{\Delta}^S$  sofort zu  $i(L_{\Delta}^S) = [n, 0, 0]$ , wobei  $n$  hier die Anzahl der Knoten des zugrundeliegenden dünnen Gitters sei. Aus der angegebenen Trägheit folgt nach [HoJo91] die positive Stabilität von  $L_{\Delta}^S$ . Eine Reihe von Aussagen, die sich aus der P-Matrix Eigenschaft, der angegebenen Trägheit sowie der positiven Stabilität (einer beliebigen Matrix, nicht notwendig unsere Diskretisierungsmatrizen) ergeben, finden sich beispielsweise in [HoJo91].

**Definition 4.6 (diskretes Vergleichsprinzip)**

*Eine  $(n \times n)$ -Matrix  $A$  erfüllt das diskrete Vergleichsprinzip, wenn für alle Vektoren  $z \in \mathbf{R}^n$  gilt:*

$$A \cdot z \geq 0 \implies z \geq 0$$

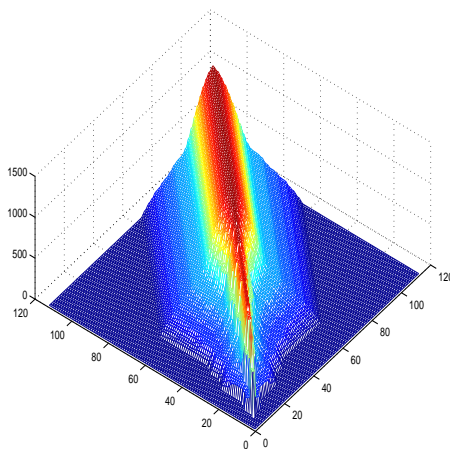


Abbildung 11: Eigenwerte der Hauptminoren der Diskretisierungsmatrix des Laplace–Operators auf dünnen Gittern in zwei Dimensionen auf Level 5, symmetrisierte Ausgabe

*Dabei sind die Vergleiche stets komponentenweise zu verstehen. Das diskrete Vergleichsprinzip ist identisch mit der Eigenschaft der Inversmonotonie (vgl. [GrRo92], Definition 1.6) von Matrizen. Dabei heißt eine  $(n \times n)$ –Matrix  $A$  inversmonoton, falls gilt:*

$$A \cdot v \leq A \cdot w \implies v \leq w$$

Mit Hilfe des Begriffs Inversmonotonie kann man eine  $M$ –Matrix auch als eine inversmonotone  $L_0$ –Matrix auffassen. Zudem ist die Inversmonotonie äquivalent zu der Aussage “ $A^{-1}$  existiert und es gilt  $A^{-1} \geq 0$ ” (vgl. hierzu auch Definition 4.3). Damit ergibt sich automatisch das folgende Lemma:

#### **Lemma 4.5** diskretes Vergleichsprinzip bei $L_{\Delta}^S$

*Die Diskretisierungsmatrizen  $L_{\Delta}^S$  erfüllen nicht das diskrete Vergleichsprinzip und sind somit nicht inversmonoton.*

**Beweis:** Ein Beispiel für einen Vektor  $z$ , für den das Kriterium des diskreten Vergleichsprinzips verletzt wird, kann für Diskretisierungsmatrizen  $L_{\Delta}^S$  auf kleinen Levels ohne großen Aufwand sofort explizit angegeben werden. Damit ist für die in praktischen Anwendungen relevanten Matrizen die Inversmonotonie nicht gegeben.  $\square$

#### **4.4.2** Eigenwerte der Diskretisierungsmatrizen

Wie wir aus Abschnitt 4.4.1 bereits wissen, sind die Diskretisierungsmatrizen nicht symmetrisch, auch wenn wir selbstadjungierte Operatoren wie den

Laplace-Operator betrachten. Trotz dieser Nicht-Symmetrie besitzen die Diskretisierungsmatrizen des Laplace-Operators reelle Eigenwerte. Die Ergebnisse der numerischen Berechnung der Eigenwerte der Diskretisierungsmatrizen  $L_\Delta$  mittels MATLAB [Mat97] für verschiedene Level sind in Tabelle 2 zu finden. Dabei bedeuten  $L$  der Level,  $\lambda$  und  $\Lambda$  den kleinsten und den größten Eigenwert,  $e_\lambda = 2\pi^2 - \lambda_{min}$  die Differenz zwischen kontinuierlichem kleinsten und diskretem kleinsten Eigenwert,  $q_e$  und  $q_\Lambda$  die Quotienten zweier aufeinanderfolgender Werte  $e_\lambda$  und  $\Lambda$ ,  $\kappa$  der Quotient von größtem und kleinstem Eigenwert sowie  $q_\kappa$  der Quotient bzgl. zweier aufeinanderfolgender Werte  $\kappa$ .

Tabelle 2: Eigenwerte von  $\Delta^S$  in zwei Dimensionen,  $\bar{\Omega} = [0, 1]^2$ 

$L$	$\lambda$	$e_\lambda$	$q_e$	$\Lambda$	$q_\Lambda$	$\kappa$	$q_\kappa$
2	18.387503	1.35 <sub>0</sub>	—	69.61	—	3.79	—
3	19.273832	4.65 <sub>-1</sub>	2.90	263.30	3.78	13.66	3.60
4	19.587081	1.52 <sub>-1</sub>	3.06	1031.82	3.92	52.68	3.86
5	19.691734	4.75 <sub>-2</sub>	3.20	4103.95	3.98	208.41	3.96
6	19.724926	1.43 <sub>-2</sub>	3.32	16391.99	3.99	831.03	3.99
7	19.735029	4.18 <sub>-3</sub>	3.42	65544.00	4.00	3321.20	4.00
8	19.738011	1.20 <sub>-3</sub>	3.48	262152.00	4.00	13281.58	4.00
9	19.738868	3.41 <sub>-4</sub>	3.52	1048584.05	4.00	53122.81	4.00

Aus Tabelle 2 kann man ersehen, daß der kleinste Eigenwert mit der Ordnung  $\mathcal{O}(h^2 \log(h^{-1}))$  gegen den kontinuierlichen kleinsten Eigenwert  $2\pi^2 \approx 19.739209$  strebt. Zudem wächst der größte Eigenwert und somit auch die verallgemeinerte Kondition der Diskretisierungsmatrizen  $L_\Delta^S$  mit der Ordnung  $\mathcal{O}(h^{-2})$  analog zur Situation auf vollen Gittern. Damit ist auch im Fall dünner Gitter zu erwarten, daß ein iterativer Löser mit zunehmender Leveltiefe und feiner werdender Maschenweite eine stark wachsende Anzahl von Iterationen benötigen wird, um die numerische Lösung mit einer vorgegebenen Genauigkeit zu berechnen. Aufgrund des Zusammenhangs (4.11) zwischen der nodalen und hierarchischen Schreibweise der Finiten Differenzen Operatoren auf dünnen Gittern und wegen  $\mathbf{H} = \mathbf{E}^{-1}$  ist klar, daß beide Darstellungen dieselben Eigenwerte besitzen, da wegen (4.11) die beiden Darstellungen durch ähnliche Matrizen beschrieben werden. Die explizite Berechnung der Eigenwerte des Laplace-Operators in drei Dimensionen mittels MATLAB [Mat97] zeigt, daß der kleinste Eigenwert hier mit der Ordnung  $\mathcal{O}(h^2 \log(h^{-1})^2)$  gegen  $3\pi^2 \approx 29.608813$  strebt. Damit liegt die Vermutung nahe, daß in  $d$  Dimensionen der kleinste Eigenwert des Laplace-Operators mit der Ordnung  $\mathcal{O}(h^2 \log(h^{-1})^{d-1})$  gegen den kontinuierlichen kleinsten Eigenwert  $d \cdot \pi^2$  konvergiert.

Die Verteilung der Eigenwerte des Laplace-Operators in zwei Dimensionen, hier stellvertretend auf den Leveln 4 und 7, ist in Abbildung 12 zu sehen.

Dabei sind in dieser Abbildung nur die Eigenwerte angegeben, die zu inneren Gitterpunkten gehören. Eigenwerte, die Randpunkten zuzuordnen sind, sind jeweils 1 und somit nicht in der Graphik mit aufgenommen. Man erkennt, daß zwar das Spektrum der Eigenwerte in etwa übereinstimmt, jedoch die Verteilung der Eigenwerte eine andere Struktur besitzt als bei der Finite Differenzen Diskretisierung auf vollen Gittern.

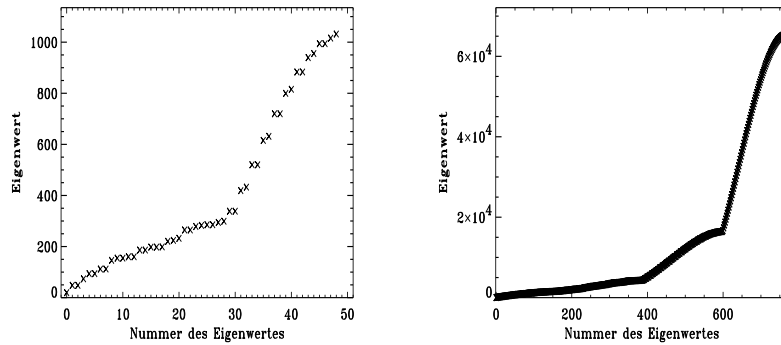


Abbildung 12: Eigenwerte der Diskretisierungsmatrizen auf dünnen Gittern für Level 4 (links) und Level 7 (rechts), Eigenwerte 1 sind weggelassen

#### 4.5 Stabilität und Konvergenz

Um Konvergenz des Verfahrens sicherzustellen, ist neben der bereits nachgewiesenen Konsistenz zudem Stabilität notwendig.

Tabelle 3: Maximumsnorm der Inversen der Diskretisierungsmatrizen

$L$	$P$	$\ (L_{\Delta}^S)^{-1}\ _{\infty}$	$\ (L_{\Delta}^S)^{-1}\ _{\infty}/L$
2	21	0.06875	0.034
3	49	0.08079	0.027
4	113	0.11418	0.029
5	257	0.14982	0.030
6	577	0.18648	0.031
7	1281	0.22337	0.032
8	2817	0.26027	0.033
9	6145	0.30011	0.033

Betrachtet man die Maximumsnorm der Inversen der Diskretisierungsmatrizen in zwei Dimensionen in nodaler Schreibweise, für die auch die Konsistenzaussagen durchgeführt wurden, so erhält man die in Abbildung 13 angegebenen Kurven. Die einzelnen Normen wurden hierbei mittels MATLAB [Mat97] errechnet. In Tabelle 3 sind die Werte sowohl für die Maximumsnorm als auch die mit dem Level skalierte Maximumsnorm enthalten.



Hieraus erkennt man, daß die Maximumsnorm im prä-asymptotischen Bereich ein logarithmisches Wachstumsverhalten aufweist. Unter Verwendung dieses Verhaltens können wir die Abschätzung

$$\|(L_\Delta^S)^{-1}\|_\infty \leq C_1 \cdot L = C_1 \cdot \log(h^{-1}) \quad (4.33)$$

annehmen. Wir erhalten damit in Analogie zur Situation auf vollen Gittern [Hac86] eine Abschätzung für die Konvergenz:

$$\begin{aligned} \|u_h - R_h u\|_\infty &\stackrel{[\text{Hac86}]}{=} \|L_\Delta^S (D_h R_h u - \tilde{R}_h \Delta u)\|_\infty \\ &\leq \|L_\Delta^S\|_\infty \cdot \underbrace{\|D_h R_h u - \tilde{R}_h \Delta u\|_\infty}_{\text{Konsistenz: } \leq C_2 \cdot h^2} \\ &\stackrel{(4.33)}{\leq} C_1 \cdot \log(h^{-1}) \cdot C_2 \cdot h^2 \\ &= C \cdot h^2 \cdot \log(h^{-1}) \end{aligned} \quad (4.34)$$

Wie auf vollen Gittern werden hierbei homogene Dirichlet-Randbedingungen vorausgesetzt. Damit ist die hier angegebene Aussage zur Konvergenz im Gegensatz zur Situation auf vollen Gittern mit einem zusätzlichen logarithmischen Faktor versehen. Betrachtet man sich die mit dem Quadrat des Levels skalierte Maximumsnorm der Inversen der Diskretisierungsmatrizen in drei Dimensionen, so erkennt man, daß sich auch hier ein konstanter Wert ergibt. Dies gibt zusammen mit den bereits gezeigten Aussagen zur Konsistenzordnung Grund zur Vermutung, daß sich eine Abschätzung der Art

$$\|(L_\Delta^S)^{-1}\|_\infty \leq C_1 \cdot L^{d-1} = C_1 \cdot \log(h^{-1})^{d-1}$$

in  $d$  Dimensionen ergibt. Diesen logarithmischen Faktor in der Konvergenz erkennt man auch in den numerischen Beispielen des Kapitels 7.

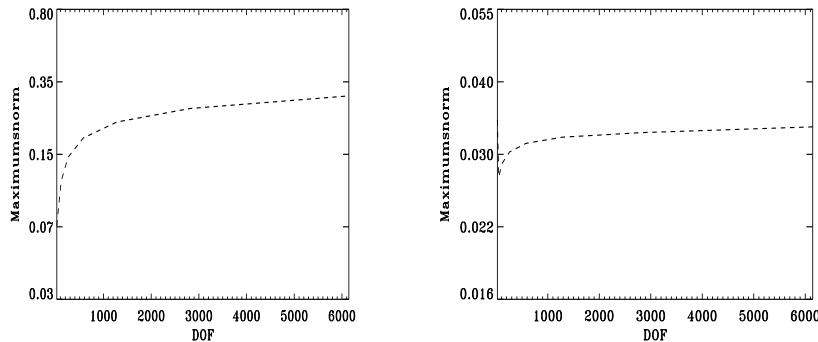


Abbildung 13:  $L_\infty$ - sowie mit dem Level skalierte  $L_\infty$ -Norm der Inversen der Diskretisierungsmatrizen auf sukzessiven Levels

#### 4.6 Finite Differenzen höherer Ordnung auf dünnen Gittern

Finite Differenzen auf dünnen Gittern werden (in nodaler Schreibweise) dadurch realisiert, daß zwischen einer Basistransformation von nodaler zu hierarchischer, stückweiser linearer Basis in  $d-1$  Koordinatenrichtungen und der zugehörigen Rücktransformation ein Finite Differenzen Stern angewendet wird, wobei jeweils die lokale Maschenweite als Schrittweite für die Finite Differenz genommen wird. Durch den Austausch der Sterne mit solchen Sternen, die auf vollen Gittern eine höhere Konsistenzordnung besitzen, lassen sich analog im Fall dünner Gitter Finite Differenzen Operatoren mit höherer Konsistenz erreichen. Beispiele für Sterne höherer Ordnung für die Diskretisierung zweiter Ableitungen sind etwa

$$\frac{1}{12 \cdot h^2} \begin{bmatrix} 1 & -16 & 30 & -16 & 1 \end{bmatrix} \quad (4.35)$$

oder

$$\frac{1}{h^2} \begin{bmatrix} \frac{1}{90} & -\frac{3}{20} & \frac{3}{2} & -\frac{49}{18} & \frac{3}{2} & -\frac{3}{20} & \frac{1}{90} \end{bmatrix} , \quad (4.36)$$

die von 4. bzw. 6. Ordnung konsistent sind. Dabei ist zu beachten, daß diese Sterne lediglich im Inneren des jeweiligen Gebiets anwendbar sind, in Randnähe jedoch andere Sterne (z.B. (4.5)) verwendet werden müssen. Diskretisiert man zweite Ableitungen auf dünnen Gittern unter Verwendung obiger Sterne höherer Ordnung, so übertragen sich die Konsistenzordnungen wie schon im Fall des Sterns (4.5). Dabei sind die entsprechenden Ordnungen wie schon im Fall voller Gitter nur an inneren Gitterpunkten zu erreichen. Neben dem Problem, daß an randnahen Punkten obige Sterne nicht verwendet werden können, hat man zudem das Problem, daß man eine höhere Glattheit einer Lösung voraussetzen muß, wobei die Voraussetzungen bereits im Fall des Sterns (4.5) von der Lösung einer partiellen Differentialgleichung oft nicht erfüllt werden. Nachfolgend bezeichne

$$\mathbf{D}_{ii,4} \quad \text{und} \quad \mathbf{D}_{ii,6}$$

die Anwendung obiger Sterne der Konsistenzordnung 4 bzw. 6 (Konsistenz bezüglich voller Gitter). Eine Aussage über die Konsistenz der entsprechenden Finiten Differenzen Operatoren auf dünnen Gittern gibt Satz 4.13.

#### Satz 4.13 Konsistenz Finiter Differenzen höherer Ordnung

*Der Operator*

$$\mathbf{D}_{jj,4}^S = \mathbf{E}_j \circ \mathbf{D}_{jj,4} \circ \mathbf{H}_j$$

*ist für innere Punkte, an denen der Stern (4.35) bezüglich der Koordinatenrichtung  $j$  angewendet werden kann, eine Finite Differenzen Diskretisierung für zweite Ableitungen auf dünnen Gittern in Koordinatenrichtung  $j$*

mit  $1 \leq j \leq d$ . Unter der Voraussetzung  $u \in C^\infty(\bar{\Omega})$  besitzt der Operator die Konsistenzordnung 4 bezüglich der feinsten Maschenweite  $h$  des Gitters (Randschrittweite):

$$\|\mathbf{D}_{jj,4}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^4 u}{\partial x_j^4}\|_\infty \leq C \cdot h^4 \|u\|_{C^\infty(\bar{\Omega})} \quad (4.37)$$

Die Konsistenzordnung 4 bleibt unter der Voraussetzung  $u \in C_{mix}^{2+4 \cdot \mathbf{e}_j}(\bar{\Omega})$  erhalten. Analog stellt der Operator

$$\mathbf{D}_{jj,6}^S = \mathbf{E}_j \circ \mathbf{D}_{jj,6} \circ \mathbf{H}_j$$

eine Finite Differenzen Diskretisierung für zweite Ableitungen auf dünnen Gittern in Koordinatenrichtung  $j$  mit  $1 \leq j \leq d$  für geeignete innere Gitterpunkte dar, wobei der Stern (4.36) verwendet wird. Unter der Voraussetzung  $u \in C^\infty(\bar{\Omega})$  besitzt der Operator die Konsistenzordnung 6 bezüglich der feinsten Maschenweite  $h$  des Gitters (Randschrittweite):

$$\|\mathbf{D}_{jj,6}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^6 u}{\partial x_j^6}\|_\infty \leq C \cdot h^6 \|u\|_{C^\infty(\bar{\Omega})} \quad (4.38)$$

Dabei bleibt die Konsistenzordnung unter der Voraussetzung  $u \in C_{mix}^{2+6 \cdot \mathbf{e}_j}(\bar{\Omega})$  erhalten.

**Beweis:** Für den Beweis der Aussagen bezüglich  $u \in C^\infty(\bar{\Omega})$  sei auf Abschnitt A.9, Seite 218 verwiesen. Der Beweis für die Aussagen bezüglich  $u \in C_{mix}^{2+4 \cdot \mathbf{e}_j}(\bar{\Omega})$  für den Operator  $\mathbf{D}_{jj,4}^S$  bzw.  $u \in C_{mix}^{2+6 \cdot \mathbf{e}_j}(\bar{\Omega})$  für den Operator  $\mathbf{D}_{jj,6}^S$  verläuft analog zu den Beweisen der Sätze 4.5 und 4.9.  $\square$

## 4.7 Behandlung von Neumann–Randbedingungen

Die Diskretisierung eines Neumann–Randes kann auf vollen Gittern auf grundlegend unterschiedliche Arten erfolgen. Betrachtet man die Gleichung

$$\begin{aligned} Lu &= f \quad \text{in } \Omega^d \\ \frac{\partial u}{\partial \nu} &= g \quad \text{auf } \Gamma^d \end{aligned} \quad (4.39)$$

mit dem Operator  $L$  aus (4.1), so ist der Operator  $L$  nur im Inneren des Gebiets definiert. Folglich besteht eine sinnvolle Diskretisierung eines Neumann–Randes darin, für einen Randknoten eine einseitige Differenz wie etwa die Vorwärts– und Rückwärts–Differenz anzusetzen und die rechte Seite entsprechend zu modifizieren. Einen anderen Zugang zur Berechnung von Neumann–Problemen erhält man, wenn man die Gleichung des Operators  $L$  auf den Rand fortsetzt. Unter der Annahme, daß  $Lu = f$  auf  $\Omega^d$  gilt, kann man am Rand neben der Neumann–Ableitung ebenfalls Gleichung (4.39) an

Neumann-Rändern diskretisieren, wobei hier die rechte Seite  $f$  entsprechend mit auf der rechten Seite des resultierenden linearen Gleichungssystems zu berücksichtigen ist. Auf vollen Gittern hat man die Wahl, welche der beiden Methoden man benutzt. In beiden Fällen erhält man bei einem sinnvoll gestellten Problem vernünftige Resultate.

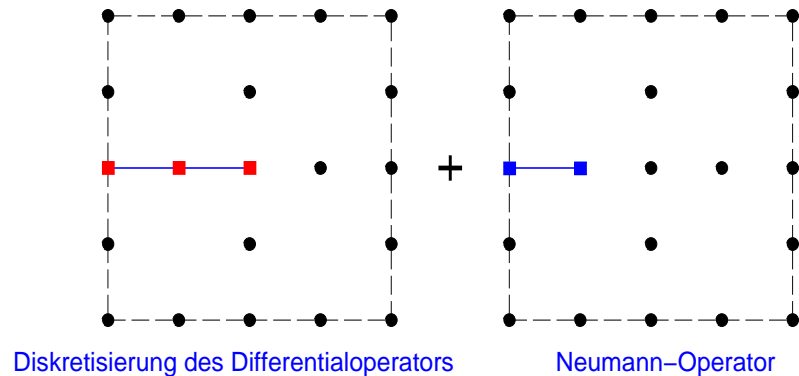


Abbildung 14: Zusammensetzung des Neumann-Operators

Bei einem dünnen Gitter hat man ebenfalls die Wahl, welchen Variante man verwendet. Wir verwenden nachfolgend Variante zwei, diskretisieren also am Rand neben der Neumann-Ableitung auch den Differentialoperator. Dabei verwenden wir einseitige Differenzen sowohl für die benötigte erste Ableitung (Neumann-Ableitung) als auch für den Anteil des Differentialoperators. Die Abbildung 14 zeigt die Vorgehensweise bei der Diskretisierung einer Neumann-Ableitung. Der Differentialoperator wird hierbei durch den (links-seitigen) Stern  $h^{-2} \begin{bmatrix} \underline{1} & -2 & 1 \end{bmatrix}$  diskretisiert, wobei die markierte  $\underline{1}$  den Randpunkt angibt. Diese einseitige Diskretisierung ist lediglich von erster Konsistenzordnung (im Gegensatz hierzu ist (4.5) von zweiter Konsistenzordnung) und stimmt damit in der Ordnung mit der verwendeten Vorwärtsdifferenz für die Diskretisierung der Neumann-Ableitung überein. Somit kann auch im numerischen Experiment maximal die Konsistenzordnung 1 erwartet werden. Wir werden im Kapitel 7 (‘‘Numerische Resultate’’) sehen, daß die numerische Konvergenzrate mit  $\mathcal{O}(h^2 \log(h^{-1}))$  jedoch größer ist als die theoretische Konsistenzordnung.

#### 4.8 Behandlung allgemeiner Gebiete

Lassen sich polygonale Gebiete durch Rechtecke zusammensetzen, so kann auf einfache Weise die Methode der Gebietszerlegung zur Lösung der vorgegebenen Problemstellung angewendet werden. Die Vorgehensweise wird hierbei in Abschnitt 4.8.1 beschrieben.

Daneben können mit dem in [BuDo97, Dor97] beschriebenen Ansatz krummlinig berandete Gebiete vermöge einer Koordinatentransformation auf rechteckige Gebiete transformiert werden. Verschiedene Möglichkeiten, wie ent-

sprechende Transformationen konstruiert werden können, finden sich in der soeben zitierten Quelle. Diesen Ansatz wollen wir in Abschnitt 4.8.2 nutzen, um Problemstellungen auf krummlinig berandeten Gebieten mittels Finiter Differenzen auf dünnen Gittern zu lösen.

Zudem kann man mittels der hierarchischen Basis beliebige Funktionen, somit auch charakteristische Funktionen, beliebig genau auflösen. Daher können mit dünnen Gittern Probleme der CSG (“Constructive Solid Geometry”) aufgrund des adaptiven Zugangs einfach realisiert werden. Der Abschnitt 4.8.3 behandelt einfach Problemstellungen der CSG und gibt Abschätzungen zur Genauigkeit der Darstellung von Objekten.

#### 4.8.1 Gebietszerlegung und Finite Differenzen auf dünnen Gittern

Problemstellungen auf polynomial berandeten Gebieten, die sich als Vereinigung von  $d$ -dimensionalen Quadern darstellen lassen, können mittels Gebietszerlegungstechniken ebenfalls auf einfache Art und Weise mit der Methode der Finiten Differenzen auf dünnen Gittern behandelt werden. Durch diese Technik lassen sich auch mit dünnen Gittern Problemstellungen wie etwa die Berechnung einer Lösung auf einem L-Gebiet (vgl. Abbildung 15) als einfachstes Beispiel behandeln.

Wir verwenden hier als Grundlage das alternierende Verfahren von H.A. Schwarz [Sch70], vgl. auch [SBG96, AbEg98]. Dieses Verfahren ist ein überlappendes Gebietszerlegungsverfahren, womit das Lösungsschema zur Lösung einer partiellen Differentialgleichung<sup>8</sup> wie nachfolgend gezeigt aufgebaut ist (vgl. auch Algorithmus 1.1.1 in [SBG96]), wenn das Grundgebiet  $\bar{\Omega}$  in zwei Teilgebiete  $\bar{\Omega}_1$  und  $\bar{\Omega}_2$  zerlegt ist und  $\Omega_1 \cap \Omega_2 \neq \emptyset$  gilt (vgl. hierzu Abbildung 15). Dabei sei  $u^\nu$  die Lösung auf dem Gebiet  $\bar{\Omega}_1$  sowie  $v^\nu$  die Lösung auf dem Gebiet  $\bar{\Omega}_2$ . Die Verallgemeinerung des Schemas auf mehr

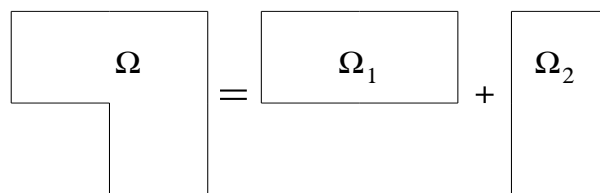


Abbildung 15: Zerlegung eines L-Gebiets in zwei rechteckige Teilgebiete

als zwei Gebiete ist dabei offensichtlich. Gegeben sei nun das Problem

$$\begin{aligned} Lu &= f && \text{in } \Omega \\ Bu &= g && \text{auf } \Gamma = \partial\bar{\Omega} \end{aligned}$$

<sup>8</sup>Die Diskretisierung einer PDE mittels Finiter Differenzen Operatoren auf dünnen Gittern wird in Abschnitt 5.2, Seite 75 beschrieben.

sowie eine Startlösung  $u^0$  auf  $\bar{\Omega}_1$  und  $\nu = 0$ .

1. Schritt: Setze  $\nu = \nu + 1$ . Lösen eines Problems auf  $\bar{\Omega}_2$ :

$$\begin{aligned} Lv^\nu &= f && \text{in } \Omega_2 \\ Bv^\nu &= g && \text{auf } \Gamma_2 \cap \Gamma = \partial\bar{\Omega}_2 \cap \partial\bar{\Omega} \\ v^\nu &= u^{\nu-1} && \text{auf } \Gamma_2 \cap \bar{\Omega}_1 \end{aligned}$$

2. Schritt: Lösen eines Problems auf  $\bar{\Omega}_1$ :

$$\begin{aligned} Lu^\nu &= f && \text{in } \Omega_1 \\ Bu^\nu &= g && \text{auf } \Gamma_1 \cap \Gamma = \partial\bar{\Omega}_1 \cap \partial\bar{\Omega} \\ u^\nu &= v^\nu && \text{auf } \Gamma_1 \cap \bar{\Omega}_2 \end{aligned}$$

Die Schritte 1 und 2 werden dabei abwechselnd so lange wiederholt ausgeführt, bis ein Abbruchkriterium erfüllt ist. Als mögliches Abbruchkriterium können dabei die Differenzen  $u^\nu - u^{\nu-1}$  und  $v^\nu - v^{\nu-1}$  verwendet werden: sind die Normen

$$\|u^\nu - u^{\nu-1}\| \quad \text{und} \quad \|v^\nu - v^{\nu-1}\| \quad (4.40)$$

jeweils kleiner als eine vorgegebene Schranke  $\varepsilon$ , so ist die numerische Lösung auf den Teilgebieten errechnet. Dabei kann man in (4.40) wahlweise die  $L_2$ -Norm, die  $L_\infty$ -Norm oder eine andere Norm verwenden.

Nach der Iteration werden die beiden numerischen Lösungen  $u^\nu$  und  $v^\nu$  auf den Teilgebieten  $\bar{\Omega}_1$  und  $\bar{\Omega}_2$  zu einer Lösung auf dem Gesamtgebiet  $\bar{\Omega}$  zusammengefaßt und dort der Fehler in den gewünschten Normen berechnet. Die Erweiterung von zwei auf  $s > 2$  Teilgebiete, die in ihrer Vereinigung das Gesamtrechnengebiet darstellen, stellt keinen grundlegenden Unterschied zu oben beschriebenen Verfahren mit zwei Gebieten dar. Nach der Berechnung der Lösung auf einem Teilgebiet  $\Omega_i$  mit  $1 \leq i \leq s$  werden die (Rand-)Daten all derjenigen Gebiete aktualisiert, die mit dem Gebiet  $\Omega_i$  überlappen. Im Abschnitt 7.3 finden sich Beispiele, die mit dieser Vorgehensweise berechnet worden sind.

#### 4.8.2 Behandlung krummlinig berandeter Gebiete

Aufgrund der Tensorproduktstruktur lassen sich Problemstellungen auf  $d$ -dimensionalen rechteckigen Gebieten auf dünnen Gittern aus gebietsgeometrischer Sicht gut behandeln. Besitzt das Gebiet einer gegebenen Aufgabenstellung keine Tensorproduktstruktur, so sind dünne Gitter nicht ohne weiteres dazu geeignet, eine Grundlage für die Diskretisierung dieser Problemstellung zu sein. Durch eine Transformation eines  $d$ -dimensionalen Würfels auf

das gegebene Gebiet der zu lösenden Problemstellung lassen sich jedoch auch auf krummlinig berandeten Gebieten dünne Gitter verwenden. Möglichkeiten für die Bestimmung einer Transformation eines  $d$ -dimensionalen Würfels auf ein krummlinig berandetes Gebiet wie etwa die transfinite Interpolation finden sich zusammen mit Anwendungen auf dünnen Gittern erstmals in [BuDo97, Dor97]. Methoden zur Erzeugung von Gebietstransformationen finden sich auch in [Bra77, Tho85, TWM85, HäTa86, AHET91, KnSt93, CaLi98].

Für die Methode der Finiten Elemente auf dünnen Gittern finden sich in [Dor97] zahlreiche Beispiele, wo auch ein Vergleich der einzelnen Transformationsmöglichkeiten durchgeführt ist. Dort wird die Koordinaten-Transformation in die Formulierung der schwachen Form auf dem Referenz-Gebiet gesteckt, was notationstechnisch gesehen zu einigen Vereinfachungen führt.

Ist eine partielle (elliptische) Differentialgleichung (2. Ordnung)

$$\sum_{i=1}^d \sum_{j=1}^d a_{ij}(\xi) \frac{\partial}{\partial \xi_i} \frac{\partial}{\partial \xi_j} u(\xi) + \sum_{i=1}^d b_i(\xi) \frac{\partial}{\partial \xi_i} u(\xi) + c(\xi) u(\xi) = f(\xi) \text{ in } P$$

(4.41)

$$u(\xi) = g(\xi) \text{ auf } \partial \bar{P}$$

auf einem krummlinig berandeten Gebiet  $P \subset \mathbb{R}^d$  gegeben und existiert eine Koordinatentransformation  $\psi$  mit

$$\begin{aligned} \psi : \bar{Q} &\rightarrow \bar{P} \\ \mathbf{x} &\mapsto \boldsymbol{\xi} = \psi(\mathbf{x}) \end{aligned}$$

wie in Abbildung 16 angegeben, so läßt sich die Differentialgleichung (4.41) in eine (elliptische) Differentialgleichung 2. Ordnung auf dem Referenzgebiet  $\bar{Q} = [0, 1]^d$  transformieren. Um die Differentialgleichung auf  $Q$  formulieren zu können, benötigen wir die zu  $\psi$  inverse Transformation  $\Xi = \psi^{-1} : \bar{P} \rightarrow \bar{Q}$ . Damit lassen sich die in (4.41) angegebenen Ableitungen bezüglich  $\boldsymbol{\xi}$  auf  $P$  angeben als Linearkombinationen von Ableitungen auf  $Q$  bezüglich  $\mathbf{x} = \Xi(\boldsymbol{\xi})$ :

$$\begin{aligned} \frac{\partial}{\partial \xi_k} &= \sum_{j=1}^d \frac{\partial}{\partial x_j} \frac{\partial x_j}{\partial \xi_k} \\ \frac{\partial^2}{\partial \xi_k \partial \xi_m} &= \sum_{j=1}^d \sum_{i=1}^d \left( \frac{\partial x_i}{\partial \xi_k} \frac{\partial x_j}{\partial \xi_m} \right) \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} + \sum_{j=1}^d \frac{\partial^2 x_j}{\partial \xi_k \partial \xi_m} \frac{\partial}{\partial x_j} \end{aligned}$$

Funktionen auf  $P$  transformieren sich vermöge  $\bar{f}(\mathbf{x}) = (f \circ \Xi)(\boldsymbol{\xi})$  zu Funktionen auf  $Q$ . Damit erhalten wir nach Transformation aus (4.41) die Diffe-

rentialgleichung

$$\begin{aligned}
& \sum_{k=1}^d \sum_{m=1}^d a_{km}(\Xi(\boldsymbol{\xi})) \left( \sum_{j=1}^d \sum_{i=1}^d \frac{\partial x_j}{\partial \xi_k} \frac{\partial x_i}{\partial \xi_m} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} + \sum_{j=1}^d \frac{\partial^2 x_j}{\partial \xi_k \partial \xi_m} \frac{\partial}{\partial x_j} \right) u(\Xi(\boldsymbol{\xi})) + \\
& + \sum_{k=1}^d \sum_{j=1}^d b_k(\Xi(\boldsymbol{\xi})) \frac{\partial x_j}{\partial \xi_k} \frac{\partial}{\partial x_j} u(\Xi(\boldsymbol{\xi})) + c(\Xi(\boldsymbol{\xi})) u(\Xi(\boldsymbol{\xi})) = \\
& = \sum_{j=1}^d \sum_{i=1}^d \underbrace{\left( \sum_{k=1}^d \sum_{m=1}^d a_{km}(\Xi(\boldsymbol{\xi})) \frac{\partial x_j}{\partial \xi_k} \frac{\partial x_i}{\partial \xi_m} \right)}_{\bar{a}_{ji}(\mathbf{x})} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} u(\mathbf{x}) + \\
& \sum_{j=1}^d \underbrace{\left( \sum_{k=1}^d \sum_{m=1}^d a_{km}(\Xi(\boldsymbol{\xi})) \frac{\partial^2 x_j}{\partial \xi_k \partial \xi_m} \right)}_{\bar{b}_{j,1}(\mathbf{x})} \frac{\partial}{\partial x_j} u(\mathbf{x}) \\
& \sum_{j=1}^d \underbrace{\left( \sum_{k=1}^d b_k(\Xi(\boldsymbol{\xi})) \frac{\partial x_j}{\partial \xi_k} \right)}_{\bar{b}_{j,2}(\mathbf{x})} \frac{\partial}{\partial x_j} u(\mathbf{x}) + \bar{c}(\mathbf{x}) u(\mathbf{x}) \\
& = \sum_{j=1}^d \sum_{i=1}^d \bar{a}_{ji}(\mathbf{x}) \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} u(\mathbf{x}) + \sum_{j=1}^d \bar{b}_j(\mathbf{x}) \frac{\partial}{\partial x_j} u(\mathbf{x}) + \bar{c}(\mathbf{x}) u(\mathbf{x}) \\
& = \bar{f}(\mathbf{x}) \tag{4.42}
\end{aligned}$$

auf  $Q$ , wobei  $\bar{b}_j(\mathbf{x}) = \bar{b}_{j,1}(\mathbf{x}) + \bar{b}_{j,2}(\mathbf{x})$  gesetzt wird.

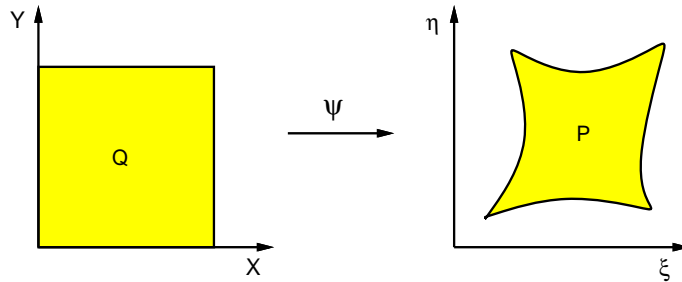


Abbildung 16: Transformation des Referenzgebiets  $Q$  auf das krummlinig berandete Gebiet  $P$  vermöge  $\psi$

In (4.42) wird eine Differentialgleichung auf  $P$  in eine Differentialgleichung auf  $Q$  transformiert. Aus implementiertechnischer Sicht ist der in (4.42) gegangene Weg jedoch unpraktisch, da die inverse Transformation  $\Xi$  in der Regel nicht bekannt ist. Jedoch lassen sich die partiellen Ableitungen auf  $Q$  analog zu (4.42) als Linearkombinationen der partiellen Ableitungen auf  $P$



formulieren. In Matrixschreibweise enthält die Matrix, die diese Linearkombination ausdrückt, partielle Ableitungen der Transformation  $\psi$ .

Für die weitere Betrachtung beschränken wir uns aus rein notationstechnischen Gründen auf den zweidimensionalen Fall ( $d = 2$ ) und setzen hier zur besseren Lesbarkeit  $\xi = \xi_1$  sowie  $\eta = \xi_2$ . Dann lautet die genannte Linearkombination in Matrixschreibweise

$$\begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \\ \frac{\partial^2 u}{\partial x^2} \\ \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial y^2} \end{pmatrix} = A \cdot \begin{pmatrix} \frac{\partial u}{\partial \xi} \\ \frac{\partial u}{\partial \eta} \\ \frac{\partial^2 u}{\partial \xi^2} \\ \frac{\partial^2 u}{\partial \xi \partial \eta} \\ \frac{\partial^2 u}{\partial \eta^2} \end{pmatrix}$$

mit der Transformationsmatrix

$$A \equiv \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \eta}{\partial x} & 0 & 0 & 0 \\ \frac{\partial \xi}{\partial y} & \frac{\partial \eta}{\partial y} & 0 & 0 & 0 \\ \frac{\partial^2 \xi}{\partial x^2} & \frac{\partial^2 \eta}{\partial x^2} & \left(\frac{\partial \xi}{\partial x}\right)^2 & 2\frac{\partial \xi}{\partial x} \frac{\partial \eta}{\partial x} & \left(\frac{\partial \eta}{\partial x}\right)^2 \\ \frac{\partial^2 \xi}{\partial x \partial y} & \frac{\partial^2 \eta}{\partial x \partial y} & \frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial y} & \frac{\partial \xi}{\partial x} \frac{\partial \eta}{\partial y} + \frac{\partial \xi}{\partial y} \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial y} \\ \frac{\partial^2 \xi}{\partial y^2} & \frac{\partial^2 \eta}{\partial y^2} & \left(\frac{\partial \xi}{\partial y}\right)^2 & 2\frac{\partial \xi}{\partial y} \frac{\partial \eta}{\partial y} & \left(\frac{\partial \eta}{\partial y}\right)^2 \end{pmatrix}$$

Da wir wie bereits angemerkt eine Darstellung der Ableitungen auf  $P$  suchen, die sich durch eine Linearkombination von Ableitungen auf  $Q$  ergibt, benötigen wir die Inverse  $A^{-1}$  der Matrix  $A$ , die aber aufgrund ihrer einfachen Struktur explizit angegeben werden kann. Somit können aus programmieretechnischer Sicht aus den Koeffizientenfunktionen in (4.41) sowie der punktweise betrachteten Inversen der Transformationsmatrix  $A$  einfach die Koeffizienten der transformierten Differentialgleichung vor dem Beginn des numerischen Lösens einer Problemstellung berechnet werden. Numerische Experimente zu diesem Lösungsansatz für elliptische Differentialgleichungen 2. Ordnung finden sich in Abschnitt 7.3. Die Anwendung dieser Technik auf zeitabhängige Problemstellungen stellt dabei prinzipiell kein Problem dar. Handelt es sich bei dem Grundgebiet der gegebenen Problemstellung um ein Gebiet, das nicht von der Zeit abhängig ist, so kann die genannte Vorgehensweise sofort ohne Modifikation übertragen und in jedem Zeitschritt angewendet werden, unabhängig davon, ob man eine explizite oder implizite Zeitdiskretisierung wählt. Betrachtet man Problemstellungen mit freien Oberflächen, so ist die Transformation  $\psi = \psi(t)$  ebenfalls zeitabhängig und

muß für jeden Zeitschritt neu bestimmt werden. Am prinzipiellen Vorgehen (Bestimmung der Transformation, Transformation der Gleichung und anschließendem Lösen des Problems) ändert sich jedoch auch in diesem Fall nichts. Folglich läßt sich beispielsweise das Stefan Problem (Problemstellung mit freien Oberflächen, vgl. etwa [Mei92, MaCh98]) mit der angegebenen Vorgehensweise auf dünnen Gittern lösen.

#### 4.8.3 Constructive Solid Geometry (CSG)

Dünne Gitter kann man neben der Darstellung für Funktionen auch zur Darstellung von Objekten einsetzen. Dazu definiert man sich für jedes geometrische Objekt  $F$ , das mit Hilfe dünner Gitter dargestellt werden soll, eine charakteristische Funktion  $\chi_F(\mathbf{x})$  vermöge

$$\chi_F(\mathbf{x}) = \begin{cases} 1 & , \mathbf{x} \in F \\ 0 & , \mathbf{x} \notin F \end{cases} \quad (4.43)$$

Diese Funktion kann nun mittels adaptiver dünner Gitter dargestellt werden und repräsentiert folglich das gewünschte geometrische Objekt  $F$ . Als Fehlerindikator dient das Kriterium  $\gamma_1$  in (2.31), wobei wir bei der Wahl der Maximumsnorm den maximalen Level fixieren müssen.

Bezüglich geometrischer Objekte kann man die auf diesen Objekten arbeitenden binären Operatoren (z.B. AND, OR, XOR, NOR oder COMPLEMENT) einführen. "Constructive Solid Geometry" (CSG) setzt nun unter anderem gerade diese Operationen ein, um aus einfachen geometrischen Körpern kompliziertere aufzubauen. Mit Hilfe der in Abschnitt 3.1 definierten Dünngitter-Verknüpfung läßt sich aus zwei verschiedenen dünnen Gitter ein drittes gewinnen, oder, wenn man die CSG als Anwendung betrachtet, aus zwei verschiedenen geometrischen Objekten ein drittes Objekt erhalten.

Dazu sei im folgenden angenommen, daß als Funktionswerte für eine Funktion auf einem dünnen Gitter nur die booleschen Werte 0 und 1 erlaubt seien. Verknüpft man zwei dünne Gitter mittels der in Abschnitt 3.1 eingeführten Dünngitter-Verknüpfung für CSG-Anwendungen, so sieht man sich an Gitterpunkten, die nur in einem der beiden zu verknüpfenden dünnen Gittern enthalten sind, bei der Anwendung eines beliebigen der mehrstelligen binären Operatoren dem Problem der Interpolation von Werten gegenüber. Die so erhaltenen Werte liegen in dem Intervall  $[0, 1]$  und müssen vermöge der Abbildung<sup>9</sup>

$$B(v) = \begin{cases} 0 & , v < \frac{1}{2} \\ 1 & , v \geq \frac{1}{2} \end{cases} \quad (4.44)$$

<sup>9</sup>Bei dieser Abbildung wird der Wert  $v = \frac{1}{2}$  auf 1 projiziert. Eine Projektion auf den Wert 0 ist ebenso denkbar.

auf einen der beiden booleschen Werte abgebildet werden. Dabei wird diese Abbildung angewandt auf Funktionswerte, nicht jedoch auf die Koeffizienten der  $d$ -dimensionalen hierarchischen Basis.

Mit Hilfe der Abbildung  $B$  können die binären Operatoren AND, OR, XOR, NOR und COMPLEMENT auf Funktionswerten auf dem Gitter  $G_3 = G_1 \circ^S G_2$  angewendet werden. Für jeden der binären Operatoren berechnet man hierzu die Werte  $B(v_1(\mathbf{x}))$  und  $B(v_2(\mathbf{x}))$  für  $v_1$  und  $v_2$  an allen Gitterpunkten  $\mathbf{x}$  des Gitters  $G_3$  und wendet dann auf diese Werte den binären Operator an. Hierarchisiert man das Ergebnis und löscht diejenigen Gitterpunkte beginnend von den Blättern, deren hierarchischer Überschuß 0 ist, so erhält man das dünne Gitter für das gesuchte geometrische Objekt.

Nachfolgend demonstrieren zwei Beispiele in zwei Dimensionen das beschriebene Vorgehen. Dabei ist die Erweiterung auf beliebige Dimensionen  $d$  klar, lediglich für eine einfachere graphische Darstellung wird an dieser Stelle auf den zweidimensionalen Fall zurückgegriffen.

Beginnen wollen wir mit der binären Operation OR. In den beiden Abbildungen 19 und 20 sind zwei Funktionen (Funktionen 1 und 2) zusammen mit ihren adaptiven dünnen Gittern gezeigt, wobei die Auflösung der Funktionen auf eine Leveltiefe von maximal 10 pro Raumdimension beschränkt ist. Das Ergebnis der Operation OR ist in Abbildung 21 zu sehen.

Als zweites Beispiel betrachten wir den binären Operator AND. Hierfür sind in den Abbildungen 22 und 23 die Operanden (Funktionen 3 und 4), wiederum mit zugehörigen dünnen Gittern der maximalen Leveltiefe 10 pro Raumdimension, angegeben. Das Ergebnis der binären Operation findet sich in Abbildung 24.

Nach diesen Beispielen, die die Vorgehensweise veranschaulichen sollen, betrachten wir an dieser Stelle den Aufwand, der zur Darstellung einiger einfacher geometrischer Körper notwendig ist. Hierzu wählen wir in drei Raumdimensionen die Körper Rechteck, Zylinder und Kugel aus.

**Beispiel 1:** (Rechteck): Die charakteristische Funktion für das Rechteck lautet

$$\chi_1(\mathbf{x}) = \begin{cases} 1 & , \quad |x| < \frac{2}{\pi}, |y| < \frac{3}{2\pi}, |z| < \frac{5}{2\pi} \\ 0 & , \quad \mathbf{x} \notin F \end{cases} \quad (4.45)$$

Das Rechteck ist so gewählt, daß die Flächen parallel zum Grundgebiet  $\Omega^3 = [-1, 1]^3$  liegen. Zu beachten ist hierbei, daß in dem hier gewählten Beispiel die Eckpunkte des Quaders nicht auf Dünngitter-Punkten zu liegen kommen. Somit muß aus allen drei Richtungen unabhängig voneinander jeweils ein logarithmischer Kostenfaktor aufgewendet werden, um das Rechteck bzw. seine charakteristische Funktion aufzulösen. Würde man die beiden Seiten bezüglich einer Raumdimension des Quaders exakt durch Punkte ei-

nes dünnen Gitters darstellen können, so hätte man einen konstanten Aufwand in dieser Richtung und würde so auf ein Problem mit einer um eins verminderten Dimension kommen. Diese Aussage ist allgemein für beliebige  $d$ -dimensionale Quader gültig. Tabelle 4 zeigt die Anzahl der Punkte, die zum Auflösen der charakteristischen Funktion  $\chi_1(\mathbf{x})$  für einen vorgegebenen maximalen Level  $L$  nötig sind. Dabei bedeuten im einzelnen:  $L$  der maximale Level,  $P$  die Anzahl benötigter Punkte im (adaptiven) dünnen Gitter,  $a$  eine Vergleichszahl,  $q_{PL3}$  der Quotient aus  $P$  dividiert durch  $L^3$ . Dabei fällt auf, daß die Anzahl benötigter Punkte in diesem Beispiel für jeden Level offenbar a priori angegeben werden kann (ohne Beweis): auf dem Level  $L$  benötigt man hier  $(2L + 1)^3$  Punkte. Somit folgt, daß der Aufwand  $\mathcal{O}(L^3)$  ist, oder genauer, exakt  $8 \cdot L^3 + 12 \cdot L^2 + 6 \cdot L + 1$  beträgt. Dies verdeutlicht auch die Spalte  $q_{PL3}$ , deren Werte für wachsende Level  $L$  gegen den Faktor 8 konvergieren. Zudem sind die Fehler, gemessen in der  $L_1$ -Norm und der  $L_2$ -Norm, angegeben.

Tabelle 4: maximaler Level und Anzahl nötiger Punkte für Beispiel 1

$L$	$L^3$	$P$	$a$	$a^3$	$q_{PL3}$	$L_1^S$	$L_2^S$
2	8	125	5	125	15.63	3.30 <sub>-1</sub>	8.10 <sub>-1</sub>
3	27	343	7	343	12.70	1.10 <sub>-1</sub>	2.94 <sub>-1</sub>
4	64	729	9	729	11.39	-7.17 <sub>-3</sub>	-2.00 <sub>-2</sub>
5	125	1331	11	1331	10.65	4.49 <sub>-3</sub>	1.25 <sub>-2</sub>
6	216	2197	13	2197	10.17	-1.54 <sub>-2</sub>	-4.31 <sub>-2</sub>
7	343	3375	15	3375	9.84	1.13 <sub>-2</sub>	3.14 <sub>-2</sub>
8	512	4913	17	4913	9.60	-2.01 <sub>-2</sub>	-5.62 <sub>-3</sub>
9	729	6859	19	6859	9.41	1.24 <sub>-2</sub>	3.46 <sub>-3</sub>
10	1000	9261	21	9261	9.26	7.43 <sub>-4</sub>	2.07 <sub>-3</sub>
11	1331	12167	23	12167	9.14	1.35 <sub>-3</sub>	3.75 <sub>-3</sub>
12	1728	15625	25	15625	9.04	5.11 <sub>-4</sub>	1.42 <sub>-3</sub>
13	2197	19683	27	19683	8.96	9.35 <sub>-5</sub>	2.60 <sub>-4</sub>
14	2744	24389	29	24389	8.89	-8.77 <sub>-6</sub>	2.44 <sub>-5</sub>
15	3375	29791	31	29791	8.83	4.23 <sub>-5</sub>	1.18 <sub>-4</sub>
17	4913	42875	35	42875	8.73	-2.66 <sub>-6</sub>	7.41 <sub>-6</sub>
20	8000	68921	41	68921	8.62	-2.59 <sub>-6</sub>	7.22 <sub>-6</sub>
25	15625	132651	51	132651	8.49		
30	27000	226981	61	226981	8.41		

**Beispiel 2:** (Zylinder) Das zweite Beispiel ist ein zur  $z$ -Achse achsenparalleler Zylinder mit Radius  $r = \pi^{-1}$  auf  $\bar{\Omega}^3 = [0, 1]^3$ . Die Höhenbegrenzung bezüglich der  $z$ -Achse sind hierbei  $\frac{1}{3\pi} < z < 1 - \frac{1}{3\pi}$ , die Achse des Zylinders ist  $(\frac{1}{2}, \frac{1}{2}, z)$ . Tabelle 5 zeigt für die einzelnen Level die verwendete Anzahl an Punkten in dem entsprechenden adaptiven dünnen Gitter.  $q_{PL}$  stellt dabei den Quotienten aus  $P$  und der Größe  $2^L \cdot L^2$  dar. Da dieser Quotient in etwa

einer Konstanten entspricht, ist der Aufwand in diesem Fall  $\mathcal{O}(2^L \cdot L^2)$ . Abbildung 5 zeigt eine Darstellung des Zylinders auf einem adaptiven dünnen Gitter mit maximalem Level 7. Deutlich erkennbar ist dabei, daß in Achsenrichtung des Zylinders relativ wenig Gitterpunkte (erkennbar an den zur Darstellung notwendigen zweidimensionalen "Schichten", in denen jeweils ein Kreis aufgelöst ist) verwendet werden (hieraus ergibt sich einer der beiden Faktoren  $L$  in der Ordnungsangabe der verwendeten Punkte).

Tabelle 5: maximaler Level und Anzahl nötiger Punkte für Beispiel 2

$L$	$2^L \cdot L^2$	$P$	$q_{PL}$	$L_1^S$	$L_2^S$
3	72	225	3.13	1.72 <sub>-2</sub>	1.74 <sub>-2</sub>
4	256	931	3.64	3.56 <sub>-2</sub>	3.43 <sub>-2</sub>
5	800	2997	3.75	5.55 <sub>-3</sub>	5.51 <sub>-3</sub>
6	2304	9075	3.94	3.57 <sub>-3</sub>	3.57 <sub>-3</sub>
7	6272	23829	3.80	2.36 <sub>-3</sub>	2.35 <sub>-3</sub>
8	16384	61875	3.78	6.59 <sub>-4</sub>	6.57 <sub>-4</sub>
9	41472	156349	3.77	1.55 <sub>-3</sub>	1.55 <sub>-3</sub>
10	102400	393851	3.85	1.05 <sub>-3</sub>	1.05 <sub>-3</sub>
11	247808	963585	3.89	7.49 <sub>-4</sub>	7.47 <sub>-4</sub>

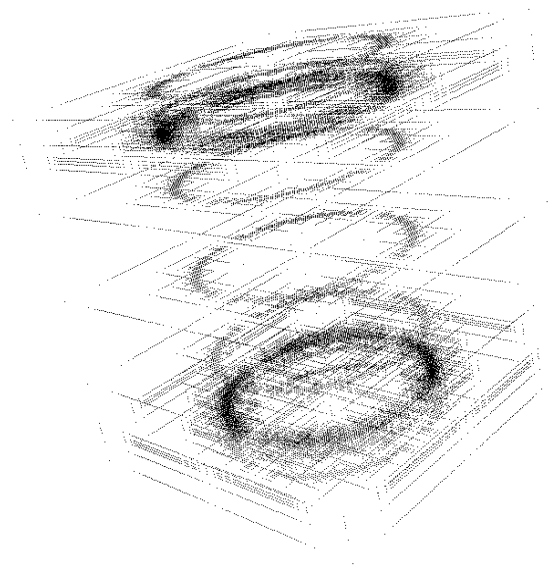


Abbildung 17: dünnes Gitter für den Zylinder aus Beispiel 2 (Level 7)

**Beispiel 3:** (Kugel) In diesem Fall betrachten wir zwei Kugeln mit den Radien  $r_1 = \frac{1}{2}$  und  $r_2 = \frac{2}{\pi}$  auf  $\bar{\Omega}^3 = [-1, 1]^3$ , wobei der Schwerpunkt der Kugeln im Ursprung liegt. In der nachfolgenden Tabelle 6 bezeichnen  $P_1$  und  $P_2$  die Punkte bezüglich der Radien  $r_1$  und  $r_2$  sowie  $q_1$  und  $q_2$  die Quotienten  $P_1$  dividiert durch  $L \cdot 2^{2L}$  bzw.  $P_2$  dividiert durch  $L \cdot 2^{2L}$ . Für die

Kugel mit Radius  $r_2$  sind zudem die Fehler, gemessen in der  $L_1$ -Norm und der  $L_2$ -Norm, angegeben. Man erkennt, daß der Aufwand zur Darstellung von Kugeln mittels dünner Gitter bei  $\mathcal{O}(L \cdot 2^{2L})$  liegt.

Tabelle 6: maximaler Level und Anzahl nötiger Punkte für Beispiel 3

$L$	$L \cdot 2^{2L}$	$P_1$	$q_1$	$P_2$	$q_2$	$L_1^S$	$L_2^S$
2	32	27	0.84	125	3.91	$2.06_{-1}$	$1.04_{-1}$
3	192	125	0.65	275	1.43	$-1.85_{-1}$	$-8.54_{-2}$
4	1024	721	0.70	1645	1.61	$7.49_{-2}$	$3.67_{-2}$
5	5120	3753	0.73	8673	1.69	$1.02_{-2}$	$4.92_{-3}$
6	24576	19499	0.79	36953	1.50	$-3.01_{-3}$	$-1.45_{-3}$
7	114688	103225	0.90	178127	1.55	$2.81_{-4}$	$1.35_{-4}$
8	524288	493245	0.94	830799	1.58	$3.29_{-4}$	$1.58_{-4}$

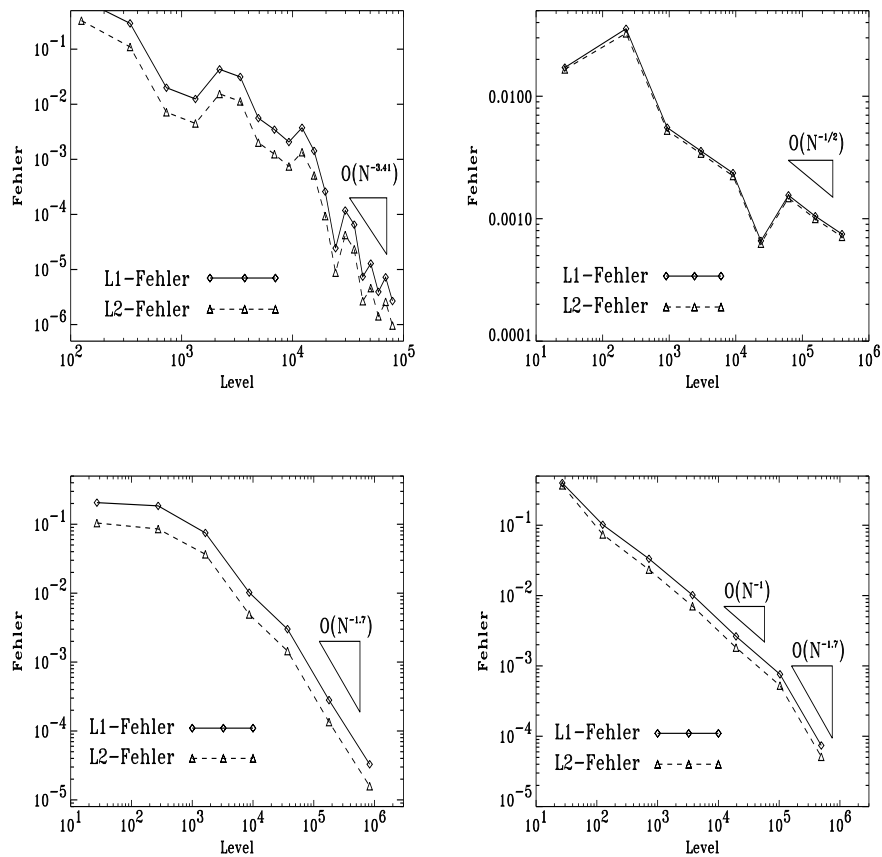


Abbildung 18: Konvergenz in der  $L_1$ -Norm und der  $L_2$ -Norm für die Beispiele 1 (Rechteck, oben links), 2 (Zylinder, oben rechts) und 3 (Kugel mit Radius  $r = 1/\pi$ , links unten, Kugel mit Radius  $r = 1/2$ , unten rechts)

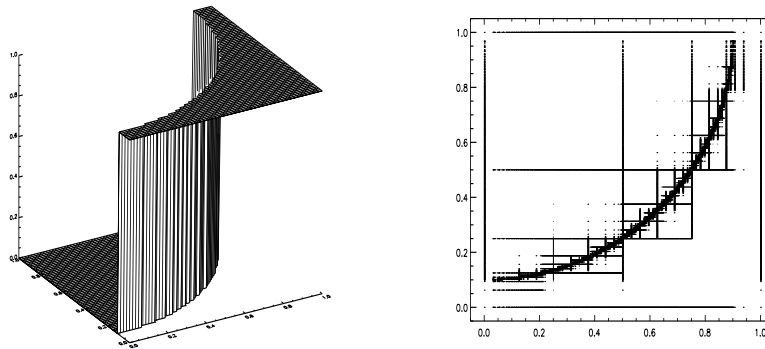


Abbildung 19: Beispiel 1: Funktion 1 zusammen mit einem adaptiven Gitter (maximaler Level eines Gitterpunkts: 10)

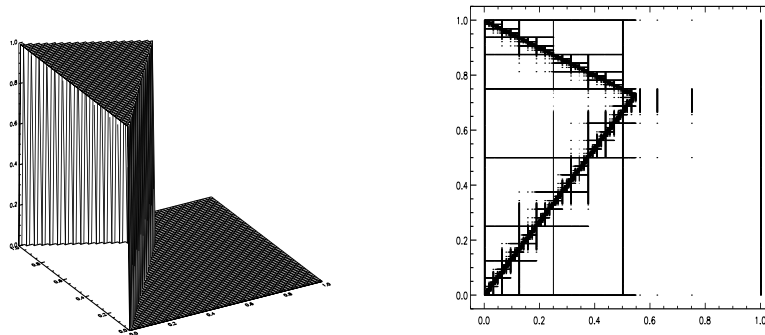


Abbildung 20: Beispiel 2: Funktion 2 zusammen mit einem adaptiven Gitter (maximaler Level eines Gitterpunkts: 10)

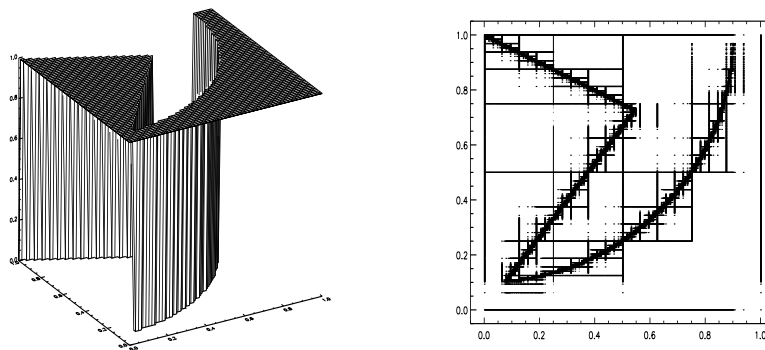


Abbildung 21: Beispiel 1: Ergebnisfunktion der Operation OR der beiden Funktionen 1 und 2 zusammen mit einem adaptiven Gitter (maximaler Level eines Gitterpunkts: 10)

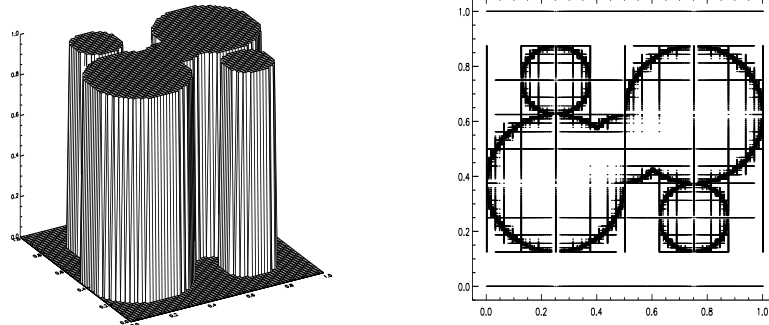


Abbildung 22: Beispiel 2: Funktion 3 zusammen mit einem adaptiven Gitter (maximaler Level eines Gitterpunkts: 10)

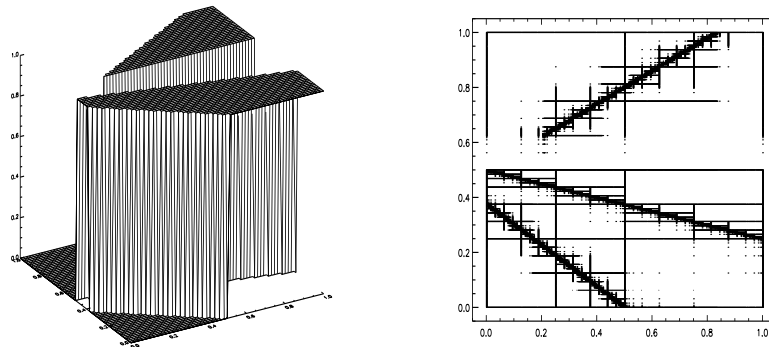


Abbildung 23: Beispiel 2: Funktion 4 zusammen mit einem adaptiven Gitter (maximaler Level eines Gitterpunkts: 10)

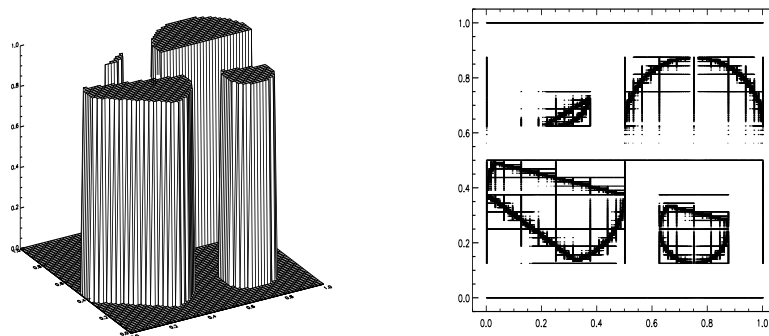


Abbildung 24: Beispiel 2: Ergebnisfunktion der Operation AND der beiden Funktionen 3 und 4 zusammen mit einem adaptiven Gitter (maximaler Level eines Gitterpunkts: 10)



## 5 Anmerkungen zur Lösung partieller Differentialgleichungen

### 5.1 Das unidirektionale Prinzip – Ausnutzung der Tensorproduktstruktur

Wie bereits aus Formel (2.12) und der anschließenden Diskussion zu entnehmen ist, können die  $d$ -dimensionalen Operatoren  $\mathbf{H}$  in (2.16) sowie deren Inverse  $\mathbf{E}$  vermöge  $d$  aufeinanderfolgender, eindimensionaler Operatoren dargestellt werden. Ebenso bezieht sich die Anwendung der Operatoren  $\mathbf{D}_{jj}^S$  bzw.  $\mathbf{D}_j^{\pm,0}$  lediglich auf eindimensionale Strukturen. Somit können die Operatoren  $\mathbf{D}_{jj}^S$ ,  $\Delta^S$ ,  $\mathbf{D}_j^{S,\pm}$  sowie  $\mathbf{D}_j^{S,0}$  aus eindimensionalen Operatoren zusammengesetzt werden, was die Implementierung der vorgestellten Diskretisierungsmethode erheblich vereinfacht. Mit Hilfe der Algebra für Operatoren auf Funktionen lassen sich so beliebige (elliptische, parabolische oder hyperbolische) Differentialoperatoren aufbauen und mittels eindimensionaler Operatoren diskretisieren.

Damit ist wie im Fall der Finiten Elemente die Anwendung der Diskretisierungsmatrix in  $\mathcal{O}(N)$  Operationen möglich, da sie lediglich aus der Komposition obiger Operatoren aufgebaut ist. Dabei bezeichnet  $N$  hier die Anzahl der Punkte in einem dünnen Gitter. Dabei sei angemerkt, daß die Diskretisierungsmatrizen im Fall der Finiten Differenzen wie auch im Fall der Finiten Elemente keine Bandstruktur besitzen und somit nicht dünn besiedelt im Sinne  $\mathcal{O}(1)$  Einträge pro Matrixzeile sind.

### 5.2 Diskretisierung elliptischer Differentialoperatoren 2. Ordnung

Zur Diskretisierung einer elliptischen Differentialgleichung

$$\begin{aligned} Lu &= f && \text{in } \Omega = (0, 1)^d \\ u &= g_D && \text{auf } \Gamma_D \\ \frac{\partial u}{\partial n} &= g_N && \text{auf } \Gamma_N \end{aligned} \tag{5.1}$$

auf dem  $d$ -dimensionalen Gebiet  $\bar{\Omega} = [0, 1]^d$ , die entweder auf dem  $d$ -dimensionalen Hyperwürfel gegeben ist oder wie in Abschnitt 4.8.2 auf diesen transformiert worden ist, verwenden wir nachfolgend die zuvor eingeführten Finiten Differenzen Operatoren auf dünnen Gittern. Zudem benützen wir die Multiplikation, skalare Multiplikation, Addition und Subtraktion auf Dünngitter-Funktionen und wenden damit die in Abschnitt 3.2 eingeführte Algebra für Operatoren auf Funktionen an. Der Operator  $L$  in (5.1) ist ein elliptischer Differentialoperator 2. Ordnung wie in (4.1) angegeben,  $g_D$

die Dirichlet–Randbedingung auf dem Dirichletrand  $\Gamma_D$  des Grundgebiets  $\bar{\Omega}$  sowie  $\partial u/\partial n$  die Neumann–Randbedingung auf dem Neumannrand  $\Gamma_N$ .

Der Differentialoperator  $L$  in (4.1) wird dabei vermöge

$$\begin{aligned}\tilde{L}^S &= \sum_{i=1}^d \tilde{a}_{ii} \cdot \mathbf{D}_{ii}^S + \sum_{i=1}^d \sum_{\substack{j=1 \\ j \neq i}}^d \tilde{a}_{ij}(\mathbf{x}) \cdot \mathbf{D}_i^S \mathbf{D}_j^S + \sum_{i=1}^d \tilde{b}_i(\mathbf{x}) \cdot \mathbf{D}_i^S + \tilde{c}(\mathbf{x}) \\ &= \sum_{i=1}^d \tilde{a}_{ii} \cdot \mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i + \sum_{i=1}^d \tilde{b}_i(\mathbf{x}) \cdot \mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i + \\ &\quad + \sum_{i=1}^d \sum_{\substack{j=1 \\ j \neq i}}^d \tilde{a}_{ij}(\mathbf{x}) \cdot \mathbf{E}_i \circ \mathbf{D}_i \circ \mathbf{H}_i \circ \mathbf{E}_j \circ \mathbf{D}_j \circ \mathbf{H}_j + \tilde{c}(\mathbf{x})\end{aligned}\quad (5.2)$$

diskretisiert. Zudem betrachten wir auch den Operator

$$\hat{L} = \sum_{i=1}^d \sum_{j=1}^d \frac{\partial}{\partial x_i} \left( \hat{a}_{ij} \frac{\partial}{\partial x_j} \right) + \sum_{i=1}^d \hat{b}_i(\mathbf{x}) \frac{\partial}{\partial x_i} + \hat{c}(\mathbf{x}) ,$$

den wir vermöge

$$\hat{L}^S = \sum_{i=1}^d \sum_{j=1}^d \mathbf{D}_i^S \left( \hat{a}_{ij}(\mathbf{x}) \mathbf{D}_j^S \right) + \sum_{i=1}^d \hat{b}_i(\mathbf{x}) \cdot \mathbf{D}_i^S + \hat{c}(\mathbf{x}) \quad (5.3)$$

diskretisieren. Dabei dienen die Kennzeichnungen  $\tilde{\phantom{x}}$  und  $\hat{\phantom{x}}$  lediglich der Unterscheidung der beiden verschiedenen Diskretisierungen der elliptischen Differentialoperatoren zweiter Ordnung. Im Fall genügend glatter Koeffizientenfunktionen können die beiden Operatoren  $\tilde{L}$  und  $\hat{L}$  ineinander überführt werden.

Dabei stellen die einzelnen Ableitungsoperatoren  $\mathbf{D}_{ii}^S$  und  $\mathbf{D}_i^S$  die zuvor eingeführten Finiten Differenzen Operatoren auf dünnen Gittern dar. Bei der Diskretisierung der Terme  $\mathbf{D}_i^S \left( \hat{a}_{ij}(\mathbf{x}) \mathbf{D}_j^S \right)$  des Operators  $\hat{L}^S$  verwenden wir entweder für die erste der beiden durchzuführenden ersten Ableitungen (hier  $\mathbf{D}_j^S$ ) die Vorwärtsdifferenz, und für die zweite durchzuführende erste Ableitung (hier  $\mathbf{D}_i^S$ ) die Rückwärtsdifferenz (oder umgekehrt), oder wir verwenden für beide (ersten) Ableitungen die zentrale Differenz. Dabei ist darauf zu achten, daß  $\hat{a}_{ij}(\mathbf{x})$  hinreichend glatt ist. In den nachfolgenden numerischen Beispielen, die elliptische Differentialgleichungen 2. Ordnung behandeln, wurden die Ergebnisse mittels eines Jacobi-vorkonditionierten BiCGSTAB [Vor95] gewonnen.

### 5.3 Iterativer Löser für die Problemstellungen: BiCGSTAB

Die in Abschnitt 4.2 eingeführten Finiten Differenzen Operatoren führen auch für selbstadjungierte Differentialoperatoren auf nicht-symmetrische Diskretisierungsmatrizen. Dies erkennt man daran, daß zum Aufbau der Operatoren Terme der Bauart  $\mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i$  verwendet werden, wobei  $\mathbf{E}_i = \mathbf{H}_i^{-1}$ , nicht jedoch  $\mathbf{E}_i = \mathbf{H}_i^T$  gilt. Somit ist jeder Einzelsummand des Gesamtoperators nicht-symmetrisch.

Daher muß zur Lösung der linearen Gleichungssysteme, die bei der Diskretisierung von Differentialoperatoren vom Typ (4.1) entstehen, ein iterativer Löser eingesetzt werden, der nicht-symmetrische Problemstellungen behandeln kann. Für die später angeführten numerischen Ergebnisse in Kapitel 7 verwenden wir im Rahmen dieser Arbeit BiCGSTAB [Vor95], sowohl in vorkonditionierter als auch in nicht-vorkonditionierter Version. Daneben sind selbstverständlich andere iterative Löser für nicht-symmetrische Problemstellungen wie etwa CNR, GMRES, CGNE, CGNR einsetzbar, um nur einige zu nennen, die nachfolgend jedoch nicht verwendet werden. Insbesondere wird hier kein Vergleich benötigter Iterationszahlen möglicher iterativer Löser angegeben.

Die Konvergenzgeschwindigkeit des BiCGSTAB-Verfahrens hängt stark von der Kondition der Matrix der zu lösenden Problemstellung ab. Die verallgemeinerte Kondition der Diskretisierungsmatrizen des Laplace-Problems wächst dabei nach Tabelle 2 wie  $\mathcal{O}(h^{-2})$ , wobei  $h$  an dieser Stelle wiederum die feinste auftretende Schrittweite des dünnen Gitters (Randmaschenweite) bezeichnet. Somit ist zu erwarten, daß bei feineren Diskretisierungen die Anzahl der benötigten Iterationen zum Lösen der betrachteten Problemstellung erheblich ansteigt. Dieses Verhalten werden wir im Abschnitt 7.2.1 der numerischen Resultate beobachten können, wo für das Laplace-Problem L1 (vgl. Abschnitt 7.2.1 für die genaue Problemstellung) die Iterationszahlen sowohl für ein vorkonditioniertes als auch ein nicht-vorkonditioniertes BiCGSTAB-Verfahren miteinander verglichen werden. Im Gegensatz zum vollen Gitter ist im Fall dünner Gitter eine Vorkonditionierung auf sehr einfache Art und Weise möglich. Die hierzu notwendigen Schritte werden im nachfolgenden Abschnitt 5.4 vorgestellt. Ein optimales Verfahren zur Lösung der auftretenden linearen Gleichungssysteme wäre ein Mehrgitter-Verfahren, das aber für unsere Anwendung noch entwickelt werden muß.

## 5.4 Jakobi–Vorkonditionierung

### 5.4.1 Vorkonditionierung von Dirichlet–Problemen

Auf vollen Gittern hat eine Diagonalskalierung (Jakobi–Vorkonditionierung) mit der Inversen der Diskretisierungsmatrix keine Auswirkung auf die Kondition, da in diesem Fall die Diagonalmatrix lediglich aus einer Konstanten multipliziert mit der Einheitsmatrix besteht. Dieser Sachverhalt ändert sich jedoch, wenn man die entsprechenden Matrizen auf dünnen Gittern betrachtet. Eine einfache Rechnung ergibt, daß die Einzeloperatoren  $\mathbf{H}_j$  und  $\mathbf{E}_j$  keinerlei Auswirkung auf den Hauptdiagonaleintrag des Gesamtoperators  $\mathbf{D}_{jj}^S$  haben, und somit nur die Teiloperatoren  $\mathbf{D}_{jj}$  zum Hauptdiagonaleintrag beitragen. Sei nun  $h_{l_j}$  die lokale Maschenweite zum nächsten Nachbarn eines inneren Gitterpunkts  $\mathbf{x}_{1,i}$  eines dünnen Gitters in Raumdimension  $j$ . Dann ergibt sich der Hauptdiagonaleintrag zu

$$\sum_{j=1}^d \frac{2}{h_{l_j}^2}, \quad (5.4)$$

der offensichtlich kein Vielfaches der Einheitsmatrix darstellt. Somit ist zu erwarten, daß eine Diagonalskalierung mit der Inversen der Diagonalmatrix der Diskretisierungsmatrizen im Fall der Finiten Differenzen Operatoren auf dünnen Gittern eine Auswirkung auf die Kondition haben wird. In Tabelle 7 sind in zwei Raumdimensionen die mit MATLAB [Mat97] numerisch berechneten Eigenwerte der diagonalskalierten Diskretisierungsmatrizen zu finden, die ebenso wie die Eigenwerte in Abschnitt 4.4.2 reell sind. Man erkennt dabei, daß die verallgemeinerte Kondition nur noch mit der Ordnung  $\mathcal{O}(h^{-1})$  anstatt von  $\mathcal{O}(h^{-2})$  wie in Tabelle 2 wächst. Der betragsgrößte Eigenwert  $\Delta$  strebt offensichtlich gegen den konstanten Wert 2, wohingegen sich der betragskleinste Eigenwert mit steigendem Level jeweils annähernd halbiert, womit die beobachtete erste Ordnung für die verallgemeinerte Kondition zu erklären ist. Durch die Diagonalskalierung stellt sich im Fall dünner Gitter somit ein Vorkonditionierungseffekt ein, wobei die Kondition der Matrizen um eine Ordnung im Vergleich zu den nicht vorkonditionierten korrespondierenden Diskretisierungsmatrizen gesenkt wird.

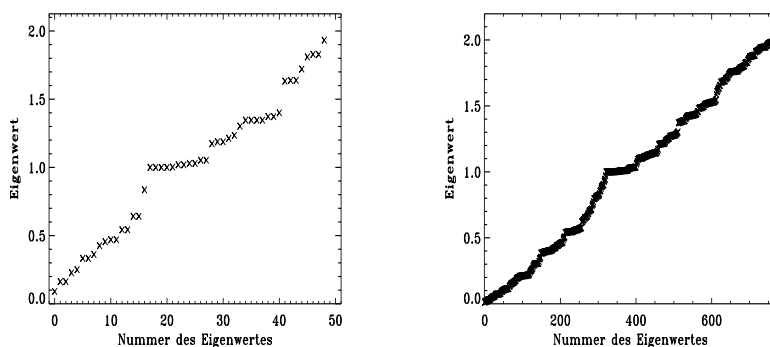
Bei den numerischen Beispielen in Kapitel 7 wird die Auswirkung der Vorkonditionierung auf die benötigte Anzahl von Iterationen untersucht werden. In der nachfolgenden Abbildung 25 sind die Spektren der diagonalskalierten Matrizen auf den Levels 4 und 7 zu sehen.

### 5.4.2 Vorkonditionierung von Neumann–Problemen

Die zuvor angegebene Vorgehensweise ist eine Vorkonditionierung für Dirichlet–Probleme. Für Neumann–Probleme kann man analog einen Jacobi–

Tabelle 7: Eigenwerte von  $\Delta^S$  in zwei Dimensionen,  $\bar{\Omega} = [0, 1]^2$ 

Level	$\lambda$	$q_\lambda$	$\Lambda$	$\kappa$	$q_\kappa$
2	4.258343 <sub>-1</sub>	—	1.400000	3.29	—
3	1.917543 <sub>-1</sub>	2.22	1.785797	9.31	2.83
4	9.018789 <sub>-2</sub>	2.13	1.932598	21.43	2.30
5	4.360928 <sub>-2</sub>	2.07	1.979593	45.39	2.12
6	2.142375 <sub>-2</sub>	2.04	1.993880	93.07	2.05
7	1.061406 <sub>-2</sub>	2.02	1.998164	188.26	2.02
8	5.281384 <sub>-3</sub>	2.01	1.999447	378.58	2.01

Abbildung 25: Eigenwerte der diagonalskalierten Diskretisierungsmatrizen auf dünnen Gittern für Level 4 (links) und Level 7 (rechts),  $d = 2$ 

Vorkonditionierer erhalten, wenn man die Diskretisierung des Neumann-Randes wie in Abschnitt 4.7 besprochen berücksichtigt. Zusätzlich zu der Größe (5.4), die bereits für innere Gitterpunkte die Hauptdiagonale darstellt, addieren wir hierzu den Ausdruck

$$\frac{1}{h_{l_j}}$$

für Neumann–Randknoten, falls der betreffende Randknoten bezüglich der  $j$ -ten Raumdimension einen Randknoten darstellt. Man beachte hierbei, daß ein Randknoten, der etwa in zwei Raumdimensionen an einer Kante des Gebiets liegt, Randknoten bezüglich einer Richtung sein kann, während er bezüglich der zweiten Richtung einen inneren Knoten darstellt. Daher wäre in diesem Fall nur bezüglich der erstgenannten Richtung der entsprechende Term für die Neumann–Ableitung zu addieren. Hierdurch erreichen wir wiederum eine Verbesserung der Ordnung der Kondition von  $\mathcal{O}(h^{-2})$  auf  $\mathcal{O}(h^{-1})$ . Ein Neumann–Beispiel, bei dem zusätzlich zu den numerischen Er-

gebissen die zugehörigen Iterationszahlen auf den einzelnen Leveln angegeben sind, ist das Problem L5, das später in dem Abschnitt 7 (“Numerische Resultate”) besprochen wird.

### 5.5 Ein selbstadaptiver Algorithmus zur Lösung elliptischer Differentialgleichungen 2. Ordnung

Neben glatten Beispielen, an denen die theoretisch zu erwartenden Konvergenzordnungen für einfache Problemstellungen nachzuweisen sind, werden wir uns in den numerischen Beispielen auch mit Funktionen auseinandersetzen, die Singularitäten aufweisen und starke Gradienten oder innere Grenzschichten bzw. Randgrenzschichten besitzen. Da dort zum Teil dünne Gitter (zumindest auf größeren Leveln) nicht in der Lage sind, die jeweilige Funktion vernünftig darzustellen, muß – ausgehend von einem Startgitter, das ein vorgegebenes reguläres dünnes Gitter ist – ein selbstadaptiver Algorithmus Gitterpunkte dort generieren, wo dies zur Fehlerreduktion der Lösung (gemessen in einer vorgegebenen Norm) notwendig ist. Die von uns gewählte Vorgehensweise ist dabei wie folgt: in Anlehnung<sup>10</sup> an Formel (2.28) für die Abschätzung des Interpolationsfehler in der Maximumsnorm auf dünnen Gitter (vgl. hierzu Lemma 2.6) berechnen wir den Ausdruck

$$F(\mathbf{x}_{1,i}) := \frac{2 \cdot D^2 u(\mathbf{x}_{1,i})}{8^d} \cdot 2^{-\|\mathbf{l}_i\|_1} \quad (5.5)$$

als Kriterium für den Fehler im Punkt  $\mathbf{x}_{1,i}$ , wobei  $D^2$  nach (2.10) definiert ist. Ist dieser Wert größer als eine vorgegebene Schranke  $\varepsilon$ , so werden – falls die entsprechenden Knoten nicht bereits im Gitter vorhanden sein sollten – die Sohnknoten bezüglich aller Raumdimensionen eingefügt. Wird für die vorgegebene Genauigkeit  $\varepsilon$  kein neuer Knoten eingefügt, gilt also für alle berechneten Werte  $F(\mathbf{x}_{1,i}) > \varepsilon$ , so wird  $\varepsilon$  mit einem vorgegebenen Faktor skaliert (bei uns analog mit dem Abfall  $1/4$  des hierarchischen Überschusses). Dieses Vorgehen wird solange durchgeführt, bis eine (untere) Schranke  $\varepsilon_{\text{ende}}$  erreicht wird und das Problem bis auf die gewünschte Genauigkeit berechnet ist. Abbildung 26 zeigt ein Flußdiagramm des besprochenen Verfeinerungs-Schemas. Der Nachweis, daß das gewählte Fehlerkriterium (5.5) einen Fehlerindikator darstellt, muß an dieser Stelle entfallen. Ein anderes mögliches Kriterium für eine adaptive Steuerung wäre etwa

$$F(\mathbf{x}_{1,i}) = C(\mathbf{l}, d) \cdot \sum_{j=1}^d D^{2\mathbf{e}_j} u(\mathbf{x}_{1,i}) \quad (5.6)$$

<sup>10</sup>Man beachte hierbei, daß die angegebene Formel für den Fehlerterm  $F(\mathbf{x}_{1,i})$  nicht mit der Formel für die Abschätzung des Interpolationsfehler in der Maximumsnorm identisch ist.

wobei  $\mathbf{e}_j = (0, \dots, 0, 1, 0, \dots, 0)$  ist und sich die 1 an der  $j$ -ten Stelle befindet ( $j$ -ter Einheitsvektor). Damit würde ein punktabhängig gewichteter Laplace-Operator die Steuerung der Adaptivität übernehmen. Dies entspricht der Betrachtung des punktwise skalierten euklidischen Skalarprodukts  $C(\mathbf{1}, d) \cdot \langle \nabla u, \nabla u \rangle$  im Finite Element Sinn. Betrachtet man in (5.6) die Ausdrücke  $D^{2\mathbf{e}_j} u(\mathbf{x}_{1,i})$  einzeln, so kann man in Abhängigkeit der Größe dieser Werte auch eine Verfeinerung wählen, die an  $u(\mathbf{x}_{1,i})$  lediglich die Sohnknoten bezüglich der Raumdimension  $j$  für  $1 \leq j \leq d$  einfügt. Wir werden in den numerischen Beispielen in Kapitel 7 ausschließlich das Kriterium (5.5) zur selbstadaptiven Steuerung verwenden.



Abbildung 26: schematische Darstellung des selbstadaptiven Algorithmus

## 6 Implementierung und Datenstrukturen

Durch die Definition mittels eines Tensorproduktansatzes ermöglichen dünne Gitter auf einfache Weise adaptive Verfeinerungen. Damit ist es sinnvoll, auch die Operatoren, beispielsweise Finite Differenzen oder die Transformationen Hierarchisierung und Enthierarchisierung, auf dünnen Gittern so zu implementieren, daß sie nicht nur auf regulären dünnen Gittern, sondern auch auf adaptiv verfeinerten dünnen Gittern anwendbar sind, also auch im nicht-äquidistanten Fall. Implementiert man folglich die einzelnen Operatoren mit schrittweitenabhängigen Koeffizienten (wo dies operatortechnisch nötig ist), so hat man auf einfache Art und Weise einen vollständig adaptiven Code für die Berechnung von partiellen Differentialgleichungen und Integralgleichungen sowie Interpolations- oder Integrations-Aufgaben. Ebenso ist die Behandlung von Fragestellungen aus dem Gebiet der “Constructive Solid Geometry” (CSG, vgl. Abschnitt 4.8.3) einfach möglich<sup>11</sup>. Eine wichtige Frage aus der Sicht des Programmierers für die Implementierung eines Dünngitter-Codes – wie auch jeden anderen voll-adaptiven Codes – ist die Wahl einer geeigneten Datenstruktur, die folgende Eigenschaften haben sollte:

- einfacher Zugriff auf beliebige Daten des dünnen Gitters, am besten mit Kosten der Ordnung  $\mathcal{O}(1)$  für einen Zugriff
- einfache Handhabbarkeit der Adaptivität
- geringer Speicheraufwand für einen einzelnen Gitterpunkt
- der Zugriff auf hierarchische Väter und Söhne sowie auf direkte Nachbarn muß einfach und schnell sein und sollte wiederum bzgl. des Aufwands  $\mathcal{O}(1)$  sein (wichtig für Finite Differenzen)

Datenstrukturen, die den genannten Anforderungen mehr oder weniger gut entsprechen, sind binäre Bäume und Hash-Tabellen, die wir in diesem Abschnitt auf ihre Anwendbarkeit bezüglich dünner Gitter untersuchen wollen.

### 6.1 Binäre Bäume

In einer Raumdimension können (adaptive) Gitter effizient in einem binären Baum gespeichert werden, wie es links in Abbildung 27 zu sehen ist. Die Erweiterung von einer auf  $d$  Dimensionen benutzt gewissermaßen eine Tensorprodukt-darstellung von Bäumen: in  $d$  Dimensionen mit  $d > 1$  erhält man binäre Bäume von binären Bäumen als Datenstruktur. Dabei stellen

<sup>11</sup> Diese Liste von möglichen Anwendungsbereichen ist selbstverständlich nicht vollständig.



die Knoten des Baumes, der  $d$  Dimensionen repräsentieren soll,  $(d - 1)$ -dimensionale Bäume dar. Ein Beispiel vermag dies zu verdeutlichen: aus-

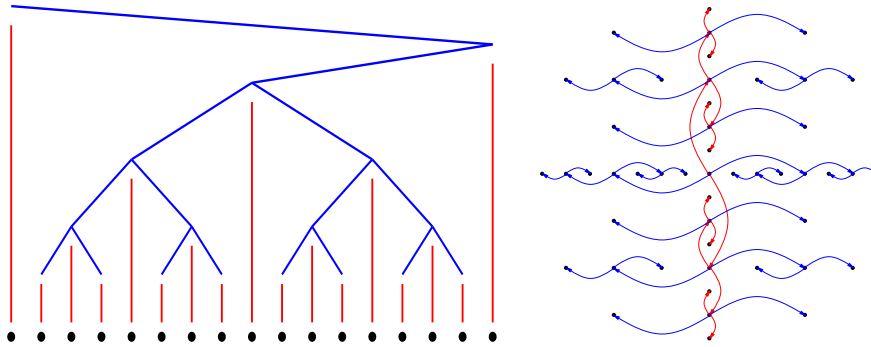


Abbildung 27: eindimensionales Gitter und seine Darstellung mit Hilfe eines binären Baums sowie zweidimensionales dünnes Gitter und seine Darstellung mit Hilfe eines binären Baums, dessen Knoten eindimensionale binäre Bäume darstellen (nur innere Gitterpunkte sind dargestellt)

gehend von eindimensionalen Bäumen, die Gitterlinien und somit eindimensionale Gitter repräsentieren, bestehen die Knoten eines Baumes, der zweidimensionale (dünne) Gitter darstellen soll, aus den  $1D$ -Bäumen. Die Knoten von binären Bäumen, die drei Dimensionen repräsentieren sollen, bestehen ihrerseits aus zweidimensionalen Baumstrukturen, die eine Gitterebene darstellen. Dieses Schema des rekursiven Aufbaus binärer Bäume zur Darstellung mehrdimensionaler Gitter wird nun solange fortgesetzt, bis ein binärer Baum vorliegt, der die vorgegebenen  $d$  Raumdimensionen auflöst. Abbildung 27 zeigt rechts eine zweidimensionale Ebene. Dort ist der Baum in  $(x_2 \equiv y)$ -Richtung dargestellt, dessen Knoten seinerseits eindimensionale binäre Bäume in die  $(x_1 \equiv x)$ -Richtung beinhalten. Dabei ist es an dieser Stelle unerheblich, ob die aufgelisteten binären Bäume volle oder dünne Gitter darstellen. Anhand der oben beschriebenen Dimensionsrekursion für binäre Baumstrukturen in mehreren Dimensionen und dem Beispiel in zwei Raumdimensionen wird bereits ersichtlich, daß eine Implementierung dieser Datenstruktur und der auf ihr arbeitenden Operatoren, besonders in höheren Dimensionen, nicht einfach ist. Zudem werden die Dimensionen beliebig, aber fest angeordnet und sind somit untereinander ausgezeichnet.

Für theoretische Zwecke sei an dieser Stelle der eindimensionale Fall betrachtet, wie er links in Abbildung 27 angeführt ist. Sei dabei  $2^n + 1$  die Anzahl der Gitterpunkte. Damit ergibt sich die Höhe des Baumes zu  $n$ , wobei die Höhe der Wurzel des Baumes als 1 definiert ist. Die Kosten für des Auffinden, Löschen und Einfügen eines beliebigen Elements sind dabei proportional zur Höhe des Baumes.

**Lemma 6.1** *Die Operationen Einfügen, Löschen und Suchen kosten im average case als auch im worst case  $\mathcal{O}(n)$  Operationen, wenn  $n$  die Höhe des Baumes ist.*

**Beweis:** (vgl. [Meh88]). □

**Lemma 6.2** *Sei  $n_i$  mit  $1 \leq i \leq d$  die maximale Höhe eines binären Baumes von binären Bäumen zur Darstellung eines  $i$ -dimensionalen Gitters. Dann sind die Kosten für das Einfügen, Löschen und Suchen von Knoten im average case als auch im worst case  $\mathcal{O}(n = \sum_{i=1}^d n_i)$  Operationen.*

**Beweis:** Der Beweis ergibt sich sofort aus der Aussage von Satz 6.1. □

Zusammenfassend kann man für die Darstellung von Gittern durch binäre Bäume sagen, daß ausgehend von einem Baum, der  $d$  Dimensionen repräsentiert, seine Knoten und Blätter aus binären Bäumen bestehen, die  $(d - 1)$  Dimensionen darstellen. Dies zieht sich rekursiv bis zu denjenigen Bäumen durch, die eine einzelne Gitterzeile und somit einen eindimensionalen Raum aufspannen. Binäre Bäume erfüllen drei der vier zuvor angeführten Eigenschaften: fängt man an der Wurzel eines Baumes an, kann man sich einfach entlang eines bestimmten Astes des Baumes zu einem gesuchten Knoten hangeln. Folglich kann ein einzelner Knoten mit einem Aufwand von  $\mathcal{O}(n)$  gefunden werden, wenn  $n$  die maximale Tiefe des gesamten Baumes ist. Sucht man alle Knoten in beliebiger Reihenfolge, so ist dies mit einem Aufwand von  $\mathcal{O}(nN)$  möglich, jedoch nicht mit  $\mathcal{O}(N)$ , wobei  $N$  die Anzahl der Knoten im Baum ist. Zudem wird Adaptivität durch diese Datenstruktur auf einfache Art und Weise unterstützt. Der benötigte Speicherplatz pro Knoten beträgt  $2d$  Zeiger für jeweils  $d$  Söhne in positive und negative Koordinatenrichtung. Damit ist schneller Zugriff auf Sohnknoten, jedoch nicht auf Vaterknoten und nächste Nachbarn gegeben. Sollen diese Knoten ebenfalls schnell erreichbar sein, so müssen jeweils weitere  $2d$  Zeiger für Vaterknoten sowie nächste Nachbarn investiert werden. Ohne diese zusätzliche Verzeigerung der Datenstruktur ist eine Implementierung der Operatoren  $\mathbf{H}$  und  $\mathbf{E}$  oder  $\mathbf{H}_j$  und  $\mathbf{E}_j$  für  $1 \leq j \leq d$  nicht einfach. Zudem bereiten in diesem Fall die Finiten Differenzen Operatoren zusätzliche Schwierigkeiten bei ihrer Implementierung. Insgesamt sind also  $6d$  Zeiger pro Knoten für eine einfache Handhabung der hier verwendeten Operatoren sinnvoll. Somit erweisen sich binäre Bäume als eine Datenstruktur, die den meisten der gewünschten Anforderungen entspricht, jedoch vergleichsweise schwierig zu implementieren ist.

## 6.2 Hash-Tabellen

Eine Alternative zu binären Bäumen sind Hash-Tabellen, die man sich vereinfacht als ein Array  $T[\cdot]$  vorstellen kann. Eine der grundlegenden Eigenschaften von Hash-Tabellen ist der direkte Zugriff auf beliebige Daten. Eine Einführung in das Themengebiet Hashing kann man in [Knu73] und [Meh88] finden.

Zur Speicherung von Daten in einer Hash-Tabelle benötigt man einen “*key*”, der für jedes abzuspeichernde Datum eine eindeutige Identifizierung mittels der Hashfunktion  $h(\textit{key})$ , die aus dem “*key*” die Position des Eintrag in der Hash-Tabelle errechnet, erlaubt. Die Hashfunktion ist dabei mittels  $h : U \rightarrow [0, \dots, A - 1]$  mit  $A \in \mathbf{N}$  definiert, wobei  $\textit{key} \in U$  und  $A$  die Länge der Hash-Tabelle ist. Der Zugriff auf ein Element geschieht dann vermöge  $T[h(\textit{key})]$ . Im Fall von  $h(\textit{key}_1) = h(\textit{key}_2)$  hat man das Problem der Kollision zu lösen. Eine Möglichkeit, Kollisionen aufzulösen, besteht darin, eine zweite Hashfunktion anzugeben, die in der Hash-Tabelle  $T[\cdot]$  einen freien Platz sucht, falls der Platz  $T[(\textit{key})]$  bereits belegt sein sollte. Diese Methode nennt man “Open Hashing” (vgl. [Knu73, Meh88]). Ein Nachteil, der insbesondere bei der Anwendung auf adaptive Strukturen auftritt, ist die Tatsache, daß Punkte nicht nur eingefügt, sondern auch aus der Hash-Tabelle entfernt werden müssen, was bei dieser Vorgehensweise der Kollisionsauflösung einen gewissen Arbeitsaufwand verursacht. Zudem taucht bei Open Hashing das Problem des “Resizing” auf: will man mehr Elemente in einer Hash-Tabelle speichern, als Plätze vorhanden sind, so ist dies nicht möglich. Man muß in diesem Fall eine neue Hash-Tabelle anlegen und alle in ihr enthaltenen Daten neu abspeichern. Wir wählen daher die Methode des Hashing mit Verkettung (vgl. [Knu73, Meh88]). Hierbei wird eine Kollision dadurch aufgelöst, daß die Hash-Tabelle als Feld von linearen Listen  $T[h(\textit{key})]$  interpretiert wird. Dabei enthält die  $i$ -te Liste diejenigen Elemente “*key*” mit  $h(\textit{key}) = i$ . Ein Beispiel hierfür findet man in Abbildung 6.2. Will man ein Element  $X$  in die Hashtabelle einfügen, so muß man zunächst  $h(X)$  berechnen und dann in der Liste  $T[h(X)]$  das Element einfügen. Die Operationen “Löschen( $X, \{M\}$ )” und “Suche( $X, \{M\}$ )” werden dabei ähnlich behandelt. Die Zeitkomplexität besteht dabei aus der Zeit zur Auswertung der Hashfunktion  $h(X)$  und der Zeit, die lineare Liste  $T[h(X)]$  zu durchsuchen. Unter der Annahme, daß die Hashfunktion für alle  $X$  in konstanter Zeit ausgewertet wird, definieren wir analog zu [Meh88] die Kosten einer Operation mit dem *key*  $X$  als  $\mathcal{O}(1 + \sigma_h(X, N))$ , wobei  $N$  die Anzahl der abgespeicherten Elemente ist und

$$\sigma_h(X, N) = \sum_{Y \in \{M\}} \sigma_h(X, Y) \quad (6.1)$$

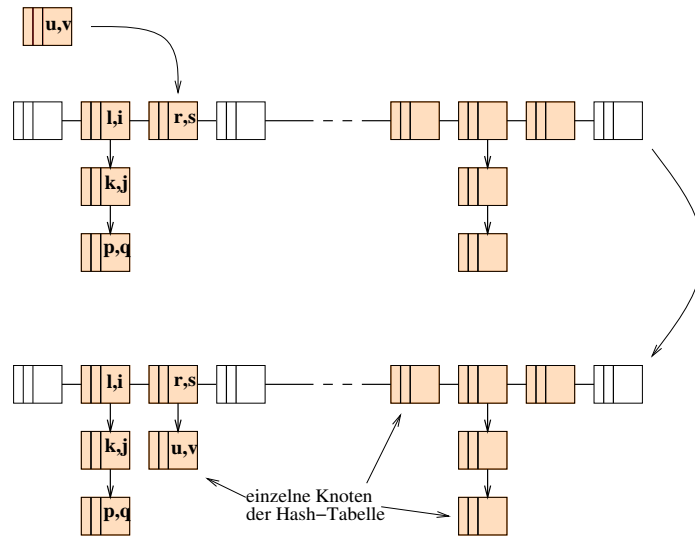


Abbildung 28: Kollisionsauflösung bei einer Hash-Tabelle mittels linearer Listen (leere Elemente entsprechen leeren Listen)

mit

$$\sigma_h(X, Y) = \begin{cases} 1, & \text{wenn } h(X) = h(Y) \wedge X \neq Y \\ 0, & \text{sonst} \end{cases} \quad (6.2)$$

gilt. Dabei stellt  $\{M\}$  die Menge der abgespeicherten Elemente dar. Eine Aussage über die Komplexität von Hashing mit Verkettung im worst case, der dann eintritt, wenn  $h(X) = i_0$  für alle  $X \in \{N\}$  ist, wird in Satz 6.1 gemacht.

**Satz 6.1** (vgl. [Meh88]) *Bei Hashing mit Verkettung ist die Zeitkomplexität der drei Operationen “Zugriff( $X, \{M\}$ )” und “Einfüge( $X, \{M\}$ )” sowie “Lösche( $X, \{M\}$ )” im schlechtesten Fall  $\mathcal{O}(N)$ , wobei  $|M| = N$  ist.*

Im mittleren Fall hingegen ist die Komplexität wesentlich besser. Dabei setzt man folgende Wahrscheinlichkeitsannahmen [Meh88] voraus:

*Annahme 1:* Die Hashfunktion  $h : U \rightarrow [0, \dots, A-1]$  verteilt das Universum der “key” gleichmäßig über das Intervall  $[0, \dots, A-1]$ , d.h. für alle  $i, i' \in [0, \dots, A-1]$  ist  $|h^{-1}(i)| = |h^{-1}(i')|$ .

*Annahme 2:* Alle Elemente von  $U$  sind als Argument einer der Operationen in der Folge gleichwahrscheinlich, d.h. das Argument der  $k$ -ten Operation ist gleich einem festen  $X \in U$  mit der Wahrscheinlichkeit  $1/|U|$ .

Diese beiden Annahmen implizieren, daß die Hashfunktion einer Gleichverteilung auf dem Intervall  $[0, \dots, A-1]$  gehorcht, d.h. daß  $h(X) = i$  mit der Wahrscheinlichkeit  $1/A$  gilt.

**Satz 6.2** (vgl. [Meh88], Satz 2, Kapitel III) *Unter den beiden obigen Annahmen hat eine Folge von  $k$  Einfüge-, Lösch- und Zugriffs-Operationen die Zeitkomplexität  $\mathcal{O}((1 + \beta/2) \cdot k)$ , wobei  $\beta = k/A$  der (maximale) Belegungsfaktor der Hash-Tabelle ist.*

**Beweis:** (vgl. [Meh88]) □

Hier stellt sich nun die Frage, ob man die gemachten Annahmen in der Realität auch erfüllen kann. Dabei erweist sich Annahme 1 als unkritisch, da man mittels der Divisionsmethode  $h(X) = \tilde{h}(X) \bmod k$  setzen kann, wenn  $k|A$  gilt,  $k$  also  $A$  restfrei teilt. Falls  $k|A$  nicht erfüllt ist, aber  $A \gg k$  gilt, so erfüllt die Divisionsmethode Annahme 1 fast und Satz 6.2 ist nach wie vor gültig. Annahme 2 hingegen ist problematisch, da man a priori im allgemeinen keine Informationen darüber besitzt, welche Elemente in einer Hash-Tabelle tatsächlich abgespeichert werden sollen. Übertragen auf den Fall der dünnen Gitter bedeutet dies beispielsweise, daß man im allgemeinen nicht weiß, wo ein selbstadaptiver Algorithmus Gitterpunkte einfügt und wo sich Bereiche des Grundgebiets befinden, in denen keine oder nur wenige Gitterpunkte benötigt werden. Somit ist Satz 6.2 diesbezüglich mit Vorsicht zu genießen<sup>12</sup>. Das Ziel muß deshalb eine Hashfunktion sein, die unabhängig von der zugrundegelegten Punktmenge der Gitterpunkte eines dünnen Gitters die Punkte möglichst gleichmäßig über die Hash-Tabelle streut. Ein Beispiel für eine Hashfunktion, die diese Forderung relativ gut zu erfüllen scheint, wird in (6.4) später in diesem Abschnitt zusammen mit einigen Anwendungen angegeben.

Der Vollständigkeit halber seien nun noch die mittleren Kosten des worst case von Hashing mit Verkettung angeführt. Dabei folgen wir wiederum den Ausführungen von [Meh88].

**Satz 6.3** (vgl. [Meh88], Satz 3, Kapitel III) *Sei  $V$  eine beliebige Teilmenge von  $U$  mit  $V = \{X_1, \dots, X_k\}$  und  $|V| = k$ . Ferner gelte die Wahrscheinlichkeit  $\text{prob}(h(X_j) = i) = 1/A$  für alle  $1 \leq j \leq k$  und  $i \in [0, \dots, A - 1]$ . Dann sind die mittleren Kosten des schlechtesten Falls von Hashing mit Verkettung  $\mathcal{O}(\log(k)/\log(\log(k)))$ , wenn  $\beta = k/A \leq 1$ .*

**Beweis:** (vgl. [Meh88]) Dabei ist die Bedingung  $\text{prob}(h(X_j) = i) = 1/A$  vergleichbar mit der vorigen Annahme 2. □

In dem Fall von Gittern, die von einer hierarchischen Tensorproduktbasis herrühren, wie es beispielsweise bei dünnen Gittern der Fall ist, ist die Angabe des Levels und eines Index innerhalb des Levels eine Möglichkeit als

<sup>12</sup>Durch den hierarchischen Ansatz haben wir es insbesondere mit bedingten Wahrscheinlichkeiten zu tun. Trotzdem ist die Zeitkomplexität des Satzes 6.2 auch in unserem Fall erreichbar.

Wahl eines “key” für jeden einzelnen Gitterpunkt des betrachteten Gitters. Eine weitere Möglichkeit ist die Speicherung der Koordinaten, die ebenfalls für jeden Punkt eindeutig sind. Damit muß man im ersten Fall  $2d$  Integer-Variablen speichern, im zweiten  $d$  REAL-Variablen<sup>13</sup>. Folglich benötigt man zur Speicherung der Information, die nötig sind, einen beliebigen Gitterpunkt zu erreichen, weniger Bytes pro Knoten als im Fall des Binärbaums. Da Knoten eines (adaptiven oder adaptiven dünnen) Gitters in Abhängigkeit der zugrundegelegten Hashfunktion “irgendwo” in einer Hash-Tabelle abgelegt werden und für ihre Speicherung in einer Hash-Tabelle keinen Bezug zueinander brauchen, wird Adaptivität durch diese Datenstruktur ebenfalls unterstützt. Dabei benötigt man in  $d$  Dimensionen mindestens  $\mathcal{O}(d)$  Operationen, falls die Hashfunktion die abzulegenden Daten vernünftig streuen soll. Für ein beliebiges, aber festes  $d$  hat man folglich  $\mathcal{O}(1)$  Operationen zur Auswertung der Hash-Funktion. Allerdings benötigt man für Hash-Tabellen einen gewissen Speicheroverhead, da der Belegungsgrad  $\beta = N/A$  einer Hash-Tabelle aus Effizienzgründen 70% – 75% nicht überschreiten sollte ([Knu73, Meh88]). Setzt man voraus, daß nur reguläre dünne Gitter verwendet werden, so kann man als Hashfunktionen die Funktion

$$h(L, \mathbf{l}, \mathbf{i}) = \begin{cases} L \times \mathbf{l} \times \mathbf{i} & \rightarrow \mathbf{N} \\ (L, \mathbf{l}, \mathbf{i}) & \mapsto h(L, \mathbf{l}, \mathbf{i}) \end{cases} \quad (6.3)$$

benutzen, wobei  $h(L, \mathbf{l}, \mathbf{i})$  einem Aufzählungsalgorithmus der Punkte eines dünnen Gitters entspricht. Dabei wird garantiert, daß keine Kollision auftreten kann, wenn nur Gitterpunkte eines dünnen Gitters vom Level  $L$  abgespeichert werden und die Länge der Hash-Tabelle größer ist als die Zahl der abzuspeichernden Dünngitter-Punkte. Sind diese Voraussetzungen erfüllt, so haben die Operationen Suchen, Einfügen und Löschen stets einen Aufwand von  $\mathcal{O}(1)$ , unabhängig davon, wie groß der Belegungsgrad der Hash-Tabelle ist. Durch die Wahl eines geeignet hohen  $L$  wird im allgemeinen eine recht gute Streuung der Gitterpunkte erreicht. Zwei Beispiele für die Anwendung dieser Hashfunktion finden sich in Abbildung 6.2. Hierin bezeichnet die  $x$ -Achse die Länge von linearen Listen, die  $y$ -Achse gibt an, wieviele Listen einer gegebenen Länge in der Hash-Tabelle existieren. In beiden Graphiken werden hierbei reguläre dünne Gitter vom Level 12 mit insgesamt 61441 Gitterpunkten verwendet, wobei die Länge  $A$  der Hash-Tabelle in der linken Graphik  $A = 70000$  und in der rechten Graphik lediglich  $A = 3000$  ist. Im ersten Fall erkennt man, daß keine Kollision auftritt (Belegungsgrad der Hash-Tabelle beträgt 87.78%), im zweiten Fall, in dem die durchschnittliche Listenlänge unter Voraussetzung einer gleichverteilenden Hashfunktion  $61441/3000 \approx 20.48$  betragen sollte, haben wir lediglich zwei verschiedene Listenlängen, nämlich 20 und 21. Somit ist die Streuung in diesem Fall annähernd optimal. Auch bei adaptiven dünnen Gittern, in denen keine star-

<sup>13</sup>REAL  $\in$  {float, double}

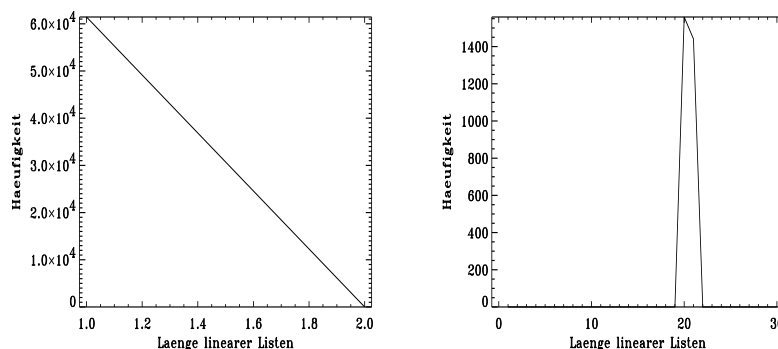


Abbildung 29: Länge von linearen Listen und deren Häufigkeit (links: Länge der Hash-Tabelle  $A = 70000$ , rechts  $A = 3000$ , Anzahl gespeicherter Punkte jeweils  $61441$ )

ken Singularitäten auftauchen, ist obige Hash-Funktion noch gut geeignet, wie Abbildung 6.2 zeigt. Allerdings treten beispielsweise bei Punktsingula-

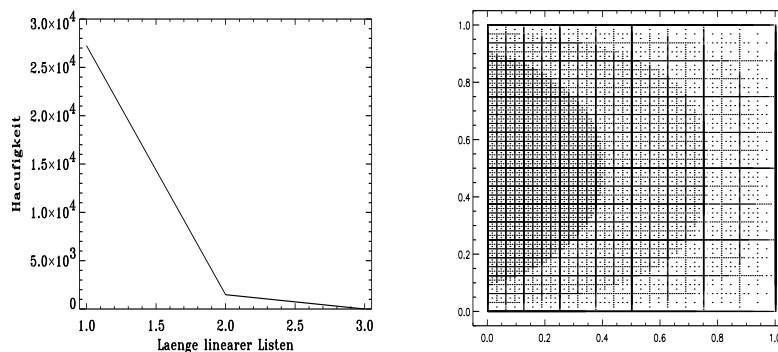


Abbildung 30: Länge von linearen Listen und deren Häufigkeit für das angegebene Gitter,  $30210$  Punkte, Länge der Hash-Tabelle  $A = 54067$ , Belegungsfaktor  $\beta = 55.88\%$

ritäten für diese Hash-Funktion Kollisionen derart auf, daß die Listenlängen stark wachsen. In diesem Fall geht der Vorteil des direkten Zugriffs verloren, da in einem Suchbefehl jeweils eine lineare Liste durchsucht werden muß. Abhilfe schafft hier etwa die Hashfunktion

$$h(\mathbf{l}, \mathbf{i}) = \sum_{j=1}^d \left( 2^{l_j} + i_j \right) \cdot \text{prim}(j) \cdot \text{prim}(43 + (d-2) \cdot 10 - j) \quad (6.4)$$

für die Dimensionen  $d = 2, 3$ , wobei  $\text{prim}(j)$  die  $j$ -te Primzahl bedeutet. Hier werden sowohl reguläre als auch adaptive dünne Gitter relativ gut gestreut.

Dabei ist zu beobachten, daß die “Optimalität” im Fall regulärer dünner Gitter nicht erhalten werden kann, jedoch im Fall adaptiver dünner Gitter auch bei starken Singularitäten die Listenlängen sehr kurz bleiben. Eine mathematische Begründung, warum sich diese Hashfunktion als geeignet erweist, muß an dieser Stelle entfallen. Zwei Beispiele hierfür finden sich in den Abbildungen 31 und 32, wo zum einen ein reguläres und zum anderen ein adaptives dünnes Gitter betrachtet wird.

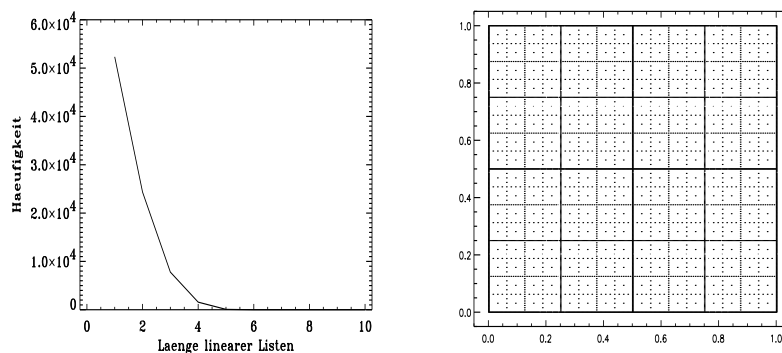


Abbildung 31: Länge von linearen Listen und deren Häufigkeit für ein reguläres dünnes Gitter, Level 13, 131.073 Punkte, Länge der Hash-Tabelle  $A = 187.000$ , Belegungsfaktor  $\beta = 70.09\%$

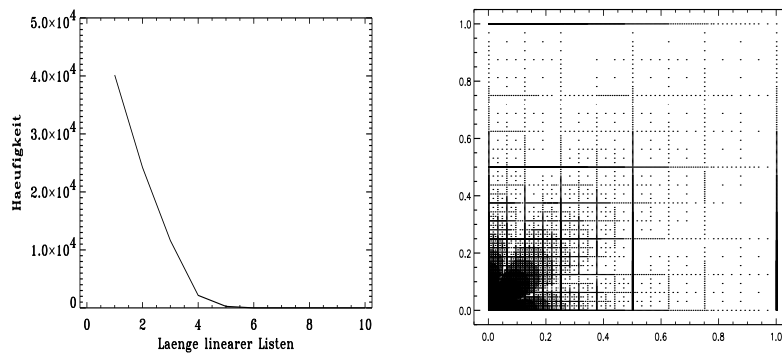


Abbildung 32: Länge von linearen Listen und deren Häufigkeit für ein adaptives Gitter, maximale Leveltiefe 10, 133.412 Punkte, Länge der Hash-Tabelle  $A = 187.000$ , Belegungsfaktor  $\beta = 71.34\%$

Vermöge einfacher Integerrechnung sind aus gegebenen Level- und Index-Informationen eines beliebigen Knotens sowohl seine hierarchischen Söhne als auch seine hierarchischen Väter einfach zu ermitteln. Zudem sind ebenfalls durch Integerrechnung die direkten Nachbarn in jede Koordinatenrichtung einfach zu ermitteln. Da zudem die Implementierung einer Hash-Tabelle relativ einfach ist, erfüllt die Datenstruktur Hash-Tabelle alle vier zuvor



gestellten Bedingungen. Damit erweist sich die Hash-Tabelle für die hier gestellten Anforderungen als die sinnvollere Datenstruktur.

### 6.3 Verzeigerte Hash-Tabellen

Im vorangegangenen Abschnitt 6.2 haben wir festgestellt, daß Hash-Tabellen für unsere Anwendungen geeigneter sind als binäre Bäume. Führen wir bei den Hash-Tabellen zusätzliche Verpointerungen zwischen bestimmten Knoten ein (etwa Pointer eines Knotens auf seine Sohnknoten oder auf seine nächsten Nachbarknoten), so kann man den Zugriff auf einen Knoten (z.B. bei einem hierarchischen Durchlauf durch alle Knoten) erheblich beschleunigen, da anstatt der Berechnung von Level- und Index-Information des gesuchten Knotens sowie der Auswertung der Hashfunktion zur Ermittlung des Knotens in der Hash-Tabelle lediglich auf einen Zeiger, der auf das gesuchte Objekt direkt verweist, zugegriffen wird. Man beachte, daß die Berechnung der Hashfunktion mindestens  $\mathcal{O}(d)$  Operationen benötigt, wenn sie die Einträge in der Hash-Tabelle ausreichend gut verteilen soll und somit jede beteiligte Koordinatenrichtung berücksichtigen muß. Zudem benötigt die Berechnung der gewünschten Level- und Index-Informationen im worst case  $\mathcal{O}(n)$  Operationen<sup>14</sup>, wobei  $n$  die maximale Leveltiefe eines Gitterpunkts ist. Der zusätzliche Speicheraufwand beläuft sich dabei auf jeweils  $2d$  Zeiger für hierarchische Söhne und Väter sowie nächste Nachbarn. Will man alle Zeiger bereitstellen, so sind folglich  $6d$  Zeiger notwendig. Mit diesem Ansatz erhält man somit eine Vermischung der Datenstrukturen Hash-Tabelle und binärer Baum.

Der Schritt des Einfügens von Zeigern kann während oder direkt nach dem Aufbau eines dünnen Gitters erfolgen. Bei adaptiver Verfeinerung oder Vergrößerung können die notwendigen Pointer sofort gesetzt bzw. bereits bestehende umgesetzt werden, was schnell und einfach durchzuführen ist. Durch diese Erweiterung der Datenstruktur kann eine bedeutende Erhöhung der Ausführungsgeschwindigkeit eines Codes erreicht werden. Zudem erlaubt die Einführung von Pointern eine programmiertechnisch einfache Handhabung von Mehrgitteralgorithmen oder auch Erzeugendensystemen, wie sie in [Gri90] verwendet werden.

---

<sup>14</sup>Ein hierarchischer Vater in eine beliebige, aber feste Koordinatenrichtung besitzt einen um 1 verminderten Level und kann daher sehr einfach ermittelt werden. Das Problem liegt im zweiten Vater: abhängig von den Level- und Index-Informationen des betrachteten Knotens und der gewählten Koordinatenrichtung kann der Vater jeden beliebigen kleineren Level besitzen, womit sich der Aufwand zur Berechnung des zweiten Vaters auf  $\mathcal{O}(n)$  Operationen belaufen kann. Ein ähnliches Problem bzgl. des Aufwands ergibt sich bei der Berechnung der nächsten Nachbarn. An dieser Stelle sei angemerkt, daß binäre Bäume aufwandstechnisch nicht besser sind, aber – wie im Abschnitt 6.1 bemerkt – schwieriger zu implementieren.

## 6.4 Kosten für die Anwendung Finiter Differenzen Operatoren auf dünnen Gittern

Wir wollen die Kosten für die Anwendung eines elliptischen Differentialoperators 2. Ordnung exemplarisch am Beispiel des Laplace-Operators studieren. Für allgemeine elliptische Differentialoperatoren 2. Ordnung kann daraus sofort der notwendige Aufwand für eine Anwendung des diskreten Operators hergeleitet werden.

Betrachten wir zunächst Operatoren in nodaler Schreibweise: um die zweite Ableitung  $\mathbf{D}_{jj}^S$  in Richtung  $j$  mit  $1 \leq j \leq d$  zu berechnen, muß zunächst eine  $(d-1)$ -dimensionale Hierarchisierung  $\mathbf{H}_j$  durchgeführt werden, anschließend wird der eindimensionale Finite Differenzenstern angewendet, und der Operator wird abgeschlossen mit der Anwendung einer  $(d-1)$ -dimensionalen Enthierarchisierung  $\mathbf{E}_j$ . Berechnet man die auftretenden Hierarchisierungssterne und Enthierarchisierungssterne gemäß Formel (2.12), so ist jeweils ein Stern mit  $3^{d-1}$  Einträgen anzuwenden. Im Falle regulärer dünner Gitter, bei denen lokal eindimensional bei der Anwendung des eindimensionalen Finite Differenzensterns stets äquidistante Gitterstrukturen auftreten, müssen die Gewichte der Finite Differenzensterne nicht berechnet werden. Somit erfordert die Anwendung des Operators  $\mathbf{D}_{jj}$  insgesamt 4 Operationen (jeweils eine Addition, Multiplikation, Subtraktion und Division). Damit ergibt sich, falls  $N$  die Anzahl der Gitterpunkte eines dünnen Gitters bezeichnet, ein Gesamtaufwand von  $(4 \cdot 3^{d-1} + 2) \cdot N$ , also  $\mathcal{O}(3^d \cdot N)$ , Operationen, womit der Aufwand exponentiell von der Dimension abhängt.

Unter Ausnutzung der bereits im Abschnitt 4.2 angesprochenen Tensorproduktstruktur (nachfolgend TP abgekürzt) der Finiten Differenzen Operatoren, wodurch ein  $d$ -dimensionaler Operator durch Konkatenation von  $d$  eindimensionalen Operatoren zusammengesetzt werden kann, verringert sich der Aufwand erheblich. Implementiert man  $d$ -dimensionale Operatoren als  $d$ -fache Anwendung eindimensionaler Operatoren, so reduziert sich der Aufwand für die Operatoren  $\mathbf{H}_i$  und  $\mathbf{E}_i$  auf nur mehr  $3 \cdot (d-1)$  Operationen, womit sich der Gesamtaufwand einer zweiten Ableitung mittels Finiter Differenzen auf dünnen Gittern durch Ausnutzung der angesprochenen Tensorproduktstruktur auf  $(6 \cdot (d-1) + 4) \cdot N$  Operationen (Ordnung  $\mathcal{O}(d \cdot N)$ ) verringert und somit nur noch linear von der Dimension abhängt. Es sei an dieser Stelle angemerkt, daß unabhängig davon, ob die Tensorproduktstruktur in der Implementierung ausgenutzt wird oder nicht, der zu erwartende Aufwand stets linear in der Anzahl der Gitterpunkte ist, also stets von der Ordnung  $\mathcal{O}(N)$ . Der gravierende Unterschied liegt jedoch in der jeweiligen Konstante verborgen. Dies zeigt auch nachfolgende Tabelle 8, in der die Konstanten für beide vorgestellten Vorgehensweise in Abhängigkeit der Dimension für  $1 \leq d \leq 9$  aufgetragen sind.

Tabelle 8: Abhängigkeit der Kosten von der Wahl der Darstellung des Operators (ohne TP:  $d$ -dimensionale Darstellung; mit TP: Darstellung vermöge eindimensionaler Operatoren)

Dimension	1	2	3	4	5	6	7	8	9
ohne TP	4	10	22	58	166	490	1462	4378	13126
mit TP	4	10	16	22	28	34	40	46	52

Um den Laplace-Operator anwenden zu können, müssen für alle  $d \geq 2$  Raumdimension die entsprechenden zweiten Ableitungen berechnet werden, wobei durch die Addition der  $d$  zweiten Ableitungen zusätzlich  $d - 1$  Additionen pro Knoten anfallen. Der Gesamtaufwand bei Ausnutzung der Tensorproduktstruktur für den Laplace-Operator beläuft sich demnach auf insgesamt

$$\left( d \cdot (6 \cdot (d - 1) + 4) + (d - 1) \right) \cdot N = (6 \cdot d^2 - d - 1) \cdot N$$

Operationen und hängt damit quadratisch von der Dimension ab. Der Faktor 6 vor  $d^2$  läßt sich auf Kosten eines erhöhten Speicherverbrauchs noch auf  $\frac{9}{2}$  senken, die Abhängigkeit von  $d^2$  jedoch bleibt.

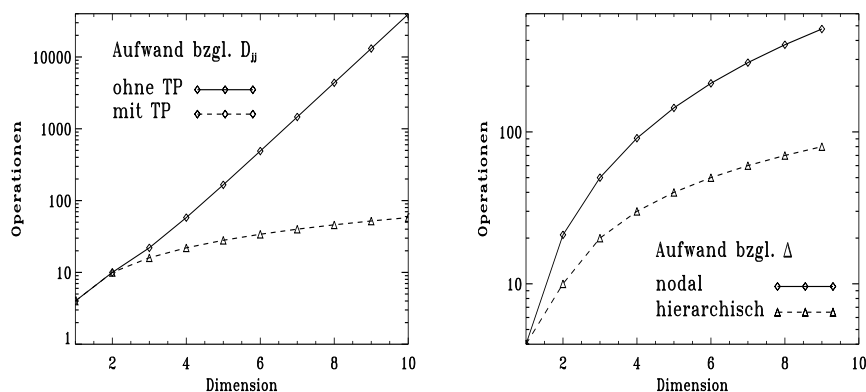


Abbildung 33: Aufwandsvergleich bei zweiter Ableitung für  $d$ -dimensionalen Operator und  $d$ -mal eindimensionalen Operator (links) sowie beim Laplace-Operator bei Verwendung des nodalen bzw. des hierarchischen Operators (rechts), unterschiedliche Skalierung der Ordinate

Betrachtet man die Operatoren in der hierarchischen Notation, so ist die Abhängigkeit des Aufwands nur noch linear in der Dimension: hier bestehen die einzelnen Operatoren aus einer eindimensionalen Enthierarchisierung, einem eindimensionalen Finite Differenzen Stern und einer abschließenden eindimensionalen Hierarchisierung. Somit ergibt sich ein Gesamtaufwand

von  $10 \cdot N$  Operationen für eine Ableitung bezüglich einer Raumdimension, für den Laplace-Operator folglich  $10 \cdot d \cdot N$  Operationen. Um in einer Implementierung die Anwendung der Diskretisierungsmatrizen möglich billig zu realisieren, sollten die Operatoren folglich in hierarchischer Notation implementiert sein.

In Abbildung 33 ist der Unterschied im Aufwand bestimmter Operatoren bei den angesprochenen Vorgehensweisen angegeben. In der linken Graphik sind die Ergebnisse aus Tabelle 8 aufgearbeitet. Im rechten Bild erkennt man, daß die Implementierung der Operatoren in hierarchischer Notation erheblich weniger Aufwand in der Berechnung verursacht.

## 6.5 Parallelisierung

### 6.5.1 Parallelisierungskonzept

Die in Abschnitt 4.8.1 angesprochene Technik der Gebietszerlegung führt auf natürliche Art und Weise auf ein paralleles Vorgehen bei der Berechnung einer Problemstellung. Darüberhinaus bietet sich die Möglichkeit, jedes einzelne (Teil-)Problem von mehr als nur einem Prozessor berechnen zu lassen. Eine sehr einfache Vorgehensweise hierzu bietet auf SMP-Multiprozessor-Maschinen "OpenMP API (application programming interface)". OpenMP stellt (vereinfacht gesprochen) dabei eine Reihe von Compiler-Direktiven, Laufzeit-Bibliotheken und Umgebungsvariablen zur Verfügung, mit denen parallele Anwendungen beschrieben und kontrolliert werden können. Detaillierte Informationen über OpenMP API [OMP] können im WWW unter

<https://toolbox.sgi.com/toolbox/>

und den dort weiter-verweisenden Links sowie im "C Language Reference Manual" (Kapitel 11: *Multiprocessing C/C++ Compiler Directives*) [SG] von Silicon Graphics gefunden werden.

Zur Parallelisierung eines Codes muß der Benutzer bei Verwendung von OpenMP (im Allgemeinen) lediglich an geeigneten Stellen innerhalb des Codes vordefinierte `#pragma`-Anweisungen einfügen sowie die Umgebungsvariable, die die Anzahl der zu verwendenden Prozessoren steuert, setzen. Die Parallelisierung der C-Schleife

```
for(i=0;i<ENDE;i++)
  { /* Schleifenanweisung */ }
```

geschieht beispielsweise vermöge der Anweisungen

```
#pragma parallel
#pragma shared ( /* shared variables */ )
```

```
#pragma local ( /* local variables */ )
#pragma pfor iterate(i=0;ENDE;1)
  for(i=0;i<ENDE;i++)
    { /* Schleifenanweisung */ }
```

Somit ist klar, daß die Parallelisierung einer Schleife schnell und einfach durchführbar ist, zumindest falls die Abarbeitung des  $k$ -ten Elements nicht die Abarbeitung eines vorangegangenen Elements benötigt.

### 6.5.2 Parallelisierung einzelner Operationen

Nachfolgend finden sich exemplarisch Ergebnisse für die Parallelisierung der skalaren Multiplikation auf dünnen Gittern, der Berechnung des Sterns  $\mathbf{D}_{00}$  (Teiloperator der zweiten Ableitung  $\mathbf{D}_{00}^S$  mittels Finiter Differenzen auf dünnen Gittern) sowie der Anwendung des iterativen Lösers BiCG-STAB. Abschließend vergleichen wir den Lösungsprozess auf 1, 2, 3 und 4 Prozessoren. Alle numerischen Ergebnisse in diesem Abschnitt wurden berechnet auf einer O200 mit 4 MIPS R10000 (180MHz) Prozessoren, IP32 mit IRIX 6.4. Bei den Ergebnissen ist zu berücksichtigen, daß beispielsweise beim Vergleich von einem zu zwei Prozessoren kein Faktor von exakt zwei in der Ausführungszeit zu erwarten ist, da beispielsweise die Thread-Verwaltung sowie eventuell notwendige Kommunikation die Ausführungszeiten beim Programmablauf von mehr als einem Prozessor beeinflussen. In den Tabellen wurden folgende Abkürzungen verwendet:

- $N$ : Anzahl der Gitterpunkte des verwendeten dünnen Gitters
- $q_N$ : Quotient der Gitterpunkt zweier aufeinanderfolgender Level
- $Z_{SM}^P$ : Zeit für 2000 skalare Multiplikationen auf dem angegebenen Level für  $P$  Prozessoren
- $q_{SM}^P$ : Quotient zweier aufeinanderfolgender Zeiten  $Z_{SM}^P$ , berechnet mittels

$$q_{SM}^P(l) = \frac{Z_{SM}^P(l-1)}{Z_{SM}^P(l)}$$

- $q_{N,SM}^P$ : dieser Quotient wird berechnet vermöge

$$q_{N,SM}^P(l) = \frac{q_{SM}^P(l)}{q_N(l)}$$

- $Z_{D_{00}}^P$ : Zeit für 500 Operationen  $\mathbf{D}_{00}$  auf dem angegebenen Level für  $P$  Prozessoren

- $q_{D_{00}}^P$ : Quotient zweier aufeinanderfolgender Zeiten  $Z_{D_{00}}^P$ , berechnet mittels

$$q_{D_{00}}^P(l) = \frac{Z_{D_{00}}^P(l-1)}{Z_{D_{00}}^P(l)}$$

- $q_{N,D_{00}}^P$ : dieser Quotient wird berechnet vermöge

$$q_{N,D_{00}}^P(l) = \frac{q_{D_{00}}^P(l)}{q_N(l)}$$

Dabei geben die beiden Quotienten  $q_{N,SM}^P$  und  $q_{N,D_{00}}^P$  an, wie sich die Ausführungszeit im Verhältnis zur erhöhten Anzahl an Gitterpunkten verhält. Wären diese Faktoren identisch 1, so würden die Ausführungszeiten linear mit der Anzahl der Punkte steigen. Wie man aus den Tabellen 9 bis 11 erkennen kann, liegen diese Faktoren stets unter der Größe 1.5, teilweise sind sie kleiner als 1. Dies zeigt wiederum die Abhängigkeit der Testläufe von der Auslastung des Rechners zum Zeitpunkt des Testens. Zugleich zeigen diese Zahlen jedoch, daß die Ausführungszeiten auch im Falle mehrerer Prozessoren in etwa mit der Anzahl der Gitterpunkte korrelieren.

Da auf kleinen Leveln die Ausführungszeiten verschwindend klein sind und ein Vergleich paralleler Algorithmen erst sinnvoll ist, wenn jeder einzelne Prozessor "genügend" Anweisungen durchzuführen hat, sind in den Tabellen 9 bis 11 lediglich die Ausführungszeiten ab Level 8 aufgenommen. An dieser Stelle muß auch der Hinweis gegeben werden, daß die angegebenen Zeiten bei jedem erneuten Testlauf anders ausfallen werden. In Abhängigkeit der Auslastung eines Rechners können hierbei geringfügige Abweichungen der hier angegebenen Ausführungszeiten auftreten, bei höheren Leveln ( $l = 13, 14, 15, \dots$ ) sind Abweichungen im Sekundenbereich möglich. Somit sind alle hier angegebenen Zahlenwerte lediglich als Richtwerte anzusehen, die das zu erwartende Verhalten tendenziell aufzeigen sollen.

Tabelle 9: Ausführungszeiten der parallelen skalaren Multiplikation und des Teiloperators  $D_{00}$  in Sekunden,  $P = 1$  Prozessor

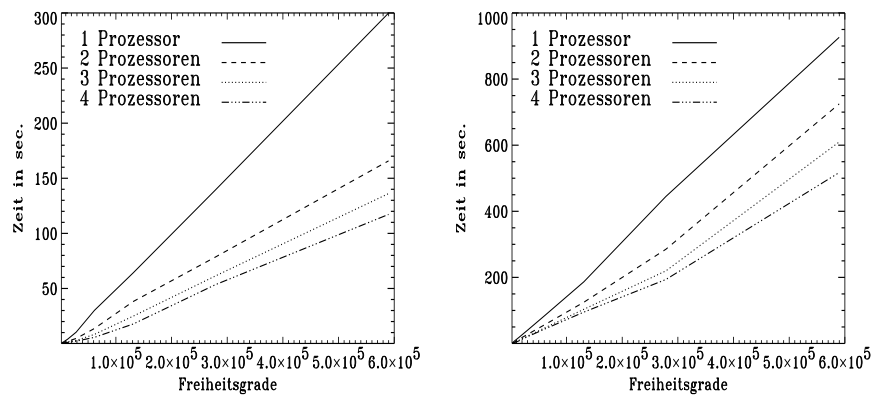
$l$	$N$	$q_N$	$Z_{SM}^1$	$q_{SM}^1$	$q_{N,SM}^P$	$Z_{D_{00}}^1$	$q_{D_{00}}^1$	$q_{N,D_{00}}^P$
8	2817	2.20	0.69	–	–	2.13	–	–
9	6145	2.18	1.58	2.29	1.05	7.19	3.38	1.55
10	13313	2.17	4.00	2.53	1.17	16.83	2.34	1.08
11	28673	2.15	10.35	2.59	1.20	38.40	2.28	1.06
12	61441	2.14	30.11	2.91	1.36	85.36	2.22	1.04
13	131073	2.13	64.16	2.13	1.00	186.60	2.19	1.03
14	278529	2.12	139.24	2.17	1.02	444.93	2.38	1.12
15	589825	2.12	299.89	2.15	1.02	926.33	2.08	0.98

Tabelle 10: Ausführungszeiten der parallelen skalaren Multiplikation und des Teiloperators  $\mathbf{D}_{00}$  in Sekunden,  $p = 2$  Prozessoren

$l$	$N$	$q_N$	$Z_{SM}^2$	$q_{SM}^2$	$q_{N,SM}^P$	$Z_{D_{00}}^2$	$q_{D_{00}}^2$	$q_{N,D_{00}}^P$
8	2817	2.20	0.35	–	–	1.41	–	–
9	6145	2.18	0.87	2.49	1.14	3.78	2.68	1.23
10	13313	2.17	2.15	2.47	1.14	11.80	3.12	1.44
11	28673	2.15	4.62	2.15	1.00	26.36	2.23	1.04
12	61441	2.14	13.67	2.96	1.38	57.39	2.18	1.02
13	131073	2.13	38.14	2.79	1.31	124.85	2.18	1.02
14	278529	2.12	78.45	2.06	0.97	285.20	2.28	1.08
15	589825	2.12	165.76	2.11	1.00	724.89	2.54	1.20

Tabelle 11: Ausführungszeiten der parallelen skalaren Multiplikation und des Teiloperators  $\mathbf{D}_{00}$  in Sekunden,  $p = 4$  Prozessoren

$l$	$N$	$q_N$	$Z_{SM}^4$	$q_{SM}^4$	$q_{N,SM}^P$	$Z_{D_{00}}^4$	$q_{D_{00}}^4$	$q_{N,D_{00}}^P$
8	2817	2.20	0.27	–	–	1.26	–	–
9	6145	2.18	0.46	1.70	0.78	2.54	2.02	0.92
10	13313	2.17	1.51	3.28	1.51	7.13	2.81	1.29
11	28673	2.15	2.19	1.45	0.67	19.68	2.76	1.28
12	61441	2.14	5.64	2.58	1.20	44.20	2.25	1.05
13	131073	2.13	17.73	3.14	1.48	94.93	2.15	1.01
14	278529	2.12	53.38	3.01	1.42	193.70	2.04	0.96
15	589825	2.12	117.21	2.20	1.04	517.11	2.67	1.26

Abbildung 34: Ausführungszeiten für skalare Multiplikation (links) und Finite Differenzen Stern  $\mathbf{D}_{00}$  (rechts) für 1, 2, 3 und 4 Prozessoren

**Definition 6.1** Unter der Effizienz  $\text{Eff}_P$  eines parallelen Programms, das auf  $P$  Prozessoren parallel arbeitet, versteht man den Ausdruck

$$\text{Eff}_P = \frac{\text{Ausführungszeit für einen Prozessor}}{P \cdot \text{Ausführungszeit für } P \text{ Prozessoren}}$$

**Definition 6.2** Unter dem Speed-Up  $\text{SU}_P$  eines parallelen Programms, das auf  $P$  Prozessoren parallel arbeitet, verstehen wir den Ausdruck

$$\text{SU}_P = \frac{\text{Ausführungszeit für einen Prozessor}}{\text{Ausführungszeit für } P \text{ Prozessoren}}$$

Damit ergibt sich der Zusammenhang  $\text{SU}_P = P \cdot \text{Eff}_P$ .

Tabelle 12: Ergebnisse der Tabellen 9 bis 11, aufbereitet für die Berechnung der Effizienz

$l$	$Z_{SM}^1$	$Z_{SM}^2$	$Z_{SM}^4$	$Z_{D_{00}}^1$	$Z_{D_{00}}^2$	$Z_{D_{00}}^4$
11	10.35	4.62	2.19	38.40	26.36	19.68
13	64.16	38.14	17.73	186.60	124.85	94.93
15	299.89	165.76	117.21	926.33	724.89	517.11

In Tabelle 12 sind Ergebnisse aus den Tabellen 9 bis 11 aufbereitet, um die Berechnung der Effizienz einfacher nachvollziehbar zu machen. Nach Definition 6.1 und Tabelle 12 ergibt sich in Abhängigkeit von der Problemgröße die in Tabelle 13 aufgelisteten Effizienzen  $\text{Eff}_P(l)$ . Dabei erkennt man, daß in der genannten Tabelle Effizienzen vorkommen, die größer als 1 sind. Diese Ergebnisse sind im wesentlichen auf Cache- und Speicher-Effekte zurückzuführen. Dieselbe Aussage gilt wegen Definition 6.2 ebenso für die Werte in Tabelle 14, in der Speed-Up's angegeben werden.

Tabelle 13: Effizienz  $\text{Eff}_P(l)$

$l$	$SM$		$D_{00}$	
	$P = 2$	$P = 4$	$P = 2$	$P = 4$
11	1.12	1.18	0.73	0.49
13	0.84	0.90	0.75	0.49
15	0.90	0.64	0.64	0.45

Nach Definition 6.2 läßt sich der Speed-Up für verschiedene parallele Algorithmen sofort aus Tabelle 13 ermitteln. Die ermittelten Zahlen für den Speed-Up finden sich in Tabelle 14.

### 6.5.3 Parallelisierung des Löser

Der Löser (vorkonditioniertes BiCGSTAB-Verfahren, das nachfolgend prec-BiCGSTAB abgekürzt wird) wurde – abgesehen von der Enthierarchie-



Tabelle 14: Speed-Up  $SU_P(l)$ 

$l$	$SM$		$D_{00}$	
	$P = 2$	$P = 4$	$P = 2$	$P = 4$
11	2.24	4.73	1.46	1.95
13	1.68	3.62	1.49	1.97
15	1.81	2.56	1.28	1.79

rung, die im Operator angewendet werden muß – komplett nach dem beschriebenen Verfahren parallelisiert. Die Ausführungszeiten für  $P \in \{1, 2, 3, 4\}$  Prozessoren finden sich in den Tabellen 15 bis 16. Gelöst wurde dabei das Laplace–Problem L1, das in dem Abschnitt 7.2.1 ausführlich beschrieben wird. In den nachfolgenden Tabellen benutzen wir zusätzlich zu den bereits vereinbarten Abkürzungen die Notationen

- $Z_{OP}^P$ : Zeit für die Lösung des Laplace–Problems L1 auf dem angegebenen Level für  $P$  Prozessoren
- $q_{OP}^P$ : Quotient zweier aufeinanderfolgender Zeiten  $Z_{OP}^P$ , berechnet mittels

$$q_{OP}^P(l) = \frac{Z_{OP}^P(l-1)}{Z_{OP}^P(l)}$$

- $q_{N,OP}^P$ : dieser Quotient wird berechnet vermöge

$$q_{N,OP}^P(l) = \frac{q_{OP}^P(l)}{q_N(l)}$$

Wie zuvor gibt der Quotient  $q_{N,OP}^P$  an, inwieweit sich die Ausführungszeit im Verhältnis zur erhöhten Anzahl an Gitterpunkten verhält.

Tabelle 15: Ausführungszeiten für das Lösen der Problemstellung L1 mittels precBiCGSTAB in Sekunden, 1 und 2 Prozessoren

$l$	$N$	$q_N$	$Z_{OP}^1$	$q_{OP}^1$	$q_{N,OP}^1$	$Z_{OP}^2$	$q_{OP}^2$	$q_{N,OP}^2$
6	577	2.25	1.88	–	–	1.80	–	–
7	1281	2.22	5.49	2.92	1.32	4.81	2.67	1.20
8	2817	2.20	19.78	3.60	1.64	16.98	3.53	1.60
9	6145	2.18	70.83	3.58	1.64	52.08	3.07	1.41
10	13313	2.17	232.83	3.29	1.52	179.77	3.45	1.59
11	28673	2.15	920.34	3.95	1.84	699.52	3.90	1.81
12	61441	2.14	4949.06	5.38	2.51	3458.33	4.94	2.31
13	131073	2.13	20325.54	4.11	1.93	15252.58	4.41	2.07
14	278529	2.12	87010.20	4.18	2.02	63690.67	4.18	1.97

Tabelle 16: Ausführungszeiten für das Lösen der Problemstellung L1 mittels precBiCGSTAB in Sekunden, 3 und 4 Prozessoren

$l$	$N$	$q_N$	$Z_{OP}^3$	$q_{OP}^3$	$q_{N,OP}^3$	$Z_{OP}^4$	$q_{OP}^4$	$q_{N,OP}^4$
6	577	2.25	1.54	–	–	1.62	–	–
7	1281	2.22	4.64	3.01	1.36	4.39	2.71	1.22
8	2817	2.20	15.83	3.41	1.55	14.98	3.41	1.55
9	6145	2.18	47.73	3.02	1.38	45.32	3.03	1.39
10	13313	2.17	157.28	3.30	1.52	148.19	3.27	1.51
11	28673	2.15	628.37	4.00	1.86	596.29	4.02	1.87
12	61441	2.14	3019.68	4.81	2.25	2845.88	4.77	2.23
13	131073	2.13	13603.54	4.50	2.12	12611.93	4.43	2.08
14	278529	2.12	56331.12	4.14	1.95	50940.00	4.04	1.91

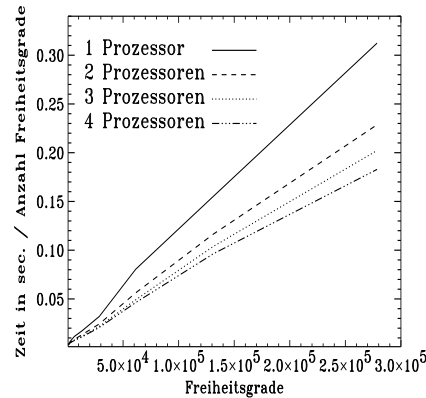


Abbildung 35: Ausführungszeiten für das Lösen der Problemstellung L1 mittels Jacobi-vorkonditioniertem BiCGSTAB (precBiCGSTAB) für 1, 2, 3 und 4 Prozessoren

Tabelle 17: Effizienz  $\text{Eff}_P(l)$ 

$l$	Effizienz von BiCGSTAB			Speed-Up von BiCGSTAB		
	$P = 2$	$P = 3$	$P = 4$	$P = 2$	$P = 3$	$P = 4$
10	0.65	0.49	0.39	1.30	1.48	1.57
11	0.66	0.49	0.39	1.32	1.46	1.54
12	0.72	0.55	0.43	1.43	1.64	1.74
13	0.65	0.50	0.40	1.33	1.49	1.61
14	0.68	0.51	0.43	1.37	1.54	1.71

Die Werte der beiden Tabellen 15 und 16 sind in Abbildung 35 zusammengefaßt. Die aus den Werten ermittelten Effizienzen finden sich in Tabelle 17. Hieraus erkennt man, daß mit wachsender Leveltiefe die Effizienz und damit

der Speed-Up geringfügig zunehmen. Die Begründung hierfür ist, daß die parallelen Routinen einen immer größer werdenden Aufgabenteil zu bewältigen haben und der kommunikative Aufwand immer kleiner gegenüber der gesamten Rechenzeit wird. Dabei ist bei der Parallelisierung des Enthierarchisierungsoperators darauf zu achten, daß hier die Bearbeitungsreihenfolge der Gitterpunkte nicht unabhängig voneinander ist: zur Berechnung eines Ergebnisses auf dem Level  $l$  müssen Ergebnisse auf dem Level  $k$  mit  $k \leq l$  zur Verfügung stehen. So ergibt eine Parallelisierung, die dem Unterraumschema folgt, eine Steigerung der Werte aus Tabelle 17 auf Werte vergleichbar mit den Werten aus Tabelle 14 für den Operator  $\mathbf{D}_{00}$ , jedoch lassen sich keine Werte vergleichbar mit denen der skalaren Multiplikation für den Löser BiCGSTAB erreichen.

Abschließend wollen wir einen Iterationsschritt des iterativen Löser prec-BiCGSTAB auf Level 14 (278529 Knoten) untersuchen. Wir berechnen dabei den Mittelwert der Ausführungszeit von 10 aufeinanderfolgende Iterationen. Das Ergebnis ist in Abbildung 36 zu sehen. Hier erkennt man ebenfalls, daß die Anwendung des Operators den meisten Aufwand verursacht, wohingegen die Operationen "skalare Multiplikation", "Vektoraddition" und "Skalarprodukt" zeit-technisch gegenüber dem Operator kaum ins Gewicht fallen. In Abbildung 36 bedeutet die Abszisse die Nummer der innerhalb des prec-BiCGSTAB auszuführenden Vektor- bzw. Matrix-Operation. Im einzelnen sind dies

- die Anwendung des Operators: Operations-Nummern 4 und 10
- die Anwendung des Vorkonditionierers: Operations-Nummern 3 und 9
- eine Vektor-Addition: Operations-Nummern 2, 7, 14 und 15
- eine skalare Multiplikation: Operations-Nummern 5 und 11
- ein Skalarprodukt: Operations-Nummern 1, 6, 8, 12, 13 und 16

In Abbildung 36 ist das Ergebnis der Mittelung der Iterationen sowohl in linearer als auch in logarithmische Skalierung der Zeit enthalten. Hierdurch wird zum einen die Skalierung des diskreten Differentialoperators deutlich (linke Graphik), zum anderen wird die Skalierung der übrigen Vektor-Operatoren ersichtlich (rechte Graphik).

Betrachtet man sich die erzielten Effizienzen und Speed-Up's für 3 und 4 Prozessoren, so erkennt man, daß die Ergebnisse nicht befriedigend sind. Der Hauptgrund hierfür liegt jedoch nicht in der zugrundeliegenden Parallelisierungsstrategie, sondern ist in der verwendeten Rechnerarchitektur zu suchen. Verwendet wurde, wie bereits eingangs erwähnt, eine SGI O200 von

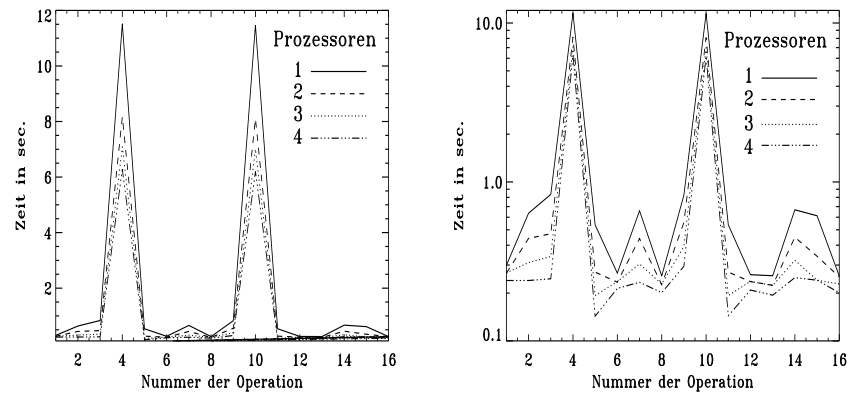


Abbildung 36: Ausführungszeiten eines precBiCGSTAB-Schrittes auf Level 14, 278529 Punkte: lineare Skalierung der Zeit (links) und logarithmische Skalierung der Zeit (rechts)

Silicon Graphics mit insgesamt 4 Prozessoren. Dabei werden zwei SGI O200 mit jeweils 2 Prozessoren durch einen Cray-Link miteinander gekoppelt. Läuft ein Programm nun auf zwei Prozessoren, die in einer SGI O200 auf einer gemeinsamen Platine schnell miteinander kommunizieren können, so ist ein guter Performance-Gewinn gegenüber der sequentiellen Version, die nur auf einem Prozessor rechnet, zu erhalten. Bei Verwendung von 3 oder 4 Prozessoren ist jedoch bei der zur Verfügung stehenden Rechnerarchitektur eine Kommunikation der Prozessoren über den Cray-Link erforderlich, der zu einer Reduktion der Gesamtperformance führt. Daher sind bessere Ergebnisse auf SMP-Maschinen zu erwarten, bei denen die vorhandenen Prozessoren alle gleich schnell miteinander kommunizieren können. Dies ist etwa bei der SGI Origin2000 der Fall, so daß beispielsweise unter Verwendung dieser Rechnerarchitektur weitere Möglichkeiten zur Optimierungen bei der Parallelisierung der einzelnen Operatoren durchgeführt werden können.

## 7 Numerische Resultate

Nachfolgend wollen wir die theoretischen Resultate für die Finiten Differenzen Operatoren auf dünnen Gitter mit numerischen Berechnungen vergleichen. Dabei werden sowohl reguläre als auch adaptive dünne Gitter betrachtet.

Die Gleichungen, mit denen wir uns in den ersten beiden Teilen (Abschnitte 7.2 und 7.3) dieses numerischen Kapitels befassen wollen, sind skalare elliptische Differentialgleichungen 2. Ordnung. Im zweiten Abschnitt (Abschnitt 7.4) behandeln wir sowohl skalare als auch gekoppelte, nichtlineare parabolische Problemstellungen. Anwendungen von dünnen Gittern auf hyperbolische Differentialgleichungen, die in dieser Arbeit nicht betrachtet werden, finden sich in [GrZu98]. Anschließend widmen wir uns in Abschnitt 7.5 der Lösung der Navier Stokes Gleichungen auf dünnen Gittern.

Sämtlich Rechnungen wurden durchgeführt auf Workstations der Firma Silicon Graphics (O2 mit MIPS R10000 (150MHz), IP32 und IRIX 6.3, O2 mit MIPS R5000 (180MHz), IP32 und IRIX 6.3, O200 mit MIPS R10000 (180MHz), IP32 mit IRIX 6.4 sowie Indigo<sup>2</sup> mit MIPS R4400 (200MHz) und IRIX 6.2) am Institut für Angewandte Mathematik der Universität Bonn.

### 7.1 Diskrete Normen auf dünnen Gittern

In den nachfolgenden numerischen Beispielen interessieren wir uns für die verursachten Fehler, die wir in der  $L_\infty$ -Norm, der  $L_2$ -Norm und der  $L_1$ -Norm messen wollen. Hierzu müssen wir zu den kontinuierlichen Normen diskrete Analoga auf dünnen Gittern bereitstellen, die eine möglichst gute Approximation der kontinuierlichen Normen darstellen. Zudem werden wir den Fehler in festgelegten Punkten betrachten.

Wie in [Bun92, Bun98] nehmen wir als Maximumsnorm den maximalen Absolutwert auf allen Gitterpunkten des dünnen Gitters. Zur Berechnung einer diskreten  $L_2$ -Norm einer diskreten Funktion  $u^{\varepsilon, \|\cdot\|}$  auf dünnen Gittern gehen wir wie folgt vor: nach einer punktwisen Multiplikation der diskreten Funktionswerte stellen wir das erhaltene Produkt dar als Linearkombination der stückweise  $d$ -linearen Ansatzfunktionen der hierarchischen Basis und integrieren diese Darstellung. Durch anschließendes Ziehen der Wurzel erhalten wir die diskrete Variante der  $L_2$ -Norm auf dünnen Gittern. Die Vorgehensweise für die Multiplikation der Funktion  $u^{\varepsilon, \|\cdot\|}$  mit sich selbst, wie es für die Berechnung der (diskreten)  $L_2$ -Norm notwendig ist, ist dabei die ab Seite 37 unten ff. vorgestellte. Damit unterscheidet sich die hier verwendete diskrete  $L_2$ -Norm geringfügig von der in [Bun92, Bun98] vorgestellten Form.

Die Berechnung der  $L_1$ -Norm kann auf dünnen Gittern nur durch einen

sehr großen programmiertechnischen Aufwand geschehen. Daher interpolieren wir die numerische Lösung (jedoch ausschließlich zu Auswertungszwecken) auf ein volles Gitter und berechnen dann eine diskrete Version der  $L_1$ -Norm. Hierzu summieren wir die Absolutbeträge einer Funktion, deren diskrete  $L_1$ -Norm zu berechnen ist, auf einem vollen Gitter auf, multiplizieren diesen Wert mit dem Volumen des Gebiets und dividieren anschließend durch die Anzahl der Gitterpunkte des zugrundeliegenden vollen Gitters.

Alle drei diskreten Normen approximieren zumindest asymptotisch die jeweiligen kontinuierlichen Normen genügend genau. Wie sich nachfolgend auch in den numerischen Ergebnissen zeigen wird, kann man unter Verwendung der angegebenen Normen auch die erwarteten Konvergenzordnungen numerisch nachweisen.

Um einen Vergleich zu Fehlermessungen zu haben, die auf vollen Gittern durchgeführt werden, interpolieren wir (wiederum ausschließlich zum Zweck der Fehlerberechnung) die numerische Lösung auf ein volles Gitter und berechnen auch dort die Maximumsnorm sowie die  $L_2$ -Norm auf die angegebene Weise. Die entsprechenden Ergebnisse sind in den Tabellen der numerischen Resultate im Fall von zweidimensionalen Beispielen in der Regel mit angegeben. In höheren Dimensionen verzichten wir auf Fehlerberechnungen, die eine Interpolation auf ein volles Gitter voraussetzen, da dort der durchzuführende Interpolationsaufwand zu groß ist. Da sich jedoch in den zweidimensionalen Beispielen die Maximumsnorm, die  $L_2$ -Norm, die Punktfehler und auch die  $L_1$ -Norm qualitativ gleich verhalten, ist zu erwarten, daß dies auch in höheren Dimensionen der Fall ist. Dies zeigen auch Auswertungen in drei Raumdimensionen auf Gittern kleinerer Level, wo die Konvergenzordnungen, gemessen in den verschiedenen Normen, wiederum qualitativ gleiche Ergebnisse liefern.

## 7.2 Elliptische Differentialgleichungen 2. Ordnung

### 7.2.1 Laplace- und Poisson-Gleichung

Die beiden einfachsten elliptischen Differentialgleichungen 2. Ordnung sind die Laplace-Gleichung

$$\Delta u = 0 \quad (7.1)$$

und die Poisson-Gleichung

$$\Delta u = f \quad (7.2)$$

zusammen mit passenden Randbedingungen. Diese beiden Gleichungen werden in zahlreichen Büchern und Artikeln über elliptische Differentialgleichungen als Referenzbeispiele für die Betrachtung numerischer Verfahren herangezogen, weswegen wir uns zu Anfang der numerischen Berechnungen ebenfalls auf diese beiden Gleichungen zurückziehen wollen.

**Beispiel L1:** Das erste Problem, das wir untersuchen, ist ein Laplace-Problem (7.1) in  $2D$  mit der exakten Lösung

$$u(x, y) = \frac{\sinh(\pi(1-x))}{\sinh(\pi)} \sin(\pi y)$$

und zugehörigen Dirichlet-Randbedingungen auf dem Gebiet  $\Omega = [0, 1]^2$ , das auf dem Gebiet der dünnen Gitter ein Standardbeispiel darstellt [Zen90, Bun92, Bun98].

Die Ergebnisse für verschiedene Level der Diskretisierung auf regulären dünnen Gittern sind in den Tabellen 18(a), 18(b) und 18(c) zu finden. Dort und auch in den weiteren Beispielen sind neben der Zahl der Punkte auch verschiedene Fehlernormen sowie die Quotienten von Ergebnissen von aufeinanderfolgenden Leveln aufgelistet. Dabei bedeuten die Bezeichnungen der einzelnen Spalten der Tabellen im einzelnen:  $l$  der Level der Dünngitter-Diskretisierung,  $P$  die Anzahl der Punkte,  $q_P := P_l/P_{l-1}$  deren Quotient,  $L_\infty^S := \|u - u_l^S\|_\infty$  die Maximumsnorm des Fehlers, wobei  $u$  die exakte Lösung und  $u_l^S$  die auf dem Level  $l$  berechnete numerische Lösung darstellt,  $q_{L_\infty^S}$  definiert als

$$q_{L_\infty^S} := \frac{\|u - u_{l-1}^S\|_\infty}{\|u - u_l^S\|_\infty}$$

ist der Quotient bzgl.  $L_\infty^S$  für sukzessive Level,  $q_{1,S}$  ist definiert vermöge  $q_{1,S} = q_{L_\infty^S}/q_P$  und ist ein Maß für Aufwand zu Genauigkeit bzgl. der Maximumsnorm,  $L_2^S$  die  $L_2$ -Norm des Fehlers,  $q_{L_2^S}$  definiert als

$$q_{L_2^S} := \frac{\|u - u_{l-1}^S\|_2}{\|u - u_l^S\|_2}$$

ist der Quotient bzgl.  $L_2^S$  für sukzessive Level,  $q_{2,S}$  ist definiert vermöge  $q_{2,S} = q_{L_2^S}/q_P$  und ist ein Maß für Aufwand zu Genauigkeit bzgl. der  $L_2$ -Norm,  $T$  ist der Punktfehler im Punkt  $(\frac{1}{\pi}, \frac{1}{\pi})$ , in zwei Dimensionen bzw. im Punkt  $(\frac{1}{\pi}, \frac{1}{\pi}, \frac{1}{\pi})$  für drei Dimensionen,  $q_T$  definiert als

$$q_T := \frac{|u(\frac{1}{\pi}, \frac{1}{\pi}) - u_{l-1}^S(\frac{1}{\pi}, \frac{1}{\pi})|}{|u(\frac{1}{\pi}, \frac{1}{\pi}) - u_l^S(\frac{1}{\pi}, \frac{1}{\pi})|}$$

in zwei Dimensionen bzw.

$$q_T := \frac{|u(\frac{1}{\pi}, \frac{1}{\pi}, \frac{1}{\pi}) - u_{l-1}^S(\frac{1}{\pi}, \frac{1}{\pi}, \frac{1}{\pi})|}{|u(\frac{1}{\pi}, \frac{1}{\pi}, \frac{1}{\pi}) - u_l^S(\frac{1}{\pi}, \frac{1}{\pi}, \frac{1}{\pi})|}$$

in drei Dimensionen ist der Quotient bzgl.  $T$  für sukzessive Level sowie  $q_3^S = q_T/q_P$ , und es ist  $M$  der Punktfehler im Schwerpunkt des jeweiligen Rechengebiets,  $q_M$  der Quotient bzgl.  $M$  auf sukzessiven Leveln (definiert analog zu  $q_T$ ) sowie  $q_4^S = q_M/q_P$ . Dabei bedeutet das hochgestellte  $S$ , daß die entsprechenden Größen auf Dünngitter-Punkten (sparse grid points) ermittelt werden.

Zudem bedeuten auf vollen Gittern  $L_\infty^V$  die Maximumsnorm des Fehlers,  $q_{L_\infty^V}$  der Quotient bzgl.  $L_\infty^V$ ,  $q_{1,V}$  ist definiert vermöge  $q_{1,V} = q_{L_\infty^V}/q_P$  und ist ein Maß für Aufwand zu Genauigkeit bzgl. der Maximumsnorm auf vollen Gittern,  $L_2^V$  die  $L_2$ -Norm des Fehlers,  $q_{L_2^V}$  der Quotient bzgl.  $L_2^V$ ,  $q_{2,V}$  ist definiert vermöge  $q_{2,V} = q_{L_2^V}/q_P$  und ist ein Maß für Aufwand zu Genauigkeit bzgl. der  $L_2$ -Norm auf vollen Gittern,  $L_1^V$  die  $L_1$ -Norm des Fehlers,  $q_{L_1^V}$  der Quotient bzgl.  $L_1^V$ , und  $q_{3,V}$  ist definiert vermöge  $q_{3,V} = q_{L_1^V}/q_P$  und ist ein Maß für Aufwand zu Genauigkeit in der  $L_1$ -Norm. Dabei wird die Lösung auf dem dünnen Gitter nur zum Zwecke der Fehlerberechnung auf ein volles Gitter interpoliert.

Anhand der Tabellen 18(a), 18(b) und 18(c) zeigt sich, daß die Konvergenzordnung bezüglich aller betrachteten Normen  $\mathcal{O}(h^2)$  ist, wobei  $h$  die feinste auftretende Maschenweite (Randmaschenweite) bedeutet. Hierbei ist zu beachten, daß sich bei der Verwendung von Finiten Elementen lediglich eine Konvergenzordnung von  $\mathcal{O}(h^2 \log(h^{-1}))$  einstellt (vgl. [Bun92], Seite 86ff.) Zudem ist ein unterschiedliches Fehlerverhalten der Fehler bei Verwendung von Finiten Differenzen und Finiten Elementen zu beobachten. Im Falle Finiten Elemente erhält man oszillierende Fehler (vgl. wiederum [Bun92], Seite 86ff.), während man im Fall Finiten Differenzen einen glatten Fehler erhält (vgl. Abbildung 38). Betrachtet man in den Tabellen zu diesem Beispiel die mit  $q_{1,S}$ ,  $q_{2,S}$ ,  $q_{1,V}$ ,  $q_{2,V}$  sowie  $q_{3,V}$  und  $q_{3,S}$  bezeichneten Spalten, die das Verhältnis von Aufwand zu Genauigkeit in den verwendeten Normen anzeigen, so zeigt sich, daß die Diskretisierung mittels dünner Gitter einer Diskretisierung höherer Ordnung entspricht. Wären die Quotienten in den



Tabelle 18: (a) Fehler für Problem L1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	1.28 <sub>-2</sub>	-	-	7.00 <sub>-3</sub>	-	-
3	49	2.33	3.15 <sub>-3</sub>	4.06	1.74	1.75 <sub>-3</sub>	4.00	1.72
4	113	2.31	7.95 <sub>-4</sub>	3.96	1.71	4.34 <sub>-4</sub>	4.03	1.74
5	257	2.27	2.00 <sub>-4</sub>	3.97	1.75	1.08 <sub>-4</sub>	4.02	1.77
6	577	2.25	5.03 <sub>-5</sub>	3.98	1.77	2.69 <sub>-5</sub>	4.01	1.78
7	1281	2.22	1.26 <sub>-5</sub>	4.00	1.80	6.72 <sub>-6</sub>	4.00	1.80
8	2817	2.20	3.15 <sub>-6</sub>	4.00	1.82	1.67 <sub>-6</sub>	4.01	1.82
9	6145	2.18	7.91 <sub>-7</sub>	3.98	1.83	4.21 <sub>-7</sub>	3.97	1.82
10	13313	2.17	1.97 <sub>-7</sub>	4.01	1.85	1.05 <sub>-7</sub>	4.01	1.85
11	28673	2.15	4.93 <sub>-8</sub>	4.00	1.86	2.63 <sub>-8</sub>	3.99	1.86
12	61441	2.14	1.23 <sub>-8</sub>	4.01	1.87	6.57 <sub>-9</sub>	4.00	1.87
13	131073	2.13	3.07 <sub>-9</sub>	4.01	1.88	1.64 <sub>-9</sub>	4.01	1.88
14	278529	2.12	7.69 <sub>-10</sub>	3.99	1.88	4.08 <sub>-10</sub>	4.02	1.90

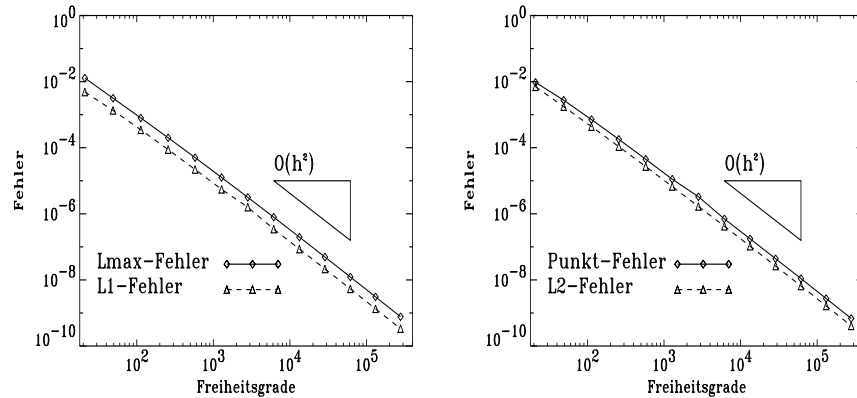
Tabelle 18: (b) Fehler für Problem L1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
2	21	-	1.28 <sub>-2</sub>	-	-	6.08 <sub>-3</sub>	-	-
3	49	2.33	3.15 <sub>-3</sub>	4.06	1.74	1.67 <sub>-3</sub>	3.64	1.56
4	113	2.31	8.01 <sub>-4</sub>	3.93	1.70	4.27 <sub>-4</sub>	3.91	1.69
5	257	2.27	2.01 <sub>-4</sub>	3.99	1.76	1.07 <sub>-4</sub>	3.99	1.76
6	577	2.25	5.05 <sub>-5</sub>	3.98	1.77	2.69 <sub>-5</sub>	3.98	1.77
7	1281	2.22	1.26 <sub>-5</sub>	4.00	1.80	6.72 <sub>-6</sub>	4.00	1.80
8	2817	2.20	3.15 <sub>-6</sub>	4.00	1.82	1.67 <sub>-7</sub>	4.01	1.82
9	6145	2.18	7.91 <sub>-7</sub>	3.98	1.83	4.21 <sub>-7</sub>	3.97	1.82
10	13313	2.17	1.97 <sub>-7</sub>	4.01	1.85	1.05 <sub>-7</sub>	4.01	1.85
11	28673	2.15	4.93 <sub>-8</sub>	4.00	1.86	2.63 <sub>-8</sub>	3.99	1.86
12	61441	2.14	1.23 <sub>-9</sub>	4.01	1.87	6.57 <sub>-9</sub>	4.00	1.87
13	131073	2.13	3.07 <sub>-9</sub>	4.01	1.88	1.64 <sub>-9</sub>	4.01	1.88
14	278529	2.12	7.69 <sub>-10</sub>	3.99	1.88	4.08 <sub>-10</sub>	4.02	1.90

genannten Spalten 1, so würde dies bedeuten, daß sich der Fehler um genau den Faktor verringert, um den die Anzahl der Punkte zunimmt, was einem  $\mathcal{O}(h^2)$ -Verhalten entsprechen würde (man beachte hierbei, daß wir uns im zweidimensionalen Fall befinden). Analog würde der Faktor  $\frac{1}{2}$  ein  $\mathcal{O}(h)$ -Verhalten widerspiegeln. Die erzielten Faktoren sind jedoch größer als 1 und gehen gegen den Wert 2, womit annähernd ein  $\mathcal{O}(h^4)$ -Verhalten bzgl. Aufwand zu Genauigkeit zu beobachten ist. Ein Vergleich der berechneten  $L_\infty$ -Normen und  $L_2$ -Normen auf dünnen und vollen Gittern (nach Projektion der Dünngitter-Funktion auf ein volles Gitter) zeigt, daß bis auf geringfügige Unterschiede bei den ersten Leveln diese Normen annähernd

Tabelle 18: (c) Fehler für Problem L1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	$4.85_{-3}$	-	-	$9.48_{-3}$	-	-
3	49	2.33	$1.35_{-3}$	3.59	1.54	$2.74_{-3}$	3.46	1.48
4	113	2.31	$3.49_{-4}$	3.87	1.68	$7.11_{-4}$	3.85	1.67
5	257	2.27	$8.79_{-5}$	3.97	1.75	$1.80_{-4}$	3.95	1.74
6	577	2.25	$2.20_{-5}$	3.99	1.77	$4.48_{-5}$	4.02	1.79
7	1281	2.22	$5.51_{-6}$	3.99	1.80	$1.12_{-5}$	4.00	1.80
8	2817	2.20	$1.59_{-6}$	4.00	1.82	$3.29_{-6}$	4.01	1.82
9	6145	2.18	$3.45_{-7}$	3.99	1.83	$7.02_{-7}$	4.69	2.15
10	13313	2.17	$8.59_{-8}$	4.02	1.85	$1.75_{-7}$	4.01	1.85
11	28673	2.15	$2.15_{-8}$	3.99	1.86	$4.38_{-8}$	4.00	1.86
12	61441	2.14	$5.38_{-9}$	4.00	1.86	$1.10_{-8}$	3.98	1.86
13	131073	2.13	$1.34_{-9}$	4.01	1.86	$2.73_{-9}$	4.03	1.89
14	278529	2.12	$3.36_{-10}$	3.99	1.88	$6.82_{-10}$	4.00	1.89

Abbildung 37: Konvergenzverhalten der jeweiligen Fehler für das Laplace-Problem L1:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)

übereinstimmen.

In Abschnitt 5.4 wurde gezeigt, daß eine Diagonalskalierung der Diskretisierungsmatrizen einen großen Einfluß auf die Kondition des resultierenden Gleichungssystems hat. Dies erkennt man auch in einem Vergleich der benötigten Iterationszahlen, die die verwendeten iterativen Löser BiCGSTAB [Vor95] und vorkonditioniertes BiCGSTAB (precBiCGSTAB) benötigen, um die jeweiligen Lösungen zu berechnen. Ein Vergleich der Iterationszahlen für das Beispiel L1 bis Level 10 ist in Tabelle 19 zu finden. Dort erkennt man, daß wie erwartet die benötigten Iterationszahlen im vorkonditionierten Fall erheblich geringer sind als im nicht-vorkonditionierten Fall.

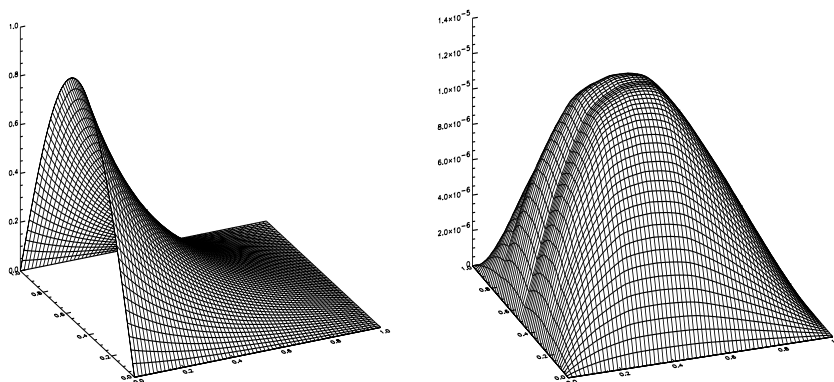


Abbildung 38: Lösung des Laplace-Problems L1 (links) sowie Fehler auf dem Level 7 (rechts, geänderte Skalierung)

Tabelle 19: Iterationszahlen für BiCGSTAB und precBiCGSTAB für das Problem L1

Level	3	4	5	6	7	8	9	10
BiCGSTAB	13	25	70	221	1.114	3.351	11.383	43.849
precBiCGSTAB	11	18	28	52	67	106	142	206

**Beispiel L2:** Das zweite Beispiel ist eine Laplace-Gleichung mit inhomogenen Dirichlet-Randbedingungen, deren exakte Lösung eine Wurzelsingularität aufweist. Gesucht ist eine Lösung der Gleichung (7.1) auf dem Gebiet  $[0, 1]^2$  mit vorgegebenen Randbedingungen, so daß

$$u(x, y) = \operatorname{Re}(z^{\frac{1}{2}}) = (x^2 + y^2)^{\frac{1}{4}} \cdot \cos\left(\frac{\arctan\left(\frac{x}{y}\right)}{2}\right)$$

die exakte Lösung darstellt. Somit verletzt die Lösungsfunktion die Glattheitsvoraussetzungen, die wir in Kapitel 2 (vgl. Anfang des Abschnitts 2.1) gefordert haben.

Aus den Tabellen 20(a) und 20(b) erkennt man klar, daß aufgrund der Singularität in der Lösung eine wesentlich kleinere Konvergenzordnung im Vergleich zur ersten Problemstellung L1 gegeben ist. Dabei sind die Reduktionsraten für die  $L_\infty$ -Norm und die  $L_2$ -Norm mit denen vergleichbar, die man mittels der Methode der Finiten Elemente (vgl. [Bun92], Seite 89ff) erhält. Aufgrund der Singularität in der Lösungsfunktion ist die Verwendung adaptiver dünner Gitter notwendig. Als Verfeinerungskriterium wird hierfür das in Abschnitt 5.5 vorgestellte Kriterium verwendet. Die Ergebnisse des selbstadaptiven Algorithmus, dessen Startgitter ein reguläres dünnes Gitter auf dem Level 3 ist, sind in den Tabellen 21(a) und 21(b) zu sehen, die

Tabelle 20: (a) Fehler für Problem L2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	1.88 <sub>-2</sub>	-	-	9.82 <sub>-3</sub>	-	-
3	49	2.33	4.69 <sub>-2</sub>	0.40	0.17	1.40 <sub>-2</sub>	0.70	0.30
4	113	2.31	4.89 <sub>-2</sub>	0.96	0.42	1.41 <sub>-2</sub>	0.99	0.43
5	257	2.27	5.45 <sub>-2</sub>	0.90	0.40	1.25 <sub>-2</sub>	1.13	0.50
6	577	2.25	5.00 <sub>-2</sub>	1.09	0.48	1.02 <sub>-2</sub>	1.23	0.55
7	1281	2.22	4.65 <sub>-2</sub>	1.08	0.49	7.96 <sub>-3</sub>	1.28	0.58
8	2817	2.20	4.05 <sub>-2</sub>	1.15	0.52	6.03 <sub>-3</sub>	1.32	0.60
9	6145	2.18	3.53 <sub>-2</sub>	1.15	0.53	4.47 <sub>-3</sub>	1.35	0.62
10	13313	2.17	3.01 <sub>-2</sub>	1.17	0.54	3.27 <sub>-3</sub>	1.37	0.63
11	28673	2.15	2.57 <sub>-2</sub>	1.17	0.54	2.37 <sub>-3</sub>	1.38	0.64
12	61441	2.14	2.17 <sub>-2</sub>	1.18	0.55	1.70 <sub>-3</sub>	1.39	0.65
13	131073	2.13	1.84 <sub>-2</sub>	1.18	0.55	1.22 <sub>-3</sub>	1.39	0.65
14	278529	2.12	1.55 <sub>-2</sub>	1.19	0.56	8.70 <sub>-4</sub>	1.40	0.66

Tabelle 20: (b) Fehler für Problem L2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$T$	$q_T$	$q_{3,S}$	$M$	$q_M$	$q_{4,S}$
2	21	-	1.65 <sub>-2</sub>	-	-	1.29 <sub>-2</sub>	-	-
3	49	2.33	3.35 <sub>-2</sub>	0.49	0.21	1.28 <sub>-2</sub>	1.01	0.43
4	113	2.31	2.55 <sub>-2</sub>	1.31	0.57	1.02 <sub>-2</sub>	1.25	0.54
5	257	2.27	1.86 <sub>-2</sub>	1.37	0.60	7.43 <sub>-3</sub>	1.37	0.60
6	577	2.25	1.33 <sub>-2</sub>	1.40	0.62	5.30 <sub>-3</sub>	1.40	0.62
7	1281	2.22	9.49 <sub>-3</sub>	1.40	0.63	3.75 <sub>-3</sub>	1.41	0.64
8	2817	2.20	6.72 <sub>-3</sub>	1.41	0.64	2.66 <sub>-3</sub>	1.41	0.64
9	6145	2.18	4.76 <sub>-3</sub>	1.41	0.65	1.88 <sub>-3</sub>	1.41	0.65
10	13313	2.17	3.37 <sub>-3</sub>	1.41	0.65	1.33 <sub>-3</sub>	1.41	0.65
11	28673	2.15	2.38 <sub>-3</sub>	1.42	0.66	9.40 <sub>-4</sub>	1.41	0.66
12	61441	2.14	1.68 <sub>-3</sub>	1.42	0.66	6.65 <sub>-4</sub>	1.41	0.66
13	131073	2.13	1.19 <sub>-3</sub>	1.41	0.66	4.70 <sub>-4</sub>	1.41	0.66
14	278529	2.12	8.41 <sub>-4</sub>	1.41	0.67	3.32 <sub>-4</sub>	1.42	0.67

Abbildung 40 zeigt das Konvergenzverhalten für die Berechnung des Beispiels L2 auf adaptiven Gittern. Dabei ist insbesondere zu erkennen, daß der Fehler gemessen in der  $L_2$ -Norm eine wesentlich höhere Konvergenzordnung auf adaptiven dünnen Gittern verglichen mit den Ergebnissen auf regulären dünnen Gittern aufweist. Man erkennt zudem, daß der Algorithmus aufgrund des schlecht angepaßten Startgitters eine Anlaufphase benötigt, in der die erzielten Resultate nur unwesentlich besser sind als diejenigen auf regulären dünnen Gittern. Ab einem bestimmten Grad der Verfeinerung zeigt sich jedoch, daß der Fehler auf dem adaptive Gitter in allen berechneten Fehlernormen um einige Größenordnungen kleiner ist.

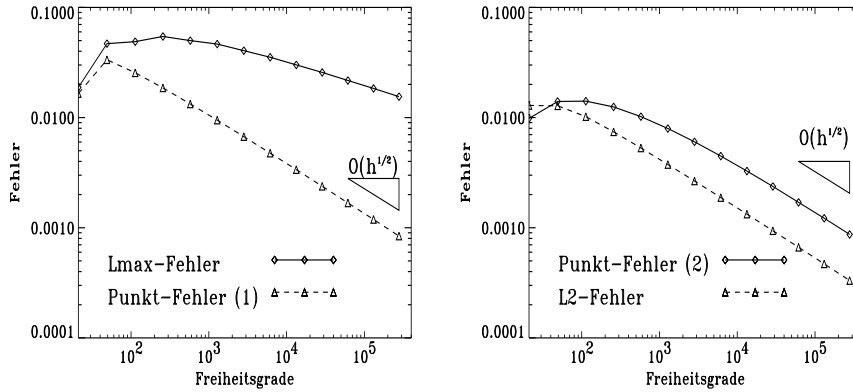


Abbildung 39: Konvergenzverhalten der jeweiligen Fehler für das Laplace-Problem  $L_2$  auf regulären dünnen Gittern:  $L_\infty^S$  und Punktfehler  $P_1 \equiv T$  (links) sowie  $L_2^S$  und Punktfehler  $P_2 \equiv M$  (rechts)

Tabelle 21: (a) Fehler für Problem  $L_2$ , adaptives dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
1	49	-	4.69 <sub>-2</sub>	-	-	1.40 <sub>-2</sub>	-	-
2	85	1.73	4.92 <sub>-2</sub>	0.95	0.55	1.43 <sub>-2</sub>	0.70	0.30
3	163	1.92	5.60 <sub>-2</sub>	0.88	0.46	1.15 <sub>-2</sub>	0.99	0.43
4	309	1.90	5.01 <sub>-2</sub>	1.12	0.59	9.39 <sub>-3</sub>	1.13	0.50
5	585	1.89	4.73 <sub>-2</sub>	1.06	0.56	7.49 <sub>-3</sub>	1.23	0.55
6	1136	1.94	4.05 <sub>-2</sub>	1.17	0.60	5.72 <sub>-3</sub>	1.28	0.58
7	2223	1.96	3.54 <sub>-2</sub>	1.14	0.58	4.30 <sub>-3</sub>	1.32	0.60
8	4394	1.98	3.05 <sub>-2</sub>	1.16	0.59	3.18 <sub>-3</sub>	1.35	0.62
9	5611	1.28	2.50 <sub>-2</sub>	1.22	0.95	1.54 <sub>-3</sub>	1.37	0.63
10	6801	1.21	2.03 <sub>-2</sub>	1.23	1.02	7.33 <sub>-4</sub>	2.10	1.74
11	8069	1.19	1.64 <sub>-2</sub>	1.24	1.04	3.59 <sub>-4</sub>	2.04	1.71
12	9033	1.12	1.25 <sub>-2</sub>	1.31	1.17	2.04 <sub>-4</sub>	1.76	1.57
13	10250	1.13	1.28 <sub>-3</sub>	9.77	8.64	1.45 <sub>-4</sub>	1.41	1.49
14	10346	1.01	4.15 <sub>-4</sub>	3.08	3.05	1.44 <sub>-4</sub>	1.01	1.00
15	16629	1.61	4.15 <sub>-4</sub>	1.00	0.62	4.25 <sub>-5</sub>	3.39	2.11
16	17264	1.04	4.15 <sub>-4</sub>	1.00	0.96	3.12 <sub>-5</sub>	1.36	1.31
17	29155	1.69	4.15 <sub>-4</sub>	1.00	0.59	9.30 <sub>-6</sub>	3.35	1.98
18	50096	1.72	4.15 <sub>-4</sub>	1.00	0.58	3.40 <sub>-6</sub>	2.74	1.59
19	52582	1.05	4.15 <sub>-4</sub>	1.00	0.95	1.31 <sub>-6</sub>	2.60	2.47
20	89190	1.70	4.15 <sub>-4</sub>	1.00	0.59	7.95 <sub>-7</sub>	1.65	0.97
21	142458	1.60	4.15 <sub>-4</sub>	1.00	0.63	7.83 <sub>-7</sub>	1.02	0.63

Im Falle der Punktfehler erhalten wir in diesem Beispiel eine Konvergenz-

Tabelle 21: (b) Fehler für Problem L2, adaptives dünnes Gitter

$l$	$P$	$q_P$	$T$	$q_T$	$q_{3,S}$	$M$	$q_M$	$q_{4,S}$
1	49	-	3.35 <sub>-2</sub>	-	-	1.28 <sub>-2</sub>	-	-
2	85	1.73	2.54 <sub>-2</sub>	1.32	0.76	9.79 <sub>-3</sub>	1.31	0.76
3	163	1.92	1.83 <sub>-2</sub>	1.39	0.72	7.27 <sub>-2</sub>	1.35	0.70
4	309	1.90	1.33 <sub>-2</sub>	1.38	0.73	5.11 <sub>-3</sub>	1.42	0.75
5	585	1.89	9.23 <sub>-3</sub>	1.44	0.76	3.49 <sub>-3</sub>	1.46	0.77
6	1136	1.94	6.50 <sub>-3</sub>	1.42	0.73	2.40 <sub>-3</sub>	1.45	0.75
7	2223	1.96	4.57 <sub>-3</sub>	1.42	0.72	1.66 <sub>-3</sub>	1.45	0.74
8	4394	1.98	3.20 <sub>-3</sub>	1.43	0.72	1.12 <sub>-3</sub>	1.48	0.75
9	5611	1.28	7.75 <sub>-4</sub>	4.13	3.23	1.33 <sub>-4</sub>	8.42	6.58
10	6801	1.21	6.40 <sub>-5</sub>	12.11	10.01	1.28 <sub>-4</sub>	1.04	0.86
11	8069	1.19	1.18 <sub>-4</sub>	0.54	0.45	1.94 <sub>-4</sub>	0.66	0.55
12	9033	1.12	1.56 <sub>-4</sub>	0.76	0.68	2.10 <sub>-4</sub>	0.92	0.82
13	10250	1.13	1.85 <sub>-4</sub>	0.84	0.74	2.21 <sub>-4</sub>	0.95	0.84
14	10346	1.01	1.85 <sub>-4</sub>	1.00	0.99	2.21 <sub>-4</sub>	1.00	0.99
15	16629	1.61	8.00 <sub>-5</sub>	2.31	1.43	7.38 <sub>-5</sub>	2.99	1.86
16	17264	1.04	1.69 <sub>-5</sub>	4.73	4.55	5.48 <sub>-5</sub>	1.35	1.30
17	29155	1.69	2.27 <sub>-6</sub>	7.44	4.40	8.60 <sub>-6</sub>	6.37	3.77
18	50096	1.72	2.55 <sub>-6</sub>	0.89	0.52	1.99 <sub>-7</sub>	43.22	25.13
19	52582	1.05	5.88 <sub>-7</sub>	4.34	4.13	1.50 <sub>-6</sub>	0.13	0.13
20	89190	1.70	7.28 <sub>-8</sub>	8.08	4.75	8.23 <sub>-8</sub>	18.22	10.72
21	142458	1.60	1.27 <sub>-7</sub>	0.28	0.17	3.19 <sub>-8</sub>	2.58	1.61

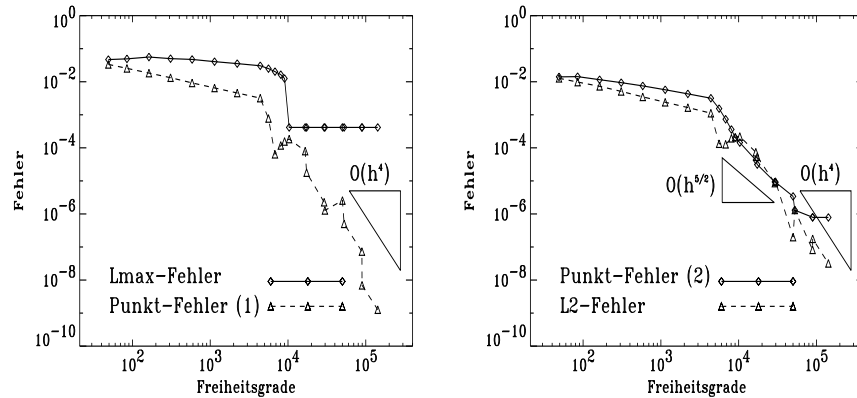


Abbildung 40: Konvergenzverhalten der jeweiligen Fehler für das Laplace-Problem L2 auf adaptiven dünnen Gittern:  $L_\infty^S$  und Punktfehler  $P_1 \equiv T$  (links) sowie  $L_2^S$  und Punktfehler  $P_2 \equiv M$  (rechts)

ordnung von  $\mathcal{O}(h^4)$ , für die  $L_2$ -Norm beobachten wir Konvergenz von der Ordnung  $\mathcal{O}(h^{5/2})$ . Ein graphischer Vergleich der Konvergenzgeschwindigkei-

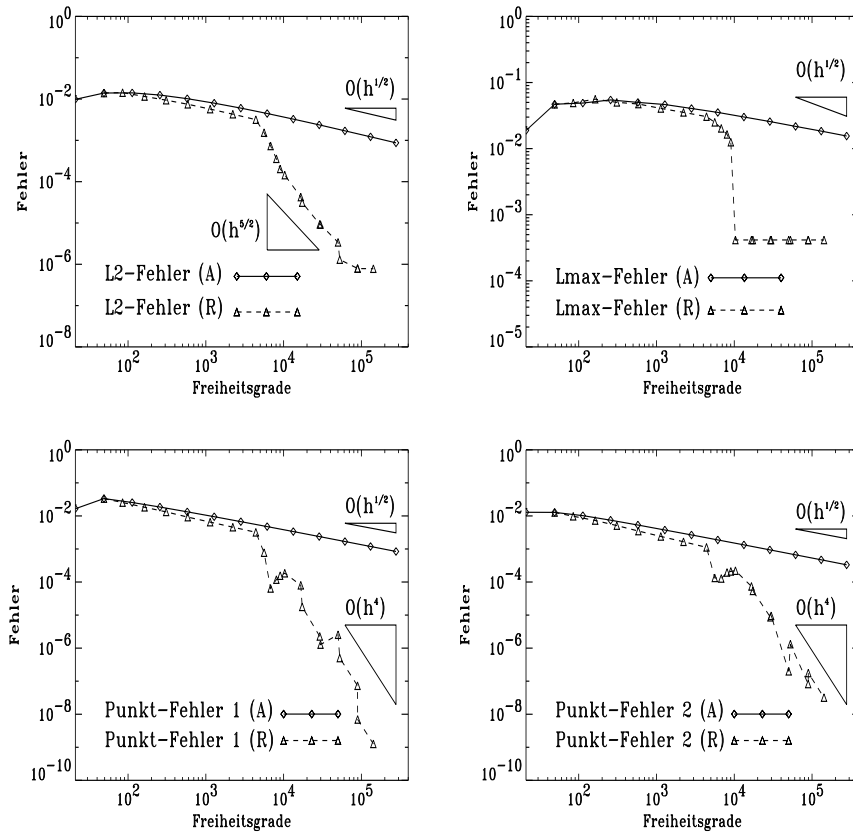


Abbildung 41: Vergleich der Konvergenzverhalten der jeweiligen Fehler auf adaptiven und regulären dünnen Gittern für das Laplace-Problem L2:  $L_2^S$ -Fehler (oben links),  $L_\infty^S$ -Norm (oben rechts), Punktfehler  $P_1 \equiv T$  (unten links) sowie Punktfehler  $P_2 \equiv M$  (unten rechts), unterschiedliche Skalierung

ten auf adaptiven und regulären dünnen Gittern findet sich schließlich in Abbildung 41.

**Beispiel L3:** Sei die exakte Lösung  $u$  des Laplace-Problems (7.1) in drei Dimensionen vermöge

$$u(x, y, z) = z \frac{\cos(4\pi(x - y)) \cdot \sinh(4\pi(x + y))}{\sinh(8\pi)}$$

gegeben. Dieses Beispiel ist entnommen aus [Bun92], wo sich auch die entsprechenden numerischen Resultate befinden, die mit der Methode der Finiten Elemente auf dünnen Gittern erzielt werden können. Die Lösungsfunktion weist in der Umgebung entlang der Kante  $(1, 1, z)$  des Hyperwürfels  $[0, 1]^3$  starke Gradienten auf, wie in Abbildung 43 zu erkennen ist. Dort ist

die Lösungsfunktion auf den Schnittflächen  $(x, y, z = 0.5)$  und  $(x, y, z = 0.95)$  zu sehen. Man erkennt dabei, daß die Gradienten mit wachsendem  $z$  steiler werden. Die numerischen Ergebnisse auf regulären dünnen Gittern zeigen für alle gemessenen Fehlernormen das typische logarithmische Verhalten, wie es auch bei der Lösung mittels Finiter Elemente zu beobachten ist. Dabei gehen jedoch die Konvergenzfaktoren schneller gegen den Faktor 4, als man normalerweise erwarten kann. Ein Grund hierfür ist möglicherweise, daß sich die Lösungsfunktion darstellen läßt vermöge  $u(x, y, z) = u_1(x, y) \cdot u_2(z)$  und somit die Variablen teilweise separiert werden können. In diesem Beispiel wurden zur Berechnung der Ergebnisse reguläre dünne Gitter verwendet. Wie man aus den Tabellen erkennen kann, haben die steilen Gradienten der Lösungsfunktion entlang der Kante  $(1, 1, z)$  keinen großen Einfluß auf das Konvergenzverhalten. Die angesprochenen numerischen Resultate befinden sich in den Tabellen 22(a) und 22(b). Die Abbildung 42 zeigt die erzielten numerischen Ergebnisse wiederum in graphischer Form.

Tabelle 22: (a) Fehler für Problem L3, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	81	-	$3.25_{-2}$	-	-	$1.17_{-2}$	-	-
3	225	2.78	$1.03_{-1}$	0.32	0.11	$1.93_{-2}$	0.61	0.22
4	593	2.64	$4.66_{-2}$	2.21	0.84	$1.02_{-2}$	1.89	0.72
5	1505	2.54	$2.05_{-2}$	2.27	0.89	$3.80_{-3}$	2.68	1.06
6	3713	2.47	$9.23_{-3}$	2.22	0.90	$1.20_{-3}$	3.16	1.28
7	8961	2.41	$2.70_{-3}$	3.42	1.42	$3.40_{-4}$	3.53	1.46
8	21249	2.37	$7.55_{-4}$	3.58	1.51	$9.02_{-5}$	3.77	1.59
9	49665	2.34	$1.98_{-4}$	3.81	1.63	$2.32_{-5}$	3.89	1.66

Tabelle 22: (b) Fehler für Problem L3, reguläres dünnes Gitter

$l$	$P$	$q_P$	$T$	$q_T$	$q_{3,S}$	$M$	$q_M$	$q_{4,S}$
2	81	-	$4.51_{-3}$	-	-	$1.54_{-2}$	-	-
3	225	2.78	$3.27_{-3}$	1.38	0.50	$1.37_{-2}$	1.12	0.40
4	593	2.64	$1.49_{-3}$	2.19	0.83	$6.75_{-3}$	2.03	0.77
5	1505	2.54	$4.79_{-4}$	3.11	1.22	$2.19_{-3}$	3.08	1.21
6	3713	2.47	$1.33_{-4}$	3.60	1.46	$6.11_{-4}$	3.58	1.45
7	8961	2.41	$3.45_{-5}$	3.86	1.60	$1.59_{-4}$	3.84	1.59
8	21249	2.37	$8.76_{-6}$	3.94	1.66	$4.03_{-5}$	3.95	1.66
9	49665	2.34	$2.20_{-6}$	3.98	1.70	$1.01_{-5}$	3.99	1.71

**Beispiel L4:** Im vierten Beispiel betrachten wir ein Laplace–Problem mit der exakten Lösung

$$u(\mathbf{x}) = \prod_{i=1}^d \sin(\pi x_i)$$



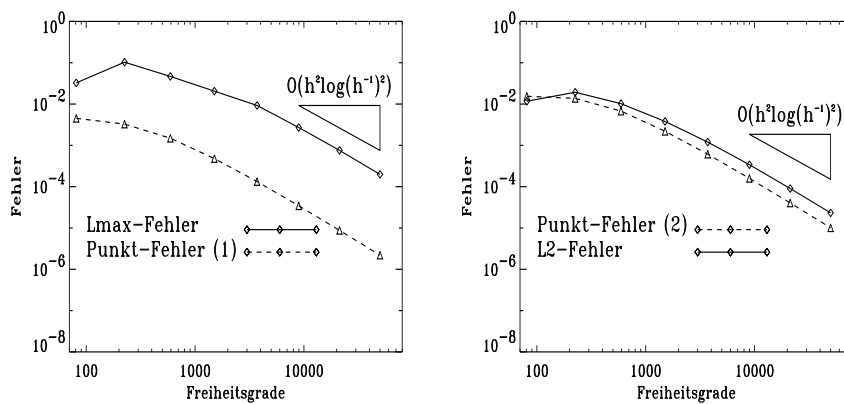


Abbildung 42: Konvergenzverhalten der jeweiligen Fehler für das Laplace-Problem L3:  $L_\infty^S$  und Punktfehler  $P_1 \equiv T$  (links) sowie  $L_2^S$  und Punktfehler  $P_2 \equiv M$  (rechts)

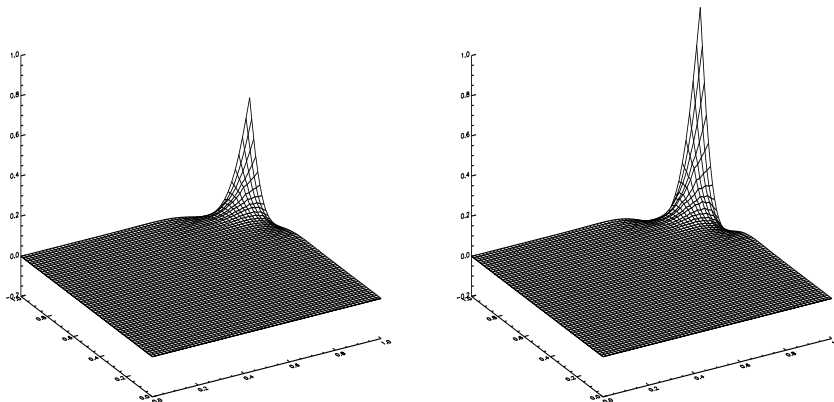


Abbildung 43: Lösungsfunktion des Problems L3: für  $z = 0.5$  (links) und  $z = 0.95$  (rechts)

und der rechten Seite

$$f(\mathbf{x}) = d \cdot \pi^2 \cdot \prod_{i=1}^d \sin(\pi x_i)$$

in  $d \in \{2, \dots, 8\}$  Dimensionen. Die entsprechenden Finite Element Lösungen auf dünnen Gittern zu den einzelnen Dimensionen befinden sich zum Vergleich der hier berechneten Ergebnisse in [BaZe92].

Die numerischen Resultate zeigen, daß mit zunehmender Dimension die Konvergenzgeschwindigkeiten abnehmen, was auf logarithmische Terme zurückzuführen ist. Derselbe Effekt ist in [BaZe92] im Finite Element Kontext zu

beobachten. Für dieses Beispiel ist unabhängig von der betrachteten Dimension der Maximumfehler und der Punktfehler im Schwerpunkt des  $d$ -dimensionalen Würfels identisch. Zudem ist insbesondere in höheren Dimensionen die Berechnung des Fehlers in der  $L_1$ -Norm, die eine Interpolation der Werte auf ein volles Gitter benötigt, sehr teuer. Daher werden in den Tabellen 23 bis 29 nachfolgend nur die Fehler bezüglich der Maximumnorm und der  $L_2$ -Norm angegeben. Die Abbildung 44 veranschaulicht und vergleicht die Ergebnisse in den einzelnen Dimensionen.

Tabelle 23: Fehler für Problem L4, reguläres dünnes Gitter,  $d = 2$ 

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
1	9	-	2.34 <sub>-1</sub>	-	-	1.17 <sub>-1</sub>	-	-
2	21	2.33	6.80 <sub>-2</sub>	3.44	1.48	3.71 <sub>-2</sub>	3.15	1.35
3	49	2.33	2.28 <sub>-2</sub>	2.97	1.27	1.21 <sub>-2</sub>	3.06	1.31
4	113	2.31	7.49 <sub>-3</sub>	3.05	1.32	3.89 <sub>-3</sub>	3.12	1.35
5	257	2.27	2.35 <sub>-3</sub>	3.19	1.41	1.21 <sub>-3</sub>	3.22	1.42
6	577	2.25	7.10 <sub>-4</sub>	3.31	1.47	3.62 <sub>-4</sub>	3.33	1.48
7	1281	2.22	2.08 <sub>-4</sub>	3.41	1.54	1.06 <sub>-4</sub>	3.42	1.54
8	2817	2.20	5.98 <sub>-5</sub>	3.48	1.58	3.03 <sub>-5</sub>	3.49	1.59
9	6145	2.18	1.69 <sub>-5</sub>	3.54	1.62	8.55 <sub>-6</sub>	3.55	1.63
10	13313	2.17	4.70 <sub>-6</sub>	3.59	1.65	2.38 <sub>-6</sub>	3.59	1.65
11	28673	2.15	1.30 <sub>-6</sub>	3.63	1.69	6.55 <sub>-7</sub>	3.63	1.69
12	61441	2.14	3.54 <sub>-7</sub>	3.66	1.71	1.79 <sub>-7</sub>	3.66	1.71
13	131073	2.13	9.62 <sub>-8</sub>	3.68	1.73	4.85 <sub>-8</sub>	3.69	1.73
14	278529	2.12	2.59 <sub>-8</sub>	3.71	1.75	1.31 <sub>-8</sub>	3.70	1.75

Tabelle 24: Fehler für Problem L4, reguläres dünnes Gitter,  $d = 3$ 

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
1	27	-	2.34 <sub>-1</sub>	-	-	8.26 <sub>-2</sub>	-	-
2	81	3.00	7.80 <sub>-2</sub>	3.00	1.00	3.29 <sub>-2</sub>	2.51	0.84
3	225	2.78	3.13 <sub>-2</sub>	2.49	0.90	1.27 <sub>-2</sub>	2.59	0.93
4	593	2.64	1.22 <sub>-2</sub>	2.58	0.98	4.74 <sub>-3</sub>	2.68	1.02
5	1505	2.54	4.45 <sub>-3</sub>	2.73	1.07	1.70 <sub>-3</sub>	2.78	1.09
6	3713	2.47	1.54 <sub>-3</sub>	2.88	1.17	5.84 <sub>-4</sub>	2.91	1.18
7	8961	2.41	5.13 <sub>-4</sub>	3.01	1.25	1.93 <sub>-4</sub>	3.03	1.26
8	21249	2.37	1.65 <sub>-4</sub>	3.11	1.31	6.17 <sub>-5</sub>	3.13	1.32
9	49665	2.34	5.16 <sub>-5</sub>	3.20	1.37	1.92 <sub>-5</sub>	3.21	1.37
10	114689	2.31	1.58 <sub>-5</sub>	3.27	1.42	5.86 <sub>-6</sub>	3.28	1.42

Tabelle 25: Fehler für Problem L4, reguläres dünnes Gitter,  $d = 4$ 

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
1	81	-	$2.34_{-1}$	-	-	$5.84_{-2}$	-	-
2	297	3.67	$8.51_{-2}$	2.75	0.75	$2.75_{-2}$	2.12	0.58
3	945	3.18	$3.86_{-2}$	2.21	0.69	$1.20_{-2}$	2.30	0.72
4	2769	2.93	$1.68_{-2}$	2.29	0.78	$4.99_{-3}$	2.40	0.82
5	7681	2.77	$6.90_{-3}$	2.44	0.88	$1.99_{-3}$	2.51	0.91
6	20481	2.67	$2.67_{-3}$	2.59	0.97	$7.59_{-4}$	2.62	0.98
7	52993	2.59	$9.82_{-4}$	2.72	1.05	$2.77_{-4}$	2.74	1.06
8	133889	2.53	$3.47_{-4}$	2.83	1.12	$9.71_{-5}$	2.94	1.16

Tabelle 26: Fehler für Problem L4, reguläres dünnes Gitter,  $d = 5$ 

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
1	243	-	$2.34_{-1}$	-	-	$4.13_{-2}$	-	-
2	1053	4.33	$9.04_{-2}$	2.59	0.60	$2.22_{-2}$	1.86	0.43
3	3753	3.56	$4.48_{-2}$	2.02	0.57	$1.06_{-2}$	2.09	0.59
4	12033	3.21	$2.13_{-2}$	2.10	0.65	$4.83_{-3}$	2.21	0.69
5	36033	2.99	$9.54_{-3}$	2.23	0.75	$2.09_{-3}$	2.32	0.78
6	102785	2.85	$4.03_{-3}$	2.37	0.83	$8.67_{-4}$	2.41	0.85
7	282625	2.75	$1.61_{-3}$	2.50	0.91	$3.44_{-4}$	2.52	0.92

Tabelle 27: Fehler für Problem L4, reguläres dünnes Gitter,  $d = 6$ 

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
1	729	-	$2.34_{-1}$	-	-	$2.92_{-2}$	-	-
2	3645	5.00	$9.46_{-2}$	2.47	0.49	$1.75_{-2}$	1.67	0.33
3	14337	3.93	$5.01_{-2}$	1.89	0.48	$9.10_{-3}$	1.92	0.49
4	49761	3.47	$2.55_{-2}$	1.96	0.56	$4.43_{-3}$	2.06	0.59
5	159489	3.21	$1.23_{-2}$	2.08	0.65	$2.05_{-3}$	2.16	0.67

Tabelle 28: Fehler für Problem L4, reguläres dünnes Gitter,  $d = 7$ 

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
1	2187	-	$2.34_{-1}$	-	-	$2.06_{-2}$	-	-
2	12393	5.67	$9.79_{-2}$	2.39	0.42	$1.36_{-2}$	1.51	0.27
3	53217	4.29	$5.46_{-2}$	1.79	0.42	$7.57_{-3}$	1.80	0.42
4	198369	3.73	$2.94_{-2}$	1.86	0.50	$3.91_{-3}$	1.94	0.52

**Beispiel L5:** Das fünfte und letzte Poisson–Problem, das wir betrachten wollen, besitzt gemischte Dirichlet– und Neumann–Randbedingungen. Die Lösung des Problems (7.2) sei

$$u(x, y) = 2^8 \cdot (x(1-x)y(1-y))^2$$

Tabelle 29: Fehler für Problem L4, reguläres dünnes Gitter,  $d = 8$ 

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
1	6561	-	$2.34_{-1}$	-	-	$1.46_{-2}$	-	-
2	41553	6.33	$1.01_{-1}$	2.32	0.37	$1.04_{-2}$	1.40	0.22
3	193185	4.65	$5.87_{-2}$	1.72	0.37	$6.17_{-3}$	1.69	0.36

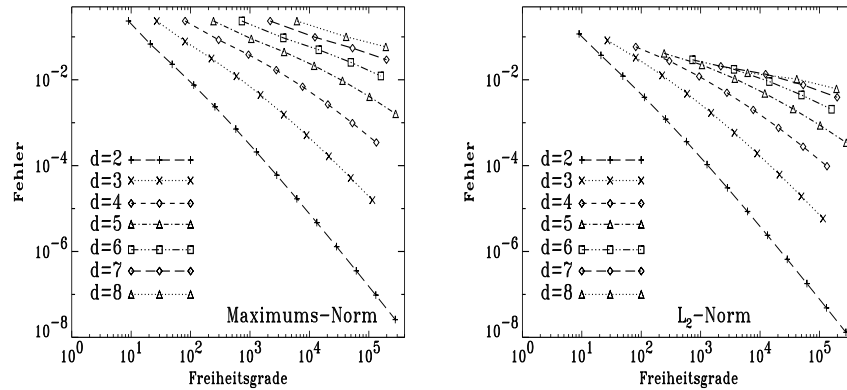


Abbildung 44: Konvergenzverhalten der jeweiligen Fehler für das Laplace-Problem L4 für die Dimensionen 2 bis 8

mit zugehöriger rechter Seite

$$f(x, y) = 2^8 \cdot \left( (2 - 12x + 12x^2)y^2(1 - y)^2 + x^2(1 - x)^2(2 - 12y + 12y^2) \right)$$

auf dem Einheitsquadrat  $\bar{\Omega} = [0, 1]^2$ . Die Randbedingungen sind homogene Dirichlet-Bedingungen auf  $\Gamma_D = \{(x, y) : x = 0 \vee y = 0\}$  sowie homogene Neumann-Randbedingungen auf  $\Gamma_N = \Gamma/\Gamma_D$ .

Die Ergebnisse auf sukzessiven Leveln finden sich in den Tabellen 30(a) und 30(b). Die Abbildung 46 zeigt die Lösungsfunktion sowie die Randbedingungen der Problemstellung.

Anhand der Tabellen 30(a) und 30(b) zeigt sich, daß die Konvergenzordnung bezüglich aller betrachteten Normen  $\mathcal{O}(h^2 \log(h^{-1}))$  ist, wobei  $h$  wie zuvor die feinste auftretende Maschenweite (Randmaschenweite) bedeutet. Hierbei ist zu beachten, daß die Diskretisierung am Neumann-Rand sowohl für den Anteil der Neumann-Ableitung als auch für den Anteil des Operators von erster Konsistenzordnung ist. Zudem ist klar, daß die Anzahl der Randpunkte im Vergleich zur Gesamtzahl der Punkte eines dünnen Gitters wesentlich stärker ins Gewicht fällt ( $\mathcal{O}(N)$  zu  $\mathcal{O}(N \log(N)^{d-1})$ , in diesem Beispiel gilt  $d = 2$ ) als bei einem vollen Gitter ( $\mathcal{O}(N)$  zu  $\mathcal{O}(N^d)$ ). Dennoch ist ein Konvergenz-Verhalten zu beobachten, daß besser als das a priori zu

Tabelle 30: (a) Fehler für Problem L5, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	6.84 <sub>-1</sub>	-	-	2.94 <sub>-1</sub>	-	-
3	49	2.33	2.02 <sub>-1</sub>	3.39	1.45	8.04 <sub>-2</sub>	3.66	1.57
4	113	2.31	6.37 <sub>-2</sub>	3.17	1.37	2.33 <sub>-2</sub>	3.45	1.49
5	257	2.27	2.03 <sub>-2</sub>	3.14	1.38	6.92 <sub>-3</sub>	3.37	1.48
6	577	2.25	6.33 <sub>-3</sub>	3.21	1.43	2.05 <sub>-3</sub>	3.38	1.50
7	1281	2.22	1.92 <sub>-3</sub>	3.30	1.49	6.01 <sub>-4</sub>	3.41	1.54
8	2817	2.20	5.69 <sub>-4</sub>	3.37	1.53	1.74 <sub>-4</sub>	3.45	1.57
9	6145	2.18	1.65 <sub>-4</sub>	3.45	1.58	4.94 <sub>-5</sub>	3.52	1.61
10	13313	2.17	4.91 <sub>-5</sub>	3.36	1.55	1.39 <sub>-5</sub>	3.55	1.64
11	28673	2.15	1.50 <sub>-5</sub>	3.27	1.52	3.85 <sub>-6</sub>	3.61	1.68
12	61441	2.14	4.44 <sub>-6</sub>	3.38	1.58	1.06 <sub>-6</sub>	3.63	1.70
13	131073	2.13	1.29 <sub>-6</sub>	3.44	1.62	2.90 <sub>-7</sub>	3.66	1.72
14	278529	2.12	3.66 <sub>-7</sub>	3.52	1.66	7.86 <sub>-8</sub>	3.69	1.74

Tabelle 30: (b) Fehler für Problem L5, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	1.85 <sub>-1</sub>	-	-	-4.25 <sub>-2</sub>	-	-
3	49	2.33	5.35 <sub>-2</sub>	3.46	1.48	-1.92 <sub>-2</sub>	2.21	0.95
4	113	2.31	1.57 <sub>-2</sub>	3.41	1.48	-2.52 <sub>-4</sub>	76.19	32.98
5	257	2.27	4.59 <sub>-3</sub>	3.42	1.51	-2.78 <sub>-3</sub>	0.09	0.04
6	577	2.25	1.36 <sub>-3</sub>	3.38	1.50	-9.10 <sub>-4</sub>	3.05	1.36
7	1281	2.22	3.98 <sub>-4</sub>	3.42	1.54	-2.83 <sub>-4</sub>	3.22	1.45
8	2817	2.20	1.15 <sub>-4</sub>	3.46	1.57	-8.47 <sub>-5</sub>	3.34	1.52
9	6145	2.18	3.25 <sub>-5</sub>	3.54	1.62	-2.47 <sub>-5</sub>	3.43	1.57
10	13313	2.17	9.09 <sub>-6</sub>	3.58	1.65	-7.04 <sub>-6</sub>	3.51	1.62
11	28673	2.15	2.52 <sub>-6</sub>	3.61	1.68	-1.98 <sub>-6</sub>	3.56	1.65
12	61441	2.14	6.91 <sub>-7</sub>	3.65	1.70	-5.49 <sub>-7</sub>	3.61	1.69
13	131073	2.13	1.88 <sub>-7</sub>	3.68	1.73	-1.51 <sub>-7</sub>	3.64	1.71
14	278529	2.12	5.09 <sub>-8</sub>	3.69	1.74	-4.12 <sub>-8</sub>	3.67	1.73

erwartende Verhalten erster Ordnung ist. Dieser Effekt ist vergleichbar mit der Shortley–Weller–Diskretisierung bei vollen Gittern<sup>15</sup>. Auch dort ist lediglich eine Konsistenzordnung von 1 nachzuweisen (sowohl in der Theorie als auch in der Praxis), dennoch ist die Konvergenzordnung (wiederum sowohl in der Theorie als auch in der Praxis) von zweiter Ordnung, die erste Ordnung wird also “geliftet”. Abbildung 45 zeigt die erzielten Ergebnisse wieder in graphischer Form. Tabelle 31 enthält die zur Berechnung mittels

<sup>15</sup>Dabei ist klar, daß Neumann–Randbedingungen etwas anderes darstellen als die Shortley–Weller–Diskretisierung, es geht hier auch nur um einen Effekt, der beiden gemeinsam ist.

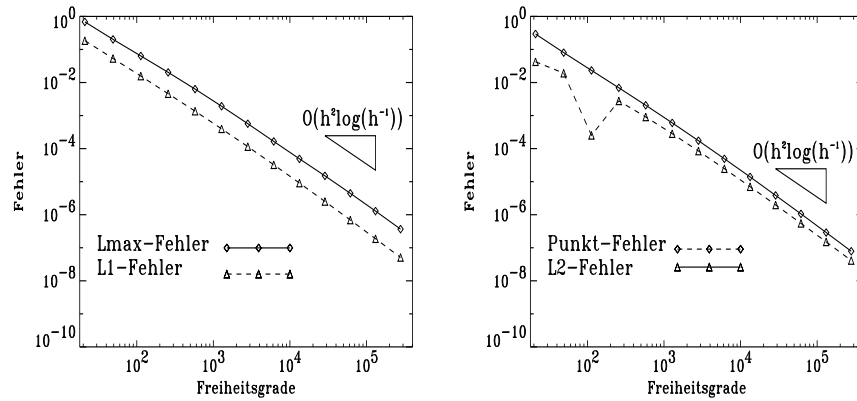


Abbildung 45: Konvergenzverhalten der jeweiligen Fehler für das Laplace-Problem L5:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)

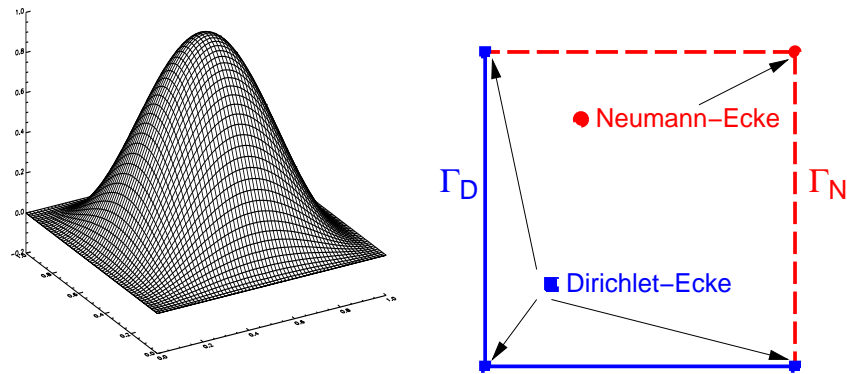


Abbildung 46: Lösung des Laplace-Problems L5 (links) sowie Randbedingungen des Problems (rechts)

eines Jacobi-vorkonditionierten BiCGSTAB-Verfahrens notwendigen Iterationszahlen. Wie schon beim Problem L1 bewirkt auch hier der (Neumann)-Vorkonditionierer eine bedeutende Reduktion der nötigen Iterationszahlen. Aufgrund der Tatsache, daß die Iterationszahlen ohne Vorkonditionierung stark ansteigen, verzichten wir im Gegensatz zur Problemstellung L1 auf einen Vergleich der Iterationszahlen zwischen nicht-vorkonditioniertem und vorkonditioniertem iterativen Löser.

Tabelle 31: Iterationszahlen für precBiCGSTAB für das Problem L5

Level	8	9	10	11	12	13	14
precBiCGSTAB	138	199	285	435	1035	1223	1825

## 7.2.2 Helmholtz–Gleichung

Wir wollen uns hier mit der Lösung der Helmholtz–Gleichung

$$\Delta u - c \cdot u = f \quad (7.3)$$

auf  $\Omega$  mit  $c > 0$  zusammen mit vorgegebenen Randbedingungen befassen.

**Beispiel H1:** Es seien

$$u(x, y) = e^{x+y} \quad \text{und} \quad c(x, y) = \frac{y}{x + 0.1}$$

sowie die rechte Seite  $f(x, y)$  entsprechend gewählt. Zudem seien Dirichlet–Randbedingungen auf dem Grundgebiet  $\bar{\Omega} = [0, 1]^2$  vorgegeben. Da die vorgegebene Problemstellung glatte Daten aufweist, rechnen wir nachfolgend auf regulären dünnen Gittern. Die numerischen Ergebnisse befinden sich in den Tabellen 32(a), 32(b) und 32(c).

Tabelle 32: (a) Fehler für das Problem H1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	1.81 <sub>-3</sub>	-	-	9.96 <sub>-4</sub>	-	-
3	49	2.33	4.56 <sub>-4</sub>	3.97	1.70	2.55 <sub>-4</sub>	3.91	1.68
4	113	2.31	1.10 <sub>-4</sub>	4.15	1.80	6.16 <sub>-5</sub>	4.14	1.79
5	257	2.27	2.59 <sub>-5</sub>	4.25	1.87	1.45 <sub>-5</sub>	4.25	1.87
6	577	2.25	6.05 <sub>-6</sub>	4.28	1.90	3.36 <sub>-6</sub>	4.32	1.92
7	1281	2.22	1.39 <sub>-6</sub>	4.35	1.96	7.70 <sub>-7</sub>	4.36	1.96
8	2817	2.20	3.16 <sub>-7</sub>	4.40	2.00	1.75 <sub>-7</sub>	4.40	2.00
9	6145	2.18	7.09 <sub>-8</sub>	4.46	2.05	3.92 <sub>-8</sub>	4.46	2.05
10	13313	2.17	1.57 <sub>-8</sub>	4.52	2.08	8.69 <sub>-9</sub>	4.51	2.08
11	28673	2.15	3.42 <sub>-9</sub>	4.59	2.13	1.89 <sub>-9</sub>	4.60	2.14
12	61441	2.14	7.12 <sub>-10</sub>	4.80	2.24	3.96 <sub>-10</sub>	4.77	2.23
13	131073	2.13	1.89 <sub>-10</sub>	3.77	1.77	1.00 <sub>-10</sub>	3.96	1.86
14	278529	2.12	5.24 <sub>-10</sub>	0.36	0.17	1.05 <sub>-10</sub>	0.95	0.45

Wie man hieraus und auch aus Abbildung 47 erkennen kann, erhalten wir hier ein Superkonvergenz–Resultat. Die auftretenden Konvergenzraten sind in diesem Beispiel größer als 4, was eine höhere Ordnung als die maximal zu erwartende Ordnung  $\mathcal{O}(h^2)$  bedeutet. Dies ist bedingt durch die Tensorprodukt–Darstellung der Lösungsfunktion. Berechnet man ein Poisson–Problem (7.2), wobei ebenfalls die Funktion  $u$  dieses Helmholtz–Beispiels die exakte Lösung ist, so ist auch hierfür ein Superkonvergenz–Verhalten zu beobachten. Ab Level 13 und insbesondere bei den Ergebnissen für Level 14 erkennt man, daß Rundungsfehler keine weitere Verbesserung der Genauigkeit der numerischen Lösung erlauben.

Tabelle 32: (b) Fehler für das Problem H1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
2	21	-	1.81 <sub>-3</sub>	-	-	8.68 <sub>-4</sub>	-	-
3	49	2.33	4.56 <sub>-4</sub>	3.97	1.70	2.43 <sub>-4</sub>	3.57	1.53
4	113	2.31	1.11 <sub>-4</sub>	4.11	1.78	6.07 <sub>-5</sub>	4.00	1.73
5	257	2.27	2.62 <sub>-5</sub>	4.24	1.87	1.44 <sub>-5</sub>	4.22	1.86
6	577	2.25	6.08 <sub>-6</sub>	4.31	1.92	3.35 <sub>-6</sub>	4.30	1.91
7	1281	2.22	1.39 <sub>-6</sub>	4.37	1.97	7.70 <sub>-7</sub>	4.35	1.96
8	2817	2.20	3.16 <sub>-7</sub>	4.40	2.00	1.75 <sub>-7</sub>	4.40	2.00
9	6145	2.18	7.09 <sub>-8</sub>	4.46	2.05	3.92 <sub>-8</sub>	4.46	2.05
10	13313	2.17	1.57 <sub>-8</sub>	4.52	2.08	8.69 <sub>-9</sub>	4.51	2.08
11	28673	2.15	3.42 <sub>-9</sub>	4.59	2.13	1.89 <sub>-9</sub>	4.60	2.14
12	61441	2.14	7.12 <sub>-10</sub>	4.80	2.24	3.96 <sub>-10</sub>	4.77	2.23
13	131073	2.13	1.89 <sub>-10</sub>	3.77	1.77	1.00 <sub>-10</sub>	3.96	1.86
14	278529	2.12	5.24 <sub>-10</sub>	0.36	0.17	1.05 <sub>-10</sub>	0.95	0.45

Tabelle 32: (c) Fehler für das Problem H1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	6.99 <sub>-4</sub>	-	-	-1.03 <sub>-3</sub>	-	-
3	49	2.33	2.02 <sub>-4</sub>	3.46	1.48	-2.81 <sub>-4</sub>	3.67	1.58
4	113	2.31	5.09 <sub>-5</sub>	3.97	1.72	-6.93 <sub>-5</sub>	4.05	1.75
5	257	2.27	1.22 <sub>-5</sub>	4.17	1.84	-1.63 <sub>-5</sub>	4.25	1.87
6	577	2.25	2.84 <sub>-6</sub>	4.30	1.91	-3.78 <sub>-6</sub>	4.31	1.92
7	1281	2.22	6.53 <sub>-7</sub>	4.35	1.96	-8.65 <sub>-7</sub>	4.37	1.97
8	2817	2.20	1.48 <sub>-7</sub>	4.41	2.00	-1.96 <sub>-7</sub>	4.41	2.00
9	6145	2.18	3.33 <sub>-8</sub>	4.44	2.04	-4.40 <sub>-8</sub>	4.45	2.04
10	13313	2.17	7.39 <sub>-9</sub>	4.51	2.08	-9.73 <sub>-9</sub>	4.52	2.08
11	28673	2.15	1.61 <sub>-9</sub>	4.59	2.13	-2.12 <sub>-9</sub>	4.59	2.13
12	61441	2.14	3.38 <sub>-10</sub>	4.76	2.22	-4.42 <sub>-10</sub>	4.80	2.25
13	131073	2.13	8.43 <sub>-11</sub>	4.01	1.88	-1.00 <sub>-10</sub>	4.42	2.08
14	278529	2.12	1.15 <sub>-12</sub>	73.30	34.58	-2.33 <sub>-11</sub>	4.29	2.02

Ebenso erkennt man aus den Resultaten, daß auch in diesem Beispiel die  $L_\infty$ -Norm und die  $L_2$ -Norm auf dünnen Gitter bzw. auf vollen Gittern (nach Projektion der Dünngitter-Funktion auf ein volles Gitter) bis auf die ersten Level identische Resultate liefern (wiederum ab Level 7 wie bereits beim Laplace-Problem L1). Damit zeigt sich auch hier, daß die Berechnung der  $L_\infty$ -Norm und der  $L_2$ -Norm auf dünnen Gittern jeweils eine vernünftige Näherung an das jeweilige kontinuierliche Analogon ist. Die numerischen Ergebnisse wurden wiederum mit einem Jacobi-vorkonditionierten BiCG-STAB gelöst, wobei lediglich der Laplace-Anteil bei der Vorkonditionierung berücksichtigt wurde. Bei Berücksichtigung auch des Helmholtz-Anteils er-



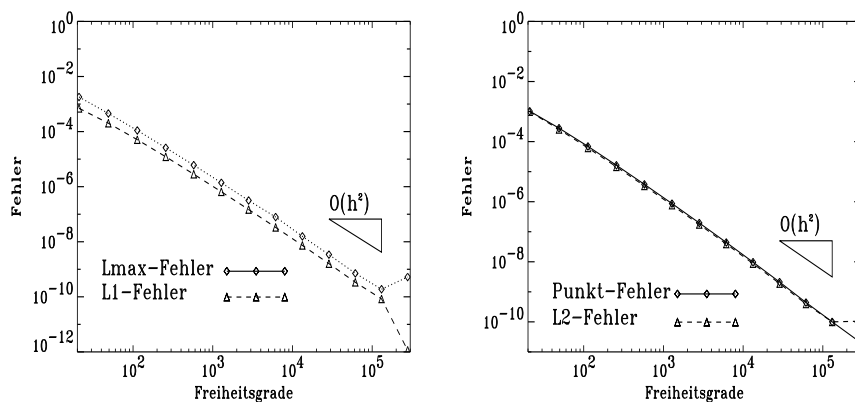


Abbildung 47: Konvergenzverhalten der jeweiligen Fehler für das Helmholtz-Problem H1:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)

geben sich lediglich geringfügige Unterschiede in den Iterationszahlen. Die so erzielten Iterationszahlen sind in Tabelle 33 nachzulesen. Auch hier zeigt sich wie schon zuvor beim Laplace-Problem L1, daß die Zahl der benötigten Iterationen moderat ansteigt.

Tabelle 33: Iterationszahlen für precBiCGSTAB für das Problem H1

Level	3	4	5	6	7	8	9	10	11
precBiCGSTAB	13	24	41	65	112	194	269	729	981

**Beispiel H2:** Es seien

$$u(x, y) = 16x(1-x)y(1-y) \left( \frac{1}{2} + \frac{\arctan[a \cdot (r^2 - (x-x_0)^2 - (y-y_0)^2)]}{\pi} \right)$$

und

$$c(x, y) = 16x(1-x)y(1-y)$$

die exakte Lösung sowie der Helmholtz-Term. Die rechte Seite  $f(x, y)$  wird entsprechend so gewählt, daß  $u(x, y)$  die Lösung der Problemstellung ist. Zudem seien wie zuvor Dirichlet-Randbedingungen auf dem Grundgebiet  $\bar{\Omega} = [0, 1]^2$  vorgegeben. Als Parameter wählen wir  $a = -20$ ,  $r = \frac{1}{4}$  und  $x_0 = y_0 = \frac{1}{2}$ . Abbildung 48 zeigt die Lösung  $u(x, y)$  sowie einen Konturplot der Lösungsfunktion. Die Lösungsfunktion weist eine starke Änderung des Gradienten auf einer Kreislinie im Grundgebiet auf, wo man eine kreisförmige innere Grenzschicht erhält. Im Fall  $a > 0$  erhält man Funktionen, die für wachsendes  $a$  immer stärkere Gradienten und damit immer stärkere Grenzschichten aufbauen, im Fall  $a < 0$  beobachtet man für fallendes  $a$  ebenfalls

eine immer stärker werdende Grenzschicht. Im Gegensatz zum Fall für  $a > 0$ , bei der die Funktion unabhängig von dem Parameter  $a$  am Punkt  $(\frac{1}{2}, \frac{1}{2})$  ihr Maximum besitzt und sich in Abhängigkeit aller beteiligter Parameter eine Grenzschicht ausbildet, tritt im Fall  $a < 0$  zudem das Problem auf, daß sich eine Grenzschicht 'von außen' und 'von innen' am Kreisring auftritt, die von den Parametern  $r$ ,  $x_0$  und  $y_0$  anhängig ist.

Dies macht den Einsatz adaptiver Gitter sinnvoll und auch notwendig, um vernünftige Approximationen an die exakte Lösung zu erhalten. Die numerischen Ergebnisse befinden sich in den Tabellen 34(a), 34(b) und 34(c), das Konvergenzverhalten ist in Abbildung 49 graphisch angegeben.

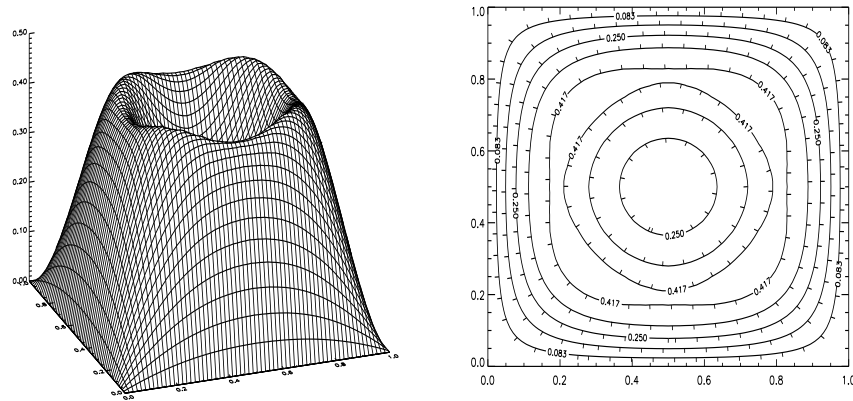


Abbildung 48: Lösungsfunktion  $u(x, y)$  des Helmholtzproblems H2 zusammen mit einem Plot der Konturlinien

Tabelle 34: (a) Fehler für das Problem H2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	9	-	6.32 <sub>-1</sub>	-	-	3.11 <sub>-1</sub>	-	-
3	33	3.67	1.05 <sub>-1</sub>	6.02	1.64	6.97 <sub>-2</sub>	4.46	1.22
4	69	2.09	5.28 <sub>-1</sub>	0.20	0.10	3.53 <sub>-1</sub>	0.20	0.10
5	209	3.03	2.08 <sub>-2</sub>	25.34	8.36	7.24 <sub>-3</sub>	48.76	16.09
6	449	2.15	3.75 <sub>-2</sub>	0.55	0.26	2.23 <sub>-2</sub>	0.32	0.15
7	1049	2.34	4.01 <sub>-3</sub>	9.35	4.00	1.78 <sub>-3</sub>	12.53	5.35
8	2453	2.34	1.57 <sub>-3</sub>	2.55	1.09	5.44 <sub>-4</sub>	3.27	1.40
9	5465	2.23	4.11 <sub>-4</sub>	3.82	1.71	1.77 <sub>-4</sub>	3.07	1.38
10	12173	2.23	1.27 <sub>-4</sub>	3.24	1.45	5.41 <sub>-5</sub>	3.27	1.47

Tabelle 34: (b) Fehler für das Problem H2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
2	9	-	6.32 <sub>-1</sub>	-	-	2.79 <sub>-1</sub>	-	-
3	33	3.67	1.05 <sub>-1</sub>	6.02	1.64	6.21 <sub>-2</sub>	4.49	1.22
4	93	2.09	5.40 <sub>-1</sub>	0.19	0.09	3.36 <sub>-1</sub>	0.18	0.09
5	209	3.03	2.08 <sub>-2</sub>	25.96	8.57	7.05 <sub>-3</sub>	47.66	15.73
6	449	2.15	3.75 <sub>-2</sub>	0.55	0.26	2.20 <sub>-2</sub>	0.32	0.15
7	1049	2.34	4.08 <sub>-3</sub>	9.19	3.93	1.74 <sub>-3</sub>	12.64	5.40
8	2453	2.34	1.57 <sub>-3</sub>	2.60	1.11	5.03 <sub>-4</sub>	3.46	1.48
9	5465	2.23	4.22 <sub>-4</sub>	3.72	1.67	1.77 <sub>-4</sub>	2.84	1.27
10	12173	2.23	1.27 <sub>-4</sub>	3.32	1.49	5.40 <sub>-5</sub>	3.28	1.47

Tabelle 34: (c) Fehler für das Problem H2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	9	-	2.52 <sub>-1</sub>	-	-	-6.21 <sub>-1</sub>	-	-
3	33	3.67	5.30 <sub>-2</sub>	4.75	1.29	-9.41 <sub>-2</sub>	6.60	1.80
4	69	2.09	2.85 <sub>-1</sub>	0.19	0.09	-4.92 <sub>-1</sub>	0.19	0.09
5	209	3.03	5.54 <sub>-3</sub>	51.44	16.98	-4.04 <sub>-3</sub>	121.78	40.19
6	449	2.15	1.82 <sub>-2</sub>	0.30	0.14	-3.60 <sub>-2</sub>	0.11	0.05
7	1049	2.34	1.31 <sub>-3</sub>	13.89	5.94	-9.49 <sub>-4</sub>	37.93	16.21
8	2453	2.34	4.17 <sub>-4</sub>	3.14	1.34	-2.71 <sub>-4</sub>	3.50	1.50
9	5465	2.23	1.30 <sub>-4</sub>	3.21	1.44	-4.20 <sub>-5</sub>	6.45	2.89
10	12173	2.23	4.01 <sub>-5</sub>	3.24	1.45	-4.34 <sub>-6</sub>	9.68	4.34

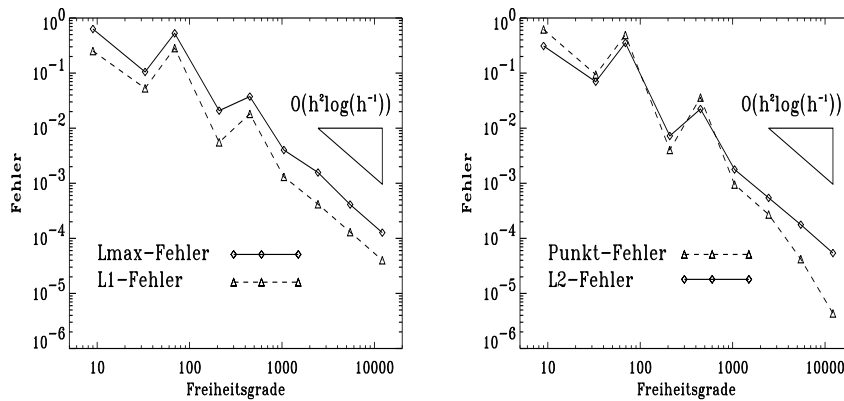


Abbildung 49: Konvergenzverhalten der jeweiligen Fehler für das Helmholtz-Problem H2 (adaptive dünne Gitter)

### 7.2.3 Allgemeine elliptische Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten

In diesem Abschnitt beschäftigen wir uns mit der numerischen Lösung von allgemeinen elliptischen Differentialgleichungen 2. Ordnung

$$-\varepsilon \Delta u + \vec{b} \cdot \nabla u + cu = f \quad (7.4)$$

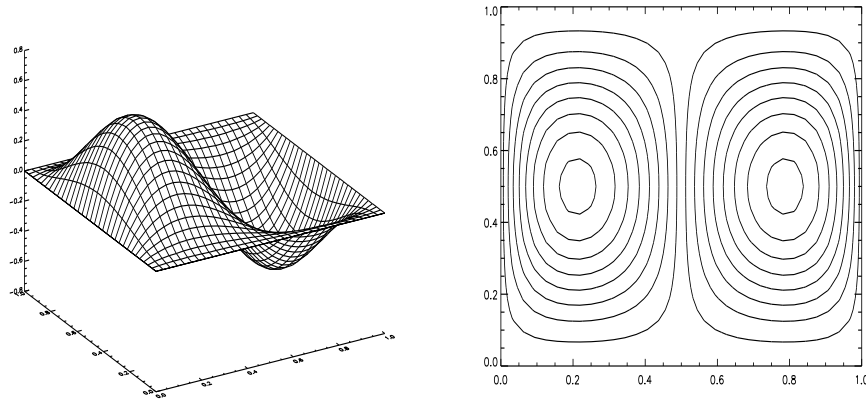


Abbildung 50: exakte Lösung des Beispiels K1 und Konturlinien der Lösung

auf  $\bar{\Omega}^d = [0, 1]^d$  zusammen mit entsprechenden Randbedingungen sowie konstanten Koeffizientenfunktionen. Als Beispiele für diesen Typ von Differentialgleichung betrachten wir zwei der drei Beispiele, die in [JMT97] als Testbeispiele angegeben werden, wobei das Beispiel K1 glatte Daten besitzt und folglich keine Adaptivität benötigt. Das Beispiel K2 besitzt eine parameterabhängige innere Grenzschicht, bei der in Abhängigkeit der Parameter Adaptivität notwendig werden kann.

**Beispiel K1:** Seien in der Gleichung (7.4)  $\vec{b} = (3, 2)^T$ ,  $c = 2$  und  $\varepsilon = 1$  sowie die Dirichlet-Randbedingungen und die rechte Seite  $f$  so gewählt, daß die Funktion

$$u(x, y) = 100(1 - x)^2 x^2 y(1 - 2y)(1 - y)$$

die exakte Lösung des gestellten Problems in zwei Dimensionen ist. Das Beispiel K1 entspricht somit dem Testbeispiel 1 in [JMT97] und besitzt sowohl eine glatte Lösung als auch eine glatte rechte Seite. In Abbildung 50 sind die Funktion sowie eine Darstellung der Höhenlinien zu sehen. Die Tabellen 35(a), 35(b) und 35(c) zeigen die numerischen Ergebnisse auf regulären dünnen Gittern. Man beachte, daß zur Diskretisierung des konvektiven Anteils in diesem Beispiel die zentrale Differenz genommen wurde. Die Konvergenzordnung ist  $\mathcal{O}(h^2 \log(h^{-1}))$ , und die erreichten Raten in der Fehlerreduktion entsprechen somit den Reduktionsraten auf dünnen Gittern, wie sie beispielsweise auch für numerische Berechnungen mittels Finiter Elemente [Gri91a, Gri91b, BaZe92, Bun92, Bal94, BDZ96, BuDo97, Dor97, Bun98] zu beobachten sind. Ein Blick auf die Spalten  $q_1^S$ ,  $q_2^S$ ,  $q_1^V$ ,  $q_2^V$ ,  $q_3^V$  und  $q_3^S$  enthüllt, daß auch hier die dünnen Gitter einem Verfahren höherer Ordnung entsprechen, da die Faktoren in den genannten Spalten deutlich größer als 1 sind.

Tabelle 35: (a) Fehler für Beispiel K1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	1.76 <sub>-1</sub>	-	-	6.93 <sub>-2</sub>	-	-
3	49	2.33	5.98 <sub>-2</sub>	2.94	1.26	2.71 <sub>-2</sub>	2.56	1.10
4	113	2.31	2.16 <sub>-2</sub>	2.77	1.20	9.39 <sub>-3</sub>	2.88	1.24
5	257	2.27	7.08 <sub>-3</sub>	3.05	1.34	3.09 <sub>-3</sub>	3.04	1.34
6	577	2.25	2.33 <sub>-3</sub>	3.04	1.35	9.73 <sub>-4</sub>	3.17	1.41
7	1281	2.22	7.05 <sub>-4</sub>	3.30	1.49	2.96 <sub>-4</sub>	3.29	1.48
8	2817	2.20	2.09 <sub>-4</sub>	3.38	1.54	8.74 <sub>-5</sub>	3.38	1.54
9	6145	2.18	6.02 <sub>-5</sub>	3.47	1.59	2.53 <sub>-5</sub>	3.46	1.59
10	13313	2.17	1.70 <sub>-5</sub>	3.53	1.63	7.17 <sub>-6</sub>	3.52	1.62
11	28673	2.15	4.76 <sub>-6</sub>	3.57	1.66	2.01 <sub>-6</sub>	3.57	1.66
12	61441	2.14	1.32 <sub>-6</sub>	3.61	1.69	5.55 <sub>-7</sub>	3.62	1.69
13	131073	2.13	3.60 <sub>-7</sub>	3.67	1.72	1.52 <sub>-7</sub>	3.65	1.71
14	278529	2.12	9.79 <sub>-8</sub>	3.68	1.73	4.14 <sub>-8</sub>	3.67	1.73

Tabelle 35: (b) Fehler für Beispiel K1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
2	21	-	1.76 <sub>-1</sub>	-	-	5.02 <sub>-2</sub>	-	-
3	49	2.33	5.99 <sub>-2</sub>	2.93	1.26	2.30 <sub>-2</sub>	2.19	0.94
4	113	2.31	2.17 <sub>-2</sub>	2.76	1.18	8.83 <sub>-3</sub>	2.60	1.12
5	257	2.27	7.38 <sub>-3</sub>	2.93	1.29	3.02 <sub>-3</sub>	2.92	1.26
6	577	2.25	2.33 <sub>-3</sub>	3.17	1.41	9.66 <sub>-4</sub>	3.13	1.39
7	1281	2.22	7.09 <sub>-4</sub>	3.29	1.48	2.95 <sub>-4</sub>	3.27	1.47
8	2817	2.20	2.09 <sub>-4</sub>	3.39	1.54	7.07 <sub>-5</sub>	4.18	1.90
9	6145	2.18	6.03 <sub>-5</sub>	3.47	1.59	2.53 <sub>-5</sub>	2.80	1.28
10	13313	2.17	1.71 <sub>-5</sub>	3.53	1.63	7.17 <sub>-6</sub>	3.52	1.62
11	28673	2.15	4.76 <sub>-6</sub>	3.57	1.66	2.01 <sub>-6</sub>	3.57	1.66
12	61441	2.14	1.32 <sub>-6</sub>	3.61	1.69	5.55 <sub>-7</sub>	3.62	1.69
13	131073	2.13	3.60 <sub>-7</sub>	3.67	1.72	1.52 <sub>-7</sub>	3.65	1.71
14	278529	2.12	9.79 <sub>-8</sub>	3.68	1.73	4.14 <sub>-8</sub>	3.67	1.73

**Beispiel K2:** Seien in der Gleichung (7.4)  $\vec{b} = (2, 3)^T$  und  $c = 2$  sowie die Randbedingungen und die rechte Seite  $f$  (in Abhängigkeit von  $\varepsilon$ ) so gewählt, daß die Funktion

$$u(x, y, \varepsilon) = xy^2 - y^2 \cdot \exp\left(\frac{2(x-1)}{\varepsilon}\right) - x \cdot \exp\left(\frac{3(y-1)}{\varepsilon}\right) + \exp\left(\frac{2(x-1) + 3(y-1)}{\varepsilon}\right)$$

die exakte Lösung des gestellten Problems ist. Dabei steuert der Parameter  $\varepsilon$  die Stärke der inneren Grenzschicht. In den Abbildungen 52 und 53

Tabelle 35: (c) Fehler für Beispiel K1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	$3.48_{-2}$	-	-	$-1.37_{-2}$	-	-
3	49	2.33	$1.70_{-2}$	2.06	0.88	$-2.91_{-2}$	0.47	0.20
4	113	2.31	$6.55_{-3}$	2.59	1.12	$-1.25_{-2}$	2.34	1.01
5	257	2.27	$2.25_{-3}$	2.91	1.28	$-4.53_{-3}$	2.75	1.21
6	577	2.25	$7.23_{-4}$	3.12	1.39	$-1.46_{-3}$	3.10	1.38
7	1281	2.22	$2.21_{-4}$	3.27	1.47	$-4.51_{-4}$	3.24	1.46
8	2817	2.20	$6.56_{-5}$	3.37	1.53	$-1.34_{-4}$	3.36	1.53
9	6145	2.18	$1.90_{-5}$	3.46	1.59	$-3.89_{-5}$	3.45	1.58
10	13313	2.17	$5.39_{-6}$	3.52	1.62	$-1.11_{-5}$	3.52	1.62
11	28673	2.15	$1.51_{-6}$	3.57	1.66	$-3.11_{-6}$	3.57	1.66
12	61441	2.14	$4.18_{-7}$	3.61	1.69	$-8.60_{-7}$	3.62	1.69
13	131073	2.13	$1.15_{-7}$	3.63	1.71	$-2.36_{-7}$	3.64	1.71
14	278529	2.12	$3.12_{-8}$	3.69	1.74	$-6.43_{-8}$	3.67	1.73

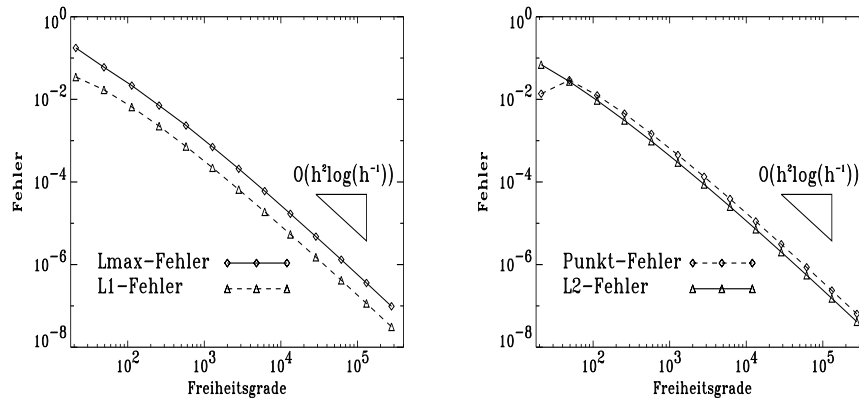


Abbildung 51: Konvergenzverhalten der jeweiligen Fehler für das Beispiel K1

sind die Funktionen für  $\varepsilon = 1$  und  $\varepsilon = 10^{-1}$  dargestellt. Im ersten Fall präsentiert sich die Funktion als glatte Funktion, bei der noch keine Grenzschicht ausgebildet ist und die aufgrund des Fehlens von singulären Stellen, Grenzschichten und anderen Problematiken einfach zu diskretisieren ist und einem Löser keine Probleme bereiten sollte. Dementsprechend erhalten wir näherungsweise dasselbe Verhalten der Fehler auf den einzelnen Levels im Fall von regulären dünnen Gitter wie schon für das Problem K1 und lassen daher die entsprechenden Tabellen für dieses Problem unberücksichtigt.

Für  $\varepsilon = 10^{-1}$  besitzt die Funktion für  $x = 1$  und  $y = 1$  typische Randschichten (boundary layers). Da diese Randschichten parallel zum Rand

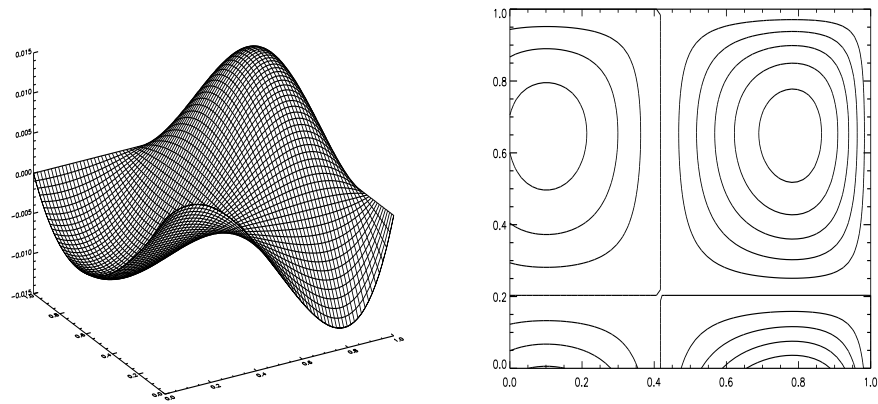
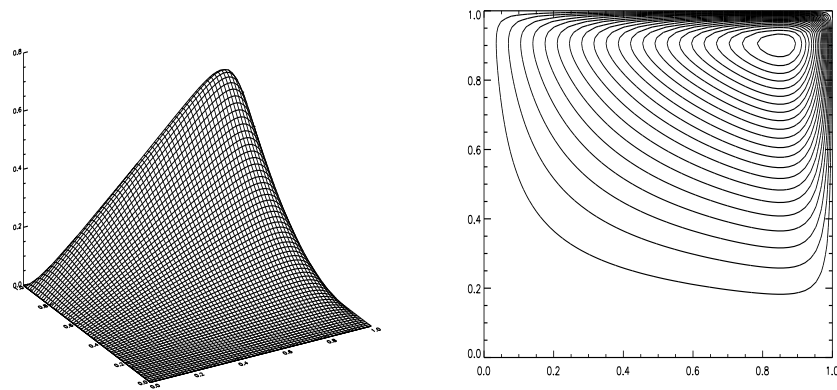
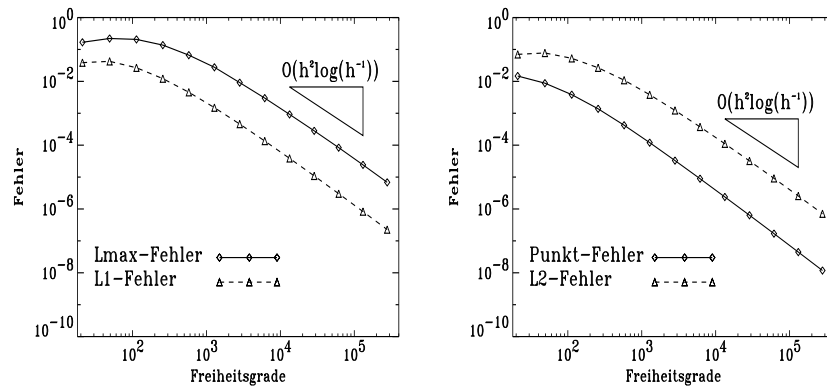
des Gebiets liegen, können diese bereits von einem regulären dünnen Gitter genügender (und in der Praxis auch berechenbarer) Leveltiefe aufgelöst und behandelt werden. Die Schwierigkeit in diesem Beispiel für dünne Gitter besteht in der Nähe der Ecke  $(1, 1)$ , bei der die Randschicht einen “Knick” macht. Dort fehlt bis zu einem gewissen, von  $\varepsilon$  abhängigen Level die notwendige Anzahl von Gitterpunkten. Bei der Wahl von  $\varepsilon = 10^{-1}$  zeigen sich auf den ersten Leveln regulärer dünner Gitter gewisse Probleme bei der Fehlerreduktion, die sich auf die soeben genannten Phänomene zurückführen lassen.

Tabelle 36: (a) Fehler für Problem K2,  $\varepsilon = 10^{-1}$ , reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	1.68 <sub>-1</sub>	-	-	7.02 <sub>-2</sub>	-	-
3	49	2.33	2.21 <sub>-1</sub>	0.76	0.33	7.77 <sub>-2</sub>	0.90	0.39
4	113	2.31	2.07 <sub>-1</sub>	1.07	0.46	5.28 <sub>-2</sub>	1.47	0.64
5	257	2.27	1.37 <sub>-1</sub>	1.51	0.67	2.67 <sub>-2</sub>	1.98	0.87
6	577	2.25	6.63 <sub>-2</sub>	2.07	0.92	1.09 <sub>-2</sub>	2.45	1.09
7	1281	2.22	2.77 <sub>-2</sub>	2.39	1.08	3.82 <sub>-3</sub>	2.62	1.18
8	2817	2.20	9.13 <sub>-3</sub>	3.03	1.38	1.23 <sub>-3</sub>	3.11	1.41
9	6145	2.18	2.99 <sub>-3</sub>	3.05	1.40	3.74 <sub>-4</sub>	3.29	1.51
10	13313	2.17	9.33 <sub>-4</sub>	3.20	1.48	1.11 <sub>-4</sub>	3.37	1.55
11	28673	2.15	2.82 <sub>-4</sub>	3.31	1.54	3.21 <sub>-5</sub>	3.46	1.61
12	61441	2.14	8.38 <sub>-5</sub>	3.37	1.57	9.14 <sub>-6</sub>	3.51	1.64
13	131073	2.13	2.41 <sub>-5</sub>	3.48	1.63	2.56 <sub>-6</sub>	3.57	1.68
14	278529	2.12	6.84 <sub>-6</sub>	3.52	1.66	7.12 <sub>-7</sub>	3.60	1.70

Tabelle 36: (b) Fehler für Problem K2,  $\varepsilon = 10^{-1}$ , reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	3.86 <sub>-2</sub>	-	-	1.47 <sub>-2</sub>	-	-
3	49	2.33	4.24 <sub>-2</sub>	0.91	0.39	8.78 <sub>-3</sub>	1.67	0.72
4	113	2.31	2.67 <sub>-2</sub>	1.59	0.69	3.84 <sub>-3</sub>	2.29	0.99
5	257	2.27	1.22 <sub>-2</sub>	2.19	0.96	1.38 <sub>-3</sub>	2.78	1.23
6	577	2.25	4.56 <sub>-3</sub>	2.68	1.19	4.22 <sub>-4</sub>	3.27	1.45
7	1281	2.22	1.50 <sub>-3</sub>	3.04	1.37	1.20 <sub>-4</sub>	3.52	1.58
8	2817	2.20	4.60 <sub>-4</sub>	3.26	1.48	3.29 <sub>-5</sub>	3.65	1.66
9	6145	2.18	1.35 <sub>-4</sub>	3.41	1.56	8.90 <sub>-6</sub>	3.70	1.70
10	13313	2.17	3.86 <sub>-5</sub>	3.50	1.61	2.38 <sub>-6</sub>	3.74	1.72
11	28673	2.15	1.09 <sub>-5</sub>	3.54	1.65	6.36 <sub>-7</sub>	3.74	1.74
12	61441	2.14	3.02 <sub>-6</sub>	3.61	1.69	1.69 <sub>-7</sub>	3.76	1.76
13	131073	2.13	8.29 <sub>-7</sub>	3.64	1.71	4.46 <sub>-8</sub>	3.79	1.78
14	278529	2.12	2.26 <sub>-7</sub>	3.67	1.73	1.18 <sub>-8</sub>	3.78	1.78

Abbildung 52: Lösung und Konturlinien des Beispiels K2 für  $\varepsilon = 1$ Abbildung 53: Lösung und Konturlinien des Beispiels K2 für  $\varepsilon = 10^{-1}$ Abbildung 54: Konvergenzverhalten beim Problem K2 auf adaptiven dünnen Gittern:  $L_{\infty}^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und Punktfehler  $T$  (rechts)



### 7.2.4 Allgemeine elliptische Differentialgleichungen 2. Ordnung mit variablen Koeffizienten

Wir wollen den Bereich der skalaren elliptischen Differentialgleichungen 2. Ordnung mit der Berechnung von Gleichungen

$$Lu = f$$

mit dem Operator

$$\tilde{L} = \sum_{i=1}^d \sum_{j=1}^d \tilde{a}_{ij}(\mathbf{x}) \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} + \sum_{i=1}^d \tilde{b}_i(\mathbf{x}) \frac{\partial}{\partial x_i} + \tilde{c}(\mathbf{x}) \quad (7.5)$$

aus (4.1) bzw. dem Operator

$$\hat{L} = \sum_{i=1}^d \sum_{j=1}^d \frac{\partial}{\partial x_i} \left( \hat{a}_{ij}(\mathbf{x}) \frac{\partial}{\partial x_j} \right) + \sum_{i=1}^d \hat{b}_i(\mathbf{x}) \frac{\partial}{\partial x_i} + \hat{c}(\mathbf{x}) \quad (7.6)$$

abschließen, wobei die beteiligten Koeffizientenfunktionen variabel seien. Dabei ist klar, wie sich die beiden Operatoren  $\tilde{L}$  und  $\hat{L}$  ineinander überführen lassen, wenn die Koeffizientenfunktionen bestimmte Glattheitsvoraussetzungen erfüllen. In diesem Fall gelten die Bedingungen

$$\begin{aligned} \tilde{a}_{ij}(\mathbf{x}) &= \hat{a}_{ij}(\mathbf{x}) \quad \text{für } 1 \leq i, j \leq d \\ \tilde{b}_i(\mathbf{x}) &= \hat{b}_i(\mathbf{x}) + \sum_{j=1}^d \frac{\partial}{\partial x_j} \hat{a}_{ij}(\mathbf{x}) \quad \text{für } 1 \leq i \leq d \\ \tilde{c}(\mathbf{x}) &= \hat{c}(\mathbf{x}) \end{aligned}$$

zur Umrechnung der Koeffizienten zwischen den beiden Operatoren.

**Beispiel V1:** Seien in zwei Dimensionen

$$\begin{aligned} \hat{a}_{11}(x, y) &= 4 + \sin(2\pi x) + \sin(2\pi y) \\ \hat{a}_{12}(x, y) &= 1 + 4y(1 - y) \sin(2\pi x) \\ \hat{a}_{21}(x, y) &= 1 + 4y(1 - y) \sin(2\pi x) \\ \hat{a}_{22}(x, y) &= 4 + \sin(2\pi xy) \\ \hat{b}_1(x, y) &= 0 \\ \hat{b}_2(x, y) &= 0 \\ \hat{c}(x, y) &= 0, \end{aligned}$$

sowie die Funktion  $u(\mathbf{x}) = u(x, y) = \sin(\pi x) \sin(\pi y)$  gegeben. Die rechte Seite  $f$  sei so gewählt, daß  $u(x, y)$  die Lösung des Problems

$$\hat{L}u = \nabla \cdot (\hat{A} \cdot \nabla)u = f$$

darstellt, wobei die Matrix  $\hat{A}$  gegeben ist vermöge

$$\hat{A} = \begin{pmatrix} \hat{a}_{11}(x, y) & \hat{a}_{12}(x, y) \\ \hat{a}_{21}(x, y) & \hat{a}_{22}(x, y) \end{pmatrix} .$$

Wie bereits für den Vergleich der Ergebnisse in mehreren Dimensionen für das Laplace–Problem L3 liegen auch hier die entsprechenden Finite Element Ergebnisse vor (vgl. [Bun98], Beispiel 6.5). Nachfolgend befinden sich die Resultate für die Berechnung mittels Finiter Differenzen auf dünnen Gittern in den Tabellen 37(a), 37(b) und 37(c), die Abbildung 55 veranschaulicht die Ergebnisse in den genannten Tabellen. Man erkennt, daß die Konvergenzgeschwindigkeit von einem logarithmischen Faktor abhängt, wie es auch bei den Finiten Elementen auf dünnen Gittern zu beobachten ist.

Tabelle 37: (a) Fehler für Problem V1, reguläres dünnes Gitter

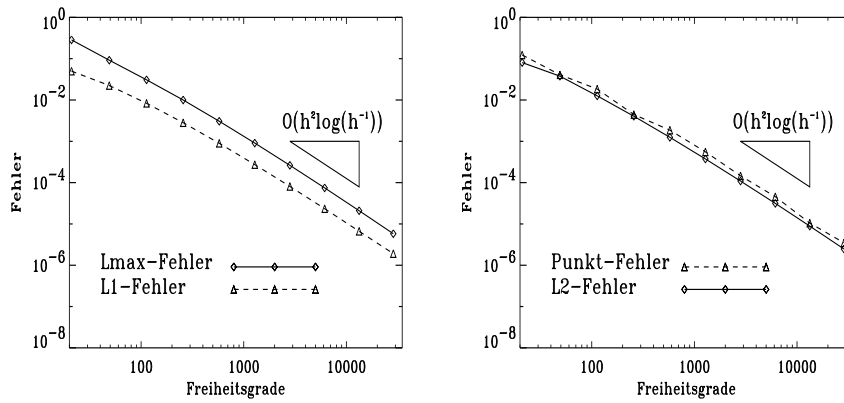
$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	2.85 <sub>-1</sub>	-	-	8.04 <sub>-2</sub>	-	-
3	49	2.33	9.18 <sub>-2</sub>	3.10	1.33	3.78 <sub>-2</sub>	2.13	0.91
4	113	2.31	3.08 <sub>-2</sub>	2.98	1.29	1.26 <sub>-2</sub>	3.00	1.30
5	257	2.27	1.00 <sub>-2</sub>	3.08	1.36	4.05 <sub>-3</sub>	3.11	1.37
6	577	2.25	3.06 <sub>-3</sub>	3.27	1.45	1.25 <sub>-3</sub>	3.24	1.44
7	1281	2.22	9.06 <sub>-4</sub>	3.38	1.48	3.74 <sub>-4</sub>	3.34	1.51
8	2817	2.20	2.62 <sub>-4</sub>	3.46	1.57	1.09 <sub>-4</sub>	3.43	1.56
9	6145	2.18	7.46 <sub>-5</sub>	3.51	1.61	3.14 <sub>-5</sub>	3.47	1.59
10	13313	2.17	2.09 <sub>-5</sub>	3.57	1.65	8.85 <sub>-6</sub>	3.55	1.64
11	28673	2.15	5.80 <sub>-6</sub>	3.60	1.67	2.47 <sub>-6</sub>	3.58	1.67

Tabelle 37: (b) Fehler für Problem V1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^V$	$q_{L_\infty^V}$	$q_1^V$	$L_2^V$	$q_{L_2^V}$	$q_2^V$
2	21	-	2.85 <sub>-1</sub>	-	-	7.83 <sub>-2</sub>	-	-
3	49	2.33	9.18 <sub>-2</sub>	3.10	1.33	3.24 <sub>-2</sub>	2.42	1.04
4	113	2.31	3.08 <sub>-2</sub>	2.98	1.29	1.15 <sub>-2</sub>	2.82	1.22
5	257	2.27	1.00 <sub>-2</sub>	3.08	1.36	3.79 <sub>-3</sub>	3.03	1.33
6	577	2.25	3.11 <sub>-3</sub>	3.22	1.43	1.19 <sub>-3</sub>	3.18	1.41
7	1281	2.22	9.27 <sub>-4</sub>	3.35	1.51	3.61 <sub>-4</sub>	3.30	1.49
8	2817	2.20	2.69 <sub>-4</sub>	3.45	1.57	1.06 <sub>-4</sub>	3.41	1.55
9	6145	2.18	7.63 <sub>-5</sub>	3.53	1.62	3.05 <sub>-5</sub>	3.48	1.60
10	13313	2.17	1.97 <sub>-5</sub>	3.87	1.78	8.36 <sub>-6</sub>	3.65	1.68
11	28673	2.15	5.50 <sub>-6</sub>	3.58	1.67	2.40 <sub>-6</sub>	3.48	1.62

Tabelle 37: (c) Fehler für Problem V1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	5.00 <sub>-2</sub>	-	-	-1.23 <sub>-1</sub>	-	-
3	49	2.33	2.24 <sub>-2</sub>	2.23	0.96	-4.08 <sub>-2</sub>	3.01	1.29
4	113	2.31	8.32 <sub>-3</sub>	2.69	1.16	-1.84 <sub>-2</sub>	2.22	0.96
5	257	2.27	2.82 <sub>-3</sub>	2.95	1.30	-4.45 <sub>-3</sub>	4.13	1.82
6	577	2.25	8.96 <sub>-4</sub>	3.15	1.40	-1.86 <sub>-3</sub>	2.39	1.06
7	1281	2.22	2.73 <sub>-4</sub>	3.28	1.48	-5.55 <sub>-4</sub>	3.35	1.51
8	2817	2.20	8.08 <sub>-5</sub>	3.38	1.54	-1.46 <sub>-4</sub>	3.80	1.73
9	6145	2.18	2.34 <sub>-5</sub>	3.45	1.58	-4.56 <sub>-5</sub>	3.20	1.47
10	13313	2.17	6.64 <sub>-6</sub>	3.52	1.62	-1.05 <sub>-5</sub>	4.34	2.00
11	28673	2.15	1.91 <sub>-6</sub>	3.48	1.62	-3.49 <sub>-6</sub>	3.01	1.40

Abbildung 55: Konvergenzverhalten der jeweiligen Fehler für das Problem V1:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)

**Beispiel V2:** Seien in drei Dimensionen die Matrix

$$\hat{A}(x, y, z) = \begin{pmatrix} x^2 e^{\pi y - 1} & x^2 & \pi^{-3} \\ 0 & \pi & 0 \\ 0 & \sin(x) & \frac{x^2 + 2y}{10} \end{pmatrix},$$

der Vektor

$$\hat{b} = \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \end{pmatrix} = \begin{pmatrix} \frac{x}{xz + 0.1} \\ \cos(e^{x+yz}) \\ xy^2 \end{pmatrix}$$

und  $\hat{c}(x, y, z) = \pi(x + y + z)$  sowie die Funktion

$$u(x, y, z) = \arctan \left( 100 \left( \frac{x + y + z}{\sqrt{3.0}} - \frac{4}{5} \right) (x - x^2)(y - y^2)(z - z^2) \right)$$

gegeben. Die rechte Seite  $f$  sei so gewählt, daß  $u(x, y, z)$  die Lösung des Dirichlet-Problems

$$\hat{L}u = \nabla \cdot (\hat{A} \cdot \nabla)u + \hat{b}\nabla u + \hat{c}u = f$$

darstellt. Damit behandeln wir in diesem Beispiel eine Problemstellung vom Typ (7.6), die aufgrund der glatten Daten in  $\hat{A}(x, y, z)$ ,  $\hat{b}(x, y, z)$  und  $\hat{c}(x, y, z)$  in die Gleichung (4.1) umgeschrieben werden kann. Die Lösungsfunktion weist am Rand des Gebiets zum Teil stärkere Gradienten auf, wie man an den Graphiken, die die Konturlinien auf verschiedenen Schnittebenen (z.B. für  $z = 0.875$ ) zeigen, erkennen kann. Nachfolgend befinden sich in den Tabellen 38(a) und 38(b) die Ergebnisse auf sukzessiven Leveln. Abbildung 56 veranschaulicht die Ergebnisse der genannten Tabellen.

Tabelle 38: (a) Fehler für Problem V2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	81	-	$7.52_{-2}$	-	-	$2.32_{-2}$	-	-
3	225	2.78	$3.82_{-2}$	1.97	0.71	$1.24_{-2}$	1.87	0.67
4	593	2.64	$2.06_{-2}$	1.85	0.70	$5.89_{-3}$	2.11	0.80
5	1505	2.54	$9.05_{-3}$	2.28	0.90	$2.43_{-3}$	2.42	0.95
6	3713	2.47	$3.70_{-3}$	2.45	0.99	$9.30_{-4}$	2.61	1.06
7	8961	2.41	$1.32_{-3}$	2.80	1.16	$3.32_{-4}$	2.80	1.16
8	21249	2.37	$4.79_{-4}$	2.76	1.16	$1.12_{-4}$	2.96	1.25
9	49665	2.34	$1.59_{-4}$	3.01	1.29	$3.67_{-5}$	3.05	1.30

Tabelle 38: (b) Fehler für Problem V2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$T$	$q_T$	$q_{3,S}$	$M$	$q_M$	$q_{4,S}$
2	81	-	$2.89_{-2}$	-	-	$3.39_{-2}$	-	-
3	225	2.78	$1.16_{-2}$	2.49	0.90	$2.53_{-2}$	1.34	0.48
4	593	2.64	$2.88_{-3}$	4.03	1.53	$8.86_{-3}$	2.86	1.08
5	1505	2.54	$4.00_{-4}$	7.20	2.83	$3.13_{-3}$	2.83	1.11
6	3713	2.47	$2.24_{-4}$	1.79	0.72	$1.19_{-3}$	2.63	1.06
7	8961	2.41	$3.59_{-4}$	0.62	0.26	$4.13_{-4}$	2.88	1.20
8	21249	2.37	$1.47_{-4}$	2.44	1.03	$1.33_{-4}$	3.11	1.31
9	49665	2.34	$4.81_{-5}$	3.06	1.31	$4.05_{-5}$	3.28	1.40

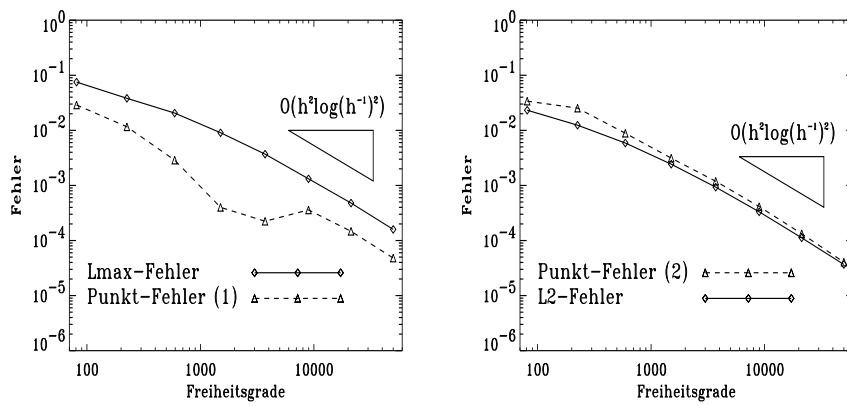


Abbildung 56: Konvergenzverhalten der jeweiligen Fehler für das Problem V2:  $L_\infty^S$  und Punktfehler  $T$  (links) sowie  $L_2^S$  und der Punktfehler  $M$  (rechts)

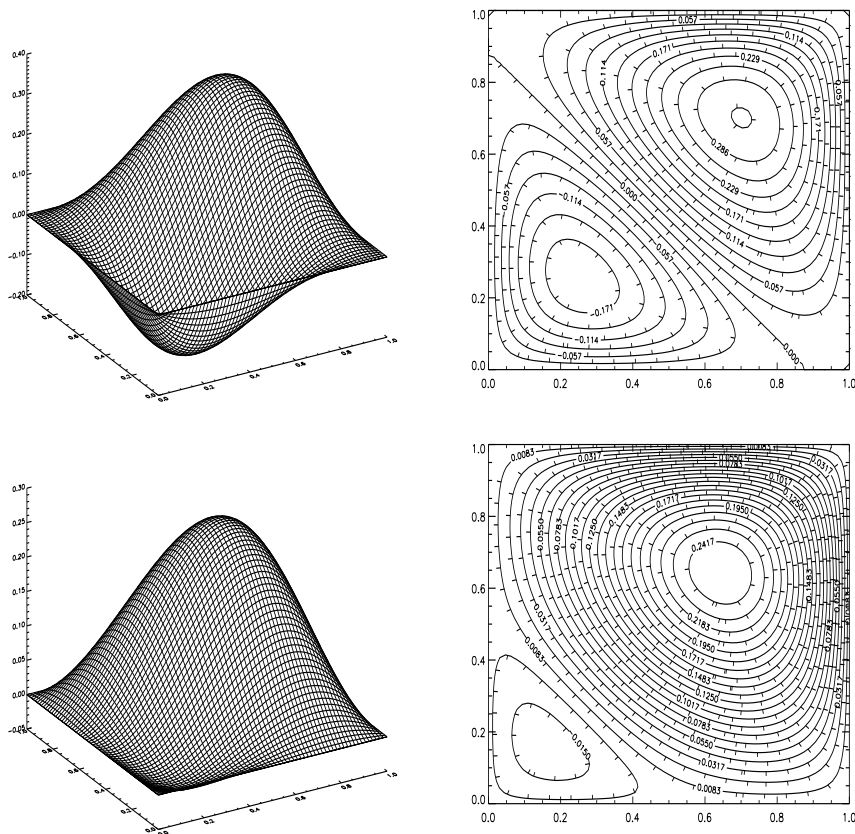


Abbildung 57: Lösungsfunktion und Konturlinien der Lösungsfunktion des Problems V2 auf den Schnittebenen  $z = 0.5$  (oben) und  $z = 0.875$  (unten)

### 7.3 Behandlung elliptischer Differentialgleichungen auf allgemeinen Gebieten

#### 7.3.1 Gebietszerlegung

In Abschnitt 4.8.1 wird das theoretische Vorgehen bei der überlappenden Schwarz-Iteration angegeben. Die dort vorgestellte Theorie wollen wir in diesem Abschnitt anhand eines numerischen Beispiels mit der Methode der Finiten Differenzen auf dünnen Gittern in der Praxis untersuchen.

**Beispiel D1:** Wir lösen das Problem (7.2) auf dem Gebiet  $\bar{\Omega}$  wie in Abbildung 58 angegeben. Die Lösung des Problems sei gegeben durch

$$u(x_1, x_2) = 4 \sin(\pi x_1) \sin(\pi x_2) \quad ,$$

die rechte Seite  $f(x_1, x_2)$  sei passend zur Problemstellung gewählt und am Rand seien Dirichlet-Randbedingungen angelegt. Dabei seien Dirichlet-Randwerte sowohl am “äußeren” Rand  $\Gamma_{D,A}$  als auch am “inneren” Rand  $\Gamma_{D,I}$  angenommen. In der Abbildung 59 wird die Zerlegung von  $\bar{\Omega}$  in die (für dieses Beispiel 4) Teilgebiete  $\bar{\Omega}_i$  (hier:  $1 \leq i \leq 4$ ) ersichtlich.

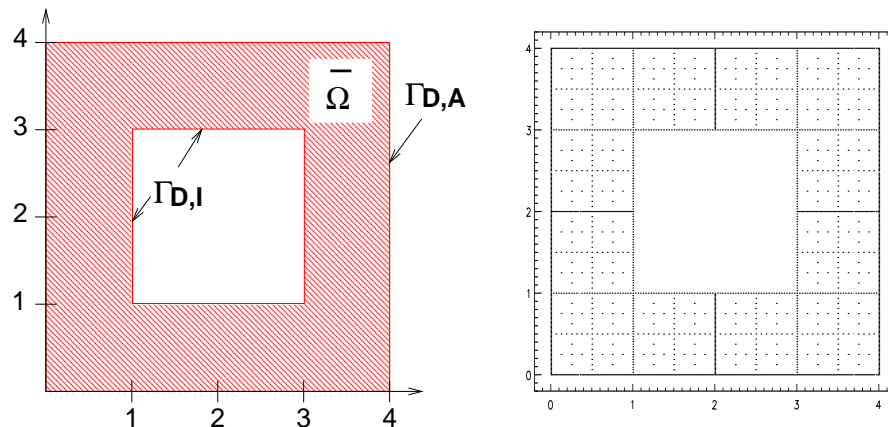


Abbildung 58: Gebiet  $\bar{\Omega}$  für das Beispiel D1 (links) sowie das Auswertungs-gitter (rechts, Level 8)

Das Vorgehen zur Lösung der Problemstellung auf dem Gebiet  $\bar{\Omega}$  ist wie folgt: auf dem Teilgebiet  $\bar{\Omega}_i$  mit  $1 \leq i \leq 4$  wird das zugehörige Teilproblem gelöst, wobei als Dirichlet-Wert von  $\bar{\Omega}_i$  der tatsächlich Dirichlet-Rand gesetzt wird, falls ein Dirichlet-Knoten des Gebiets  $\bar{\Omega}_i$  ein Dirichlet-Knoten des Gesamtgebiets  $\bar{\Omega}$  ist, und es wird ein interpolierter Wert (oder der Wert 0, falls noch nichts berechnet wurde) gesetzt, falls der betreffende Dirichlet-Knoten von  $\bar{\Omega}_i$  im Inneren von  $\bar{\Omega}$  liegt. Sobald ein Teilproblem gelöst wird, werden die Lösungen der dieses Teilgebiet überlappenden Teilgebiete aktualisiert. Die Berechnung wird abgebrochen, falls der Fehler (gemessen in der

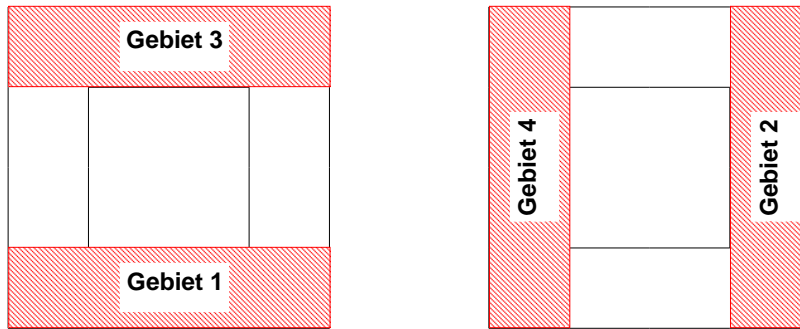
Abbildung 59: Zerlegung des Gebiets  $\bar{\Omega}$  in 4 Teilgebiete für das Beispiel D1

Tabelle 39: (a) Fehler für Problem D1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
5	224	-	9.35 <sub>-1</sub>	-	-	1.62 <sub>0</sub>	-	-
6	496	2.21	2.72 <sub>-1</sub>	3.44	1.56	5.14 <sub>-1</sub>	3.15	1.43
7	1088	2.19	9.13 <sub>-2</sub>	2.98	1.36	1.68 <sub>-1</sub>	3.06	1.40
8	2368	2.18	3.00 <sub>-2</sub>	3.04	1.39	5.31 <sub>-2</sub>	3.16	1.44
9	5120	2.16	9.40 <sub>-3</sub>	3.19	1.48	1.67 <sub>-2</sub>	3.18	1.47
10	11008	2.15	2.84 <sub>-3</sub>	3.31	1.54	5.02 <sub>-3</sub>	3.33	1.55
11	23552	2.14	8.33 <sub>-4</sub>	3.41	1.59	1.47 <sub>-3</sub>	3.41	1.60
12	50176	2.13	2.39 <sub>-4</sub>	3.49	1.64	4.20 <sub>-4</sub>	3.50	1.64
13	106496	2.12	6.75 <sub>-5</sub>	3.54	1.67	1.18 <sub>-4</sub>	3.56	1.68

Tabelle 39: (c) Fehler für Problem D1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
5	224	-	2.80 <sub>0</sub>	-	-	3.79 <sub>-1</sub>	-	-
6	496	2.21	1.26 <sub>0</sub>	2.22	1.01	1.79 <sub>-1</sub>	2.12	0.96
7	1088	2.19	4.54 <sub>-1</sub>	2.78	1.27	6.53 <sub>-2</sub>	2.74	1.25
8	2368	2.18	1.30 <sub>-1</sub>	3.49	1.60	1.83 <sub>-2</sub>	3.57	1.64
9	5120	2.16	4.70 <sub>-2</sub>	2.77	1.28	6.86 <sub>-3</sub>	2.67	1.24
10	11008	2.15	1.41 <sub>-2</sub>	3.33	1.55	2.06 <sub>-3</sub>	3.33	1.55
11	23552	2.14	4.13 <sub>-3</sub>	3.41	1.60	6.03 <sub>-4</sub>	3.42	1.60
12	50176	2.13	1.18 <sub>-3</sub>	3.50	1.64	1.73 <sub>-4</sub>	3.49	1.64
13	106496	2.12	3.34 <sub>-4</sub>	3.53	1.67	4.86 <sub>-5</sub>	3.56	1.68

Maximumsnorm und der  $L_2$ -Norm) zwischen  $k$ -ter und  $(k-1)$ -ter Iteratierter auf allen Teilgebieten kleiner als eine vorgegebene Schranke ist, die für die nachfolgenden Berechnungen auf  $10^{-10}$  gesetzt ist. Da die Lösungsfunktion der Problemstellung D1 eine glatte Funktion darstellt, kann die Berechnung auf regulären dünnen Gittern erfolgen, wobei für jedes Teilgebiet ein Gitter mit derselben Leveltiefe verwendet wird. Zudem ist es nicht erforderlich, an

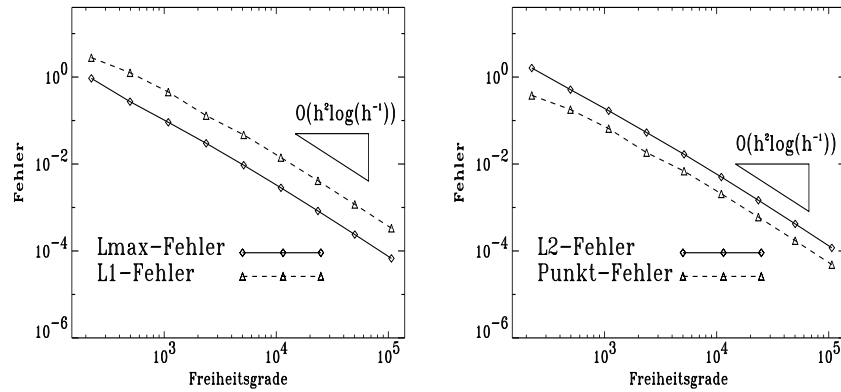


Abbildung 60: Konvergenzraten für das Beispiel D1 in der Maximumsnorm und der  $L_1$ -Norm (links) sowie der  $L_2$ -Norm und des Punktfehlers  $T$  (rechts)

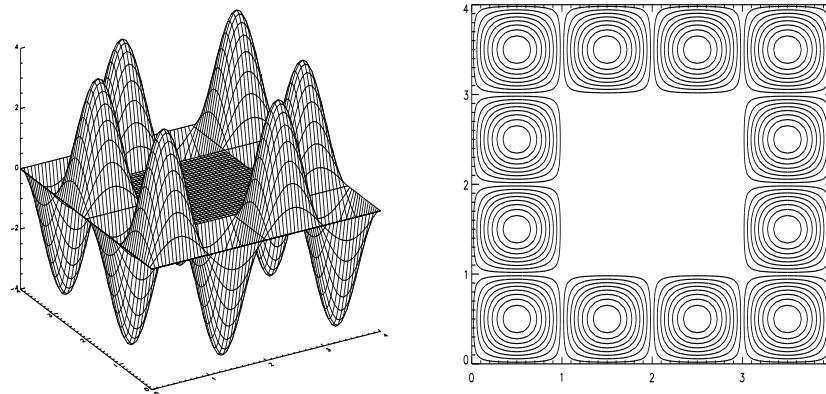


Abbildung 61: Lösung des Beispiels D1 (links) sowie Konturplot der Lösung auf  $\bar{\Omega}$  (rechts)

den Ecken feiner aufzulösen als im Inneren des Gebiets, da für dieses Beispiel keine Singularitäten in den Ecken auftreten. Zur Berechnung des Fehlers der Problemstellung auf  $\bar{\Omega}$  werden die Teilprobleme auf ein reguläres dünnes Gitter interpoliert, das zwei Level feiner ist als die regulären dünnen Gitter der einzelnen Teilgebiete. Hiermit wird sichergestellt, daß die vier einzelnen Gitter nicht zu viele Gitterpunkte an Koordinaten besitzen, die auf dem Gitter, das ausschließlich zur Berechnung des Fehlers betrachtet wird, fehlen und somit die Fehlerberechnung negativ beeinflussen würden. In diesem Beispiel reicht es aus, wenn das Auswertungsgitter für die Fehlerberechnung nur unwesentlich feiner ist als die Gitter zur Berechnung der Teilprobleme. Dies ändert sich selbstverständlich, wenn die Lösungsfunktion singuläre Stellen



oder Grenzschichten aufweist. Dort ist eine lokale Verfeinerung auch des Auswertungsgitters erforderlich, um die Fehlerberechnung nicht durch einen Fehler wegen unzureichender Interpolation zu beeinträchtigen.

Auf dem Auswertungs-Gitter, das hier zwei Level feiner als die 4 beteiligten dünnen Gitter für die Teilgebiete gewählt ist, wird nach durchgeführter Interpolation der Fehler in der Maximumsnorm, der  $L_2$ - und der  $L_1$ -Norm ermittelt sowie der Punktfehler im Punkt  $T$  berechnet. Die numerischen Ergebnisse für das Beispiel D1 finden sich in den Tabellen 39(a) und 39(b). Dabei ist zu beachten, daß aufgrund der angesprochenen Dünngitter-Interpolation, die lediglich mit einer Genauigkeit von  $\mathcal{O}(h^2 \log(h^{-1}))$  in zwei Dimensionen durchzuführen ist, auch maximal diese Ordnung in der Konvergenz zu erwarten ist. Die Konvergenzraten der einzelnen Fehler sind in Abbildung 60 graphisch aufbereitet.

Wie man aus den Tabellen erkennt, stimmt die erzielte Konvergenzrate mit der theoretisch maximal zu erwartenden Konvergenzrate überein. Verwendet man ein um vier Level feineres Auswertungsgitter, so zeigt sich, daß das Fehlerverhalten qualitativ vergleichbar ist mit dem, das auf einem um zwei Level feineren Auswertungsgitter zu erzielen ist. Bei obigem Beispiel ist der Maximumfehler sogar unabhängig von der Wahl des Auswertungsgitters, sofern das Auswertungsgitter mindestens denselben Level besitzt wie die Gitter der einzelnen Teilprobleme (ohne Beweis).

### 7.3.2 Behandlung krummlinig berandeter Gebiete

In Abschnitt 4.8.2 wird besprochen, wie Problemstellungen auf krummlinig berandeten Gebieten  $P$  behandelt werden können. In diesem Abschnitt wollen wir zum einen eine Poisson-Gleichung (Beispiel KG1) sowie eine Konvektions-Diffusion-Gleichung (Beispiel KG2) auf einem krummlinig berandeten Gebiet lösen. Nachfolgend beziehen sich die Koordinaten  $(\xi, \eta)$  stets auf das Gebiet  $P$  sowie  $(x, y)$  auf das Gebiet  $Q$ .

**Beispiel KG1:** Wir betrachten in zwei Dimensionen das Poisson-Problem (vgl. (7.2))

$$\Delta u(\xi, \eta) = f(\xi, \eta) \text{ in } (\xi, \eta) \in P \tag{7.7}$$

$$u(\xi, \eta) = g(\xi, \eta) \text{ auf } \partial P \tag{7.8}$$

auf dem Gebiet  $P$  (vgl. Abbildung 62). Dabei ist in diesem Beispiel die Transformation  $\psi : Q = [0, 1]^2 \rightarrow P$  analytisch vorgegeben. Wie in Abschnitt 4.8.2 beschrieben berechnen wir auf  $Q$  punktweise die Koeffizientenfunktionen und die rechte Seite  $f(x, y)$  der transformierten elliptischen Differentialgleichung (7.7) sowie aus (7.8) die Randbedingungen  $g(x, y)$  auf  $Q$ . Die rechte Seite

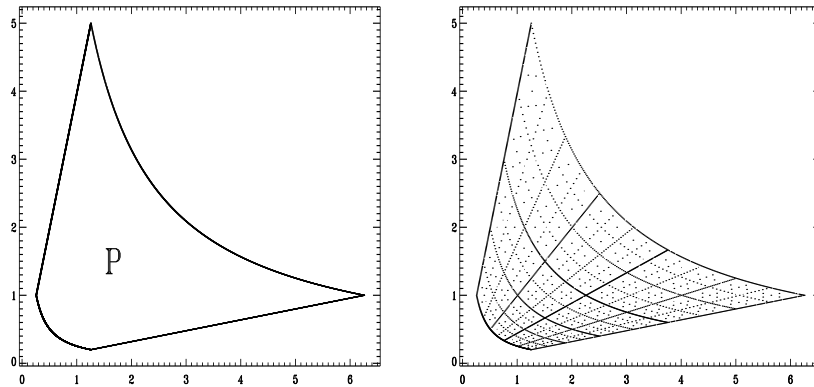


Abbildung 62: Gebiet  $P$  (links) sowie zugehöriges dünne Gitter, Level 7, (links) des Beispiels KG1

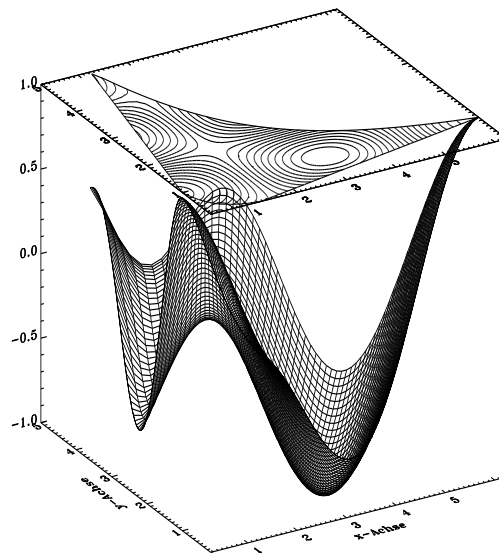


Abbildung 63: Lösungsfunktion zusammen mit Konturplot, Beispiel KG1

$f(\xi, \eta)$  auf  $P$  sowie die Randbedingungen auf  $\partial P$  seien so gegeben, daß

$$u(\xi, \eta) = \cos(\xi) \cdot \sin\left(\frac{\pi}{2}\eta\right)$$

die Lösung des Problems KG1 ist. Die Ergebnisse auf sukzessiven Leveln befinden sich in den Tabellen 40(a) und 40(b), die zugehörige graphische Darstellung findet sich in Abbildung 64. Die Lösungsfunktion zusammen mit einem Konturplot zur besseren Vorstellung der Lage der Lösungsfunktion innerhalb der Abbildung befindet sich in Abbildung 63. Aus den angesprochenen Tabellen erkennt man, daß die Konvergenzordnung wie erwartet von der Ordnung  $\mathcal{O}(h^2 \log(h^{-1}))$  ist.

Tabelle 40: (a) Fehler für Problem KG1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	2.01 <sub>-1</sub>	-	-	1.72 <sub>-1</sub>	-	-
3	49	2.33	1.32 <sub>-1</sub>	1.52	0.65	1.23 <sub>-1</sub>	1.40	0.60
4	113	2.31	7.09 <sub>-2</sub>	1.86	0.81	6.69 <sub>-2</sub>	1.84	0.80
5	257	2.27	4.71 <sub>-2</sub>	1.51	0.66	3.49 <sub>-2</sub>	1.92	0.84
6	577	2.25	2.58 <sub>-2</sub>	1.83	0.81	1.54 <sub>-2</sub>	2.27	1.01
7	1281	2.22	1.11 <sub>-2</sub>	2.32	1.05	5.83 <sub>-3</sub>	2.64	1.19
8	2817	2.20	4.18 <sub>-3</sub>	2.66	1.21	2.01 <sub>-3</sub>	2.90	1.32
9	6145	2.18	1.42 <sub>-3</sub>	2.94	1.35	6.48 <sub>-4</sub>	3.10	1.42
10	13313	2.17	4.50 <sub>-4</sub>	3.16	1.45	2.00 <sub>-4</sub>	3.24	1.49
11	28673	2.15	1.37 <sub>-4</sub>	3.28	1.53	5.95 <sub>-5</sub>	3.36	1.56

Tabelle 40: (b) Fehler für Problem KG1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	1.26 <sub>-1</sub>	-	-	9.71 <sub>-2</sub>	-	-
3	49	2.33	1.64 <sub>-1</sub>	0.77	0.33	1.22 <sub>-1</sub>	0.80	0.24
4	113	2.31	1.00 <sub>-1</sub>	1.64	0.71	5.91 <sub>-2</sub>	2.06	0.89
5	257	2.27	5.26 <sub>-2</sub>	1.90	0.84	2.64 <sub>-2</sub>	2.24	0.99
6	577	2.25	2.25 <sub>-2</sub>	2.34	1.04	9.72 <sub>-3</sub>	2.72	1.21
7	1281	2.22	8.14 <sub>-3</sub>	2.76	1.25	3.14 <sub>-3</sub>	3.10	1.39
8	2817	2.20	2.70 <sub>-3</sub>	3.03	1.38	9.29 <sub>-4</sub>	3.38	1.54
9	6145	2.18	8.52 <sub>-4</sub>	3.17	1.45	2.61 <sub>-4</sub>	3.56	1.63
10	13313	2.17	2.58 <sub>-4</sub>	3.30	1.52	7.18 <sub>-5</sub>	3.64	1.68
11	28673	2.15	7.57 <sub>-5</sub>	3.41	1.59	1.94 <sub>-5</sub>	3.70	1.72

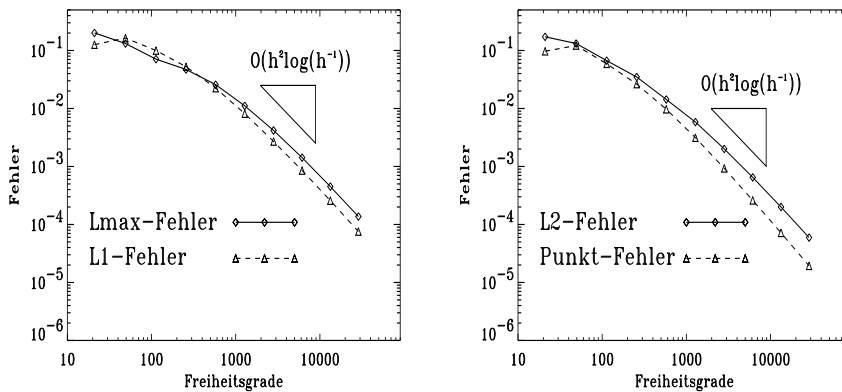


Abbildung 64: Konvergenzverhalten der jeweiligen Fehler für das Problem KG1:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)

**Beispiel KG2:** Im zweiten Beispiel zur Behandlung krummliniger Gebiete betrachten wir in zwei Dimensionen eine Konvektions–Diffusions–Gleichung (vgl. (7.5)) auf dem in Abbildung 65 gezeigten Gebiet  $P$ .

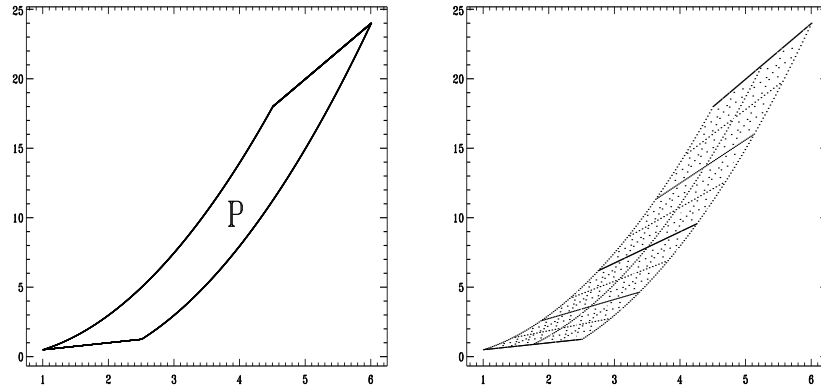


Abbildung 65: Gebiet  $P$  (links) sowie zugehöriges dünne Gitter, Level 7, (rechts) des Beispiels KG2

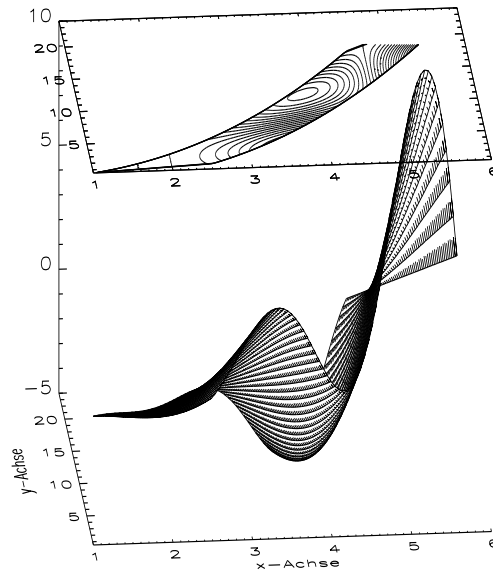


Abbildung 66: Lösungsfunktion zusammen mit Konturplot, Beispiel KG2

Die Koeffizientenfunktionen der elliptischen DGL auf  $P$  seien

$$\begin{aligned} a_{11}(\xi, \eta) &= \frac{2}{\xi}(2\eta + \xi^2) & a_{22}(\xi, \eta) &= \frac{1}{\xi^2} \left( (\eta + \xi^2)^2 + \eta^2 \right) \\ b_1(\xi, \eta) &= 0 & b_2(\xi, \eta) &= 2 \end{aligned}$$

sowie  $c(\xi, \eta) = \xi$ . Die Vorgehensweise beim Lösen ist wie zuvor beschrieben. In den Tabellen 41(a) und 41(b) sind die Ergebnisse auf den jeweiligen

Leveln aufgetragen. Die zugehörigen Graphiken, die diese Werte wieder veranschaulichen, sind in Abbildung 67 zu sehen. Deutlich erkennt man auch hier wiederum ein Konvergenzverhalten der Ordnung  $\mathcal{O}(h^2 \log(h^{-1}))$ .

Tabelle 41: (a) Fehler für Problem KG2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	1.14 <sub>0</sub>	-	-	9.81 <sub>-1</sub>	-	-
3	49	2.33	3.05 <sub>-1</sub>	3.74	1.60	2.91 <sub>-1</sub>	3.37	1.45
4	113	2.31	9.31 <sub>-2</sub>	3.28	1.42	8.53 <sub>-2</sub>	3.41	1.48
5	257	2.27	2.80 <sub>-2</sub>	3.33	1.46	2.52 <sub>-2</sub>	3.38	1.49
6	577	2.25	8.30 <sub>-3</sub>	3.37	1.50	7.47 <sub>-3</sub>	3.37	1.50
7	1281	2.22	2.43 <sub>-3</sub>	3.42	1.54	2.18 <sub>-3</sub>	3.43	1.54
8	2817	2.20	6.96 <sub>-4</sub>	3.49	1.59	6.28 <sub>-4</sub>	3.47	1.58
9	6145	2.18	1.96 <sub>-4</sub>	3.55	1.63	1.78 <sub>-4</sub>	3.61	1.66
10	13313	2.17	5.47 <sub>-5</sub>	3.58	1.65	4.99 <sub>-5</sub>	3.57	1.64
11	28673	2.15	1.51 <sub>-5</sub>	3.62	1.68	1.38 <sub>-5</sub>	3.62	1.68
12	61441	2.14	4.13 <sub>-6</sub>	3.66	1.71	3.79 <sub>-6</sub>	3.64	1.70
13	131073	2.13	1.12 <sub>-6</sub>	3.69	1.73	1.03 <sub>-6</sub>	3.68	1.73
14	278529	2.12	3.03 <sub>-7</sub>	3.70	1.74	2.79 <sub>-7</sub>	3.69	1.74

Tabelle 41: (b) Fehler für Problem KG2, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	21	-	8.62 <sub>-1</sub>	-	-	1.90 <sub>-1</sub>	-	-
3	49	2.33	3.73 <sub>-1</sub>	2.31	0.99	5.51 <sub>-2</sub>	3.45	1.48
4	113	2.31	1.29 <sub>-1</sub>	2.89	1.25	1.65 <sub>-2</sub>	3.34	1.45
5	257	2.27	4.20 <sub>-2</sub>	3.07	1.35	3.46 <sub>-3</sub>	4.77	2.10
6	577	2.25	1.29 <sub>-2</sub>	3.26	1.45	7.98 <sub>-4</sub>	4.34	1.93
7	1281	2.22	3.83 <sub>-3</sub>	3.37	1.52	1.72 <sub>-4</sub>	4.64	2.09
8	2817	2.20	1.11 <sub>-3</sub>	3.45	1.57	3.59 <sub>-5</sub>	4.79	2.18
9	6145	2.18	3.17 <sub>-4</sub>	3.50	1.61	7.18 <sub>-6</sub>	5.00	2.29
10	13313	2.17	8.89 <sub>-5</sub>	3.57	1.64	1.34 <sub>-6</sub>	5.36	2.42
11	28673	2.15	2.47 <sub>-5</sub>	3.60	1.67	2.20 <sub>-7</sub>	6.09	2.83
12	61441	2.14	6.78 <sub>-6</sub>	3.64	1.70	2.64 <sub>-8</sub>	8.33	3.89
13	131073	2.13	1.85 <sub>-6</sub>	3.66	1.72	5.39 <sub>-10</sub>	48.98	23.00
14	278529	2.12	5.01 <sub>-7</sub>	3.69	1.75	1.92 <sub>-9</sub>	0.28	0.13

### 7.3.3 Gebietszerlegung bei krummlinig berandeten Gebieten

In den beiden Abschnitten 4.8.1 und 4.8.2 wird besprochen, wie Problemstellungen zum einen mittels Gebietszerlegung und zum anderen auf krummlinig berandeten Gebieten behandelt werden können. Hier wollen wir nun beide

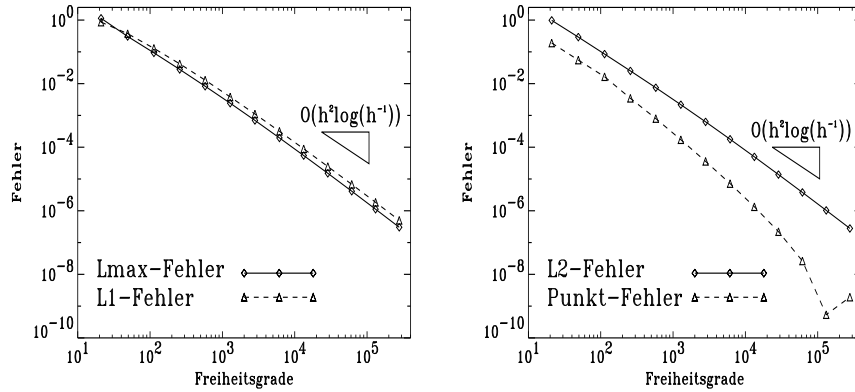


Abbildung 67: Konvergenzverhalten der jeweiligen Fehler für das Problem KG2:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)

Möglichkeiten zum Lösen partieller Differentialgleichungen kombinieren, um auf einem krummlinig berandeten L-Gebiet eine vorgegebene Problemstellung zu lösen.

**Beispiel DK1:** Sei die Konvektions-Diffusion-Gleichung (7.5) die auf dem in Abbildung 68 (links) gezeigten Gebiet (rechts in dieser Abbildung ist die Unterteilung des Gesamtgebiets in zwei Teilgebiete zu sehen) zu lösende Problemstellung. Dabei seien die Koeffizientenfunktionen wie folgt gegeben:

$$\begin{aligned}
 a_{11}(\xi, \eta) &= \frac{\xi(1+\xi)(1+\xi^2)}{\eta} & a_{12}(\xi, \eta) &= 2\xi + 1 - \xi^2 \\
 a_{21}(\xi, \eta) &= 2\xi + 1 - \xi^2 & a_{22}(\xi, \eta) &= \frac{5\xi^2\eta + 4\xi\eta + \eta}{\xi(1+\xi)} \\
 b_1(\xi, \eta) &= \frac{2\xi^2(1+\xi)}{\eta} & b_2(\xi, \eta) &= 2 \\
 c_1(\xi, \eta) &= 0
 \end{aligned}$$

Die rechte Seite sei so gewählt, daß die Funktion

$$u(\xi, \eta) = 4 \cos \left( \sqrt{\frac{\pi^2 \xi \eta}{1 + \xi}} \right) \cdot \sin \left( \sqrt{\frac{\pi^2 \eta}{\xi(1 + \xi)}} \right)$$

die Lösung der Problemstellung ist. Die Vorgehensweise zur Lösung der Problemstellung ist eine Kombination aus der Vorgehensweise für Gebietszerlegung und jeweiliger Transformation für jedes einzelne Teilgebiet. Die erzielten Ergebnisse sind den Tabellen 42(a) und 42(b) zu entnehmen, die Abbildung 70 zeigt die Ergebnisse in einer graphischen Darstellung. Wie zuvor beim Beispiel D1 als auch bei den beiden Beispielen KG1 und KG2 stellt sich auch in diesem Beispiel die Konvergenzordnung  $\mathcal{O}(h^2 \log(h^{-1}))$  ein. Die Lösungsfunktion zusammen mit einem Konturplot befindet sich in Abbildung 69.

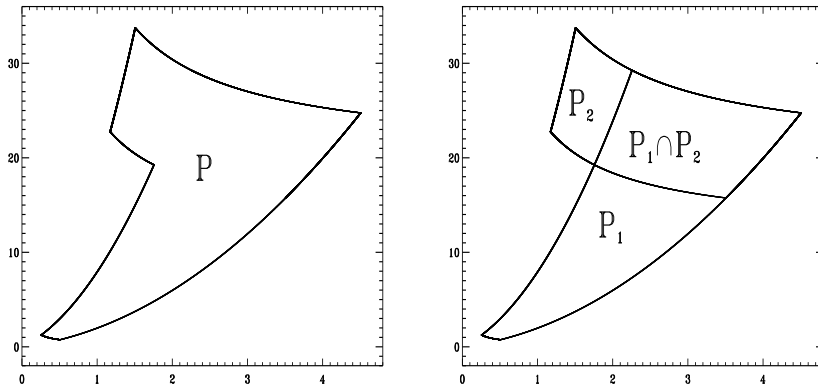


Abbildung 68: Gebiet  $P$  (links) des Beispiels DK1 sowie die Unterteilung von  $P$  in die zwei Teilgebiete  $P_1$  und  $P_2$  (rechts)

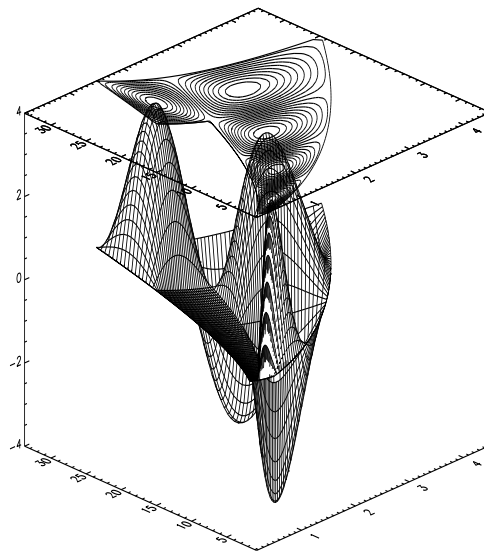


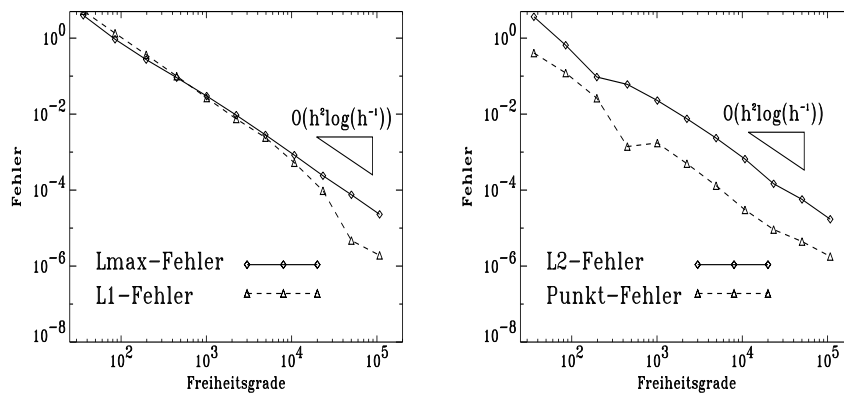
Abbildung 69: Lösungsfunktion zusammen mit Konturplot, Beispiel DK1

Tabelle 42: (a) Fehler für Problem DK1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	36	—	4.00 <sub>0</sub>	—	—	3.65 <sub>0</sub>	—	—
3	85	2.36	9.35 <sub>-1</sub>	4.28	1.81	6.51 <sub>-1</sub>	5.61	2.38
4	197	2.32	2.72 <sub>-1</sub>	3.44	1.48	9.45 <sub>-2</sub>	6.89	2.97
5	449	2.28	9.13 <sub>-2</sub>	2.98	1.31	6.05 <sub>-2</sub>	1.56	0.68
6	1009	2.25	2.99 <sub>-2</sub>	3.05	1.36	2.28 <sub>-2</sub>	2.65	1.18
7	2241	2.22	9.40 <sub>-3</sub>	3.18	1.43	7.55 <sub>-3</sub>	3.02	1.36
8	4929	2.20	2.84 <sub>-3</sub>	3.31	1.50	2.34 <sub>-3</sub>	3.23	1.47
9	10753	2.18	8.33 <sub>-4</sub>	3.41	1.56	6.56 <sub>-4</sub>	3.57	1.64
10	23297	2.17	2.39 <sub>-4</sub>	3.49	1.61	1.47 <sub>-4</sub>	4.46	2.06
11	50177	2.15	7.54 <sub>-5</sub>	3.17	1.47	5.70 <sub>-5</sub>	2.58	1.20
12	107521	2.14	2.30 <sub>-5</sub>	3.28	1.53	1.71 <sub>-5</sub>	3.33	1.56

Tabelle 42: (b) Fehler für Problem DK1, reguläres dünnes Gitter

$l$	$P$	$q_P$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
2	36	—	5.27 <sub>0</sub>	—	—	4.06 <sub>-1</sub>	—	—
3	85	2.36	1.35 <sub>0</sub>	3.90	1.65	1.21 <sub>-1</sub>	3.36	1.42
4	197	2.32	3.67 <sub>-1</sub>	3.68	1.59	2.63 <sub>-2</sub>	4.60	1.98
5	449	2.28	1.00 <sub>-1</sub>	3.67	1.61	1.40 <sub>-3</sub>	18.79	8.24
6	1009	2.25	2.66 <sub>-2</sub>	3.76	1.67	1.74 <sub>-3</sub>	0.80	0.36
7	2241	2.22	7.48 <sub>-3</sub>	3.56	1.60	4.97 <sub>-4</sub>	3.50	1.58
8	4929	2.20	2.41 <sub>-3</sub>	3.10	1.41	1.31 <sub>-4</sub>	3.79	1.72
9	10753	2.18	5.27 <sub>-4</sub>	4.57	2.10	3.02 <sub>-5</sub>	4.34	1.99
10	23297	2.17	9.63 <sub>-5</sub>	5.47	2.52	9.16 <sub>-6</sub>	3.30	1.52
11	50177	2.15	4.71 <sub>-5</sub>	2.04	0.95	4.43 <sub>-6</sub>	2.07	0.96
12	107521	2.14	1.92 <sub>-5</sub>	2.45	1.14	1.79 <sub>-6</sub>	2.47	1.15

Abbildung 70: Konvergenzverhalten der jeweiligen Fehler für das Problem DK1:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)



## 7.4 Parabolische Differentialgleichungen

Nach elliptischen Differentialgleichungen, die im vorangegangenen Abschnitt 7.2 behandelt wurden, betrachten wir in diesem Abschnitt parabolische Differentialgleichungen. Zunächst studieren wir in Abschnitt 7.4.1 skalare parabolische Differentialgleichungen anhand eines einfachen Beispiels, im zweiten Teil (Abschnitt 7.4.2) betrachten wir dann ein gekoppeltes, nichtlineares System parabolischer Differentialgleichungen. Der dritte und letzte Teil dieses Unterkapitels (Abschnitt 7.4.3) behandelt das Gray Scott Modell für Reaktion und Diffusion chemischer Spezies.

### 7.4.1 Skalare parabolische Differentialgleichungen

Wir betrachten das parabolische Anfangs–Randwertproblem

$$\frac{\partial \vec{u}(t, \mathbf{x})}{\partial t} = L\vec{u}(t, \mathbf{x}) \quad (7.9)$$

auf dem Gebiet  $\bar{\Omega}_T^{d+1} = [0, T] \times \bar{\Omega}^d$  mit den Randbedingungen  $\vec{u}(t, \mathbf{x}) = \vec{g}(t, \mathbf{x})$  für  $\mathbf{x} \in \Gamma^d = \partial\bar{\Omega}^d$  sowie der Anfangsbedingung  $\vec{u}(0, \mathbf{x}) = \vec{u}_0(\mathbf{x})$ . Die Diskretisierung zeitabhängiger Problemstellungen kann im Zusammenhang mit dünnen Gittern prinzipiell auf verschiedene Weisen erfolgen:

1. für jeden Zeitschritt bleibt das dünne Gitter (bzgl. des Orts) unverändert

Dies entspricht der Rothe–Methode (horizontale Linienmethode, vgl. [Rek82, Kač85, GrRo92]), bei der eine parabolische Randwertaufgabe mittels Zeitdiskretisierung durch eine Folge elliptischer Randwertaufgaben approximiert wird. Hiermit erhält man in der Zeit–Dimension nicht den Effekt dünner Gitter, hat somit eindimensional in Richtung der Zeit ein volles Gitter. Bei expliziter Diskretisierung der Zeit (etwa mittels explizitem Euler) muß ebenso wie auf vollen Gittern auf die Einhaltung einer CFL–Bedingung (vgl. etwa [GrRo92]) geachtet werden.

2. selbstadaptive Steuerung

Dies entspricht ebenfalls der Rothe–Methode, wobei nun von Zeitscheibe zu Zeitscheibe das jeweilige Gitter adaptiv verfeinert wird. Somit erhalten wir auch hier in Zeitrichtung keinen Dünngitter–Effekt und eine CFL–Bedingung muß auch hier berücksichtigt werden, wenn wir eine explizite Zeitdiskretisierung wählen.

Die Vorgehensweise bei der selbstadaptiven Steuerung ist wie folgt: nach dem Berechnen der neuen Iterierten wird die Gitterverfeine-

rung anhand eines zuvor ausgewählten Adaptivitätskriteriums durchgeführt. Dabei hat der Benutzer die Wahl, anhand welcher der beteiligten Größen eine adaptive Verfeinerung geschehen soll (z.B. nur anhand der  $u$ -Komponente oder nur anhand der  $v$ -Komponente, oder anhand aller Komponenten, usw.). Danach erfolgt eine Interpolation aller Größen auf das neue Gitter.

### 3. Raum-Zeit-Behandlung mittels dünnem Gitter

Dieser Ansatz wird in [ScZu98] für parabolische Differentialgleichungen knapp vorgestellt, ausführlich wird diese Methode für hyperbolische Problemstellungen in [GrZu98] untersucht. Dabei wird ein  $(d+1)$ -dimensionales dünnes Gitter verwendet, bei dem eine Dimension (z.B. die Dimension  $d+1$ ) als Zeitdimension angesehen wird. Dabei muß die Diskretisierung der Zeit durch eine Rückwärtsdifferenz erfolgen. Zudem benötigt man in Zeitrichtung eine Diskretisierung, die die stückweise konstante, hierarchische Basis (vgl. [Bal94, GrZu98]) verwendet. Für eine detaillierte Beschreibung dieses Verfahrens verweisen wir auf [GrZu98]. In der vorliegenden Arbeit wird diese Methode nicht betrachtet.

Im weiteren Verlauf wollen wir uns mit den ersten beiden angegebenen Möglichkeiten beschäftigen. Die Diskretisierung von (7.9) erfolgt dabei vermöge eines expliziten Eulerschritts

$$\frac{\partial \vec{u}(t, \mathbf{x})}{\partial t} \approx \frac{\vec{u}(t + \Delta t, \mathbf{x}) - \vec{u}(t, \mathbf{x})}{\Delta t}$$

für die Diskretisierung der Zeitableitung und der Diskretisierung (5.2) bzw. (5.3) für den elliptischen Operator  $L$  in (7.9), wobei die Zeitableitung komponentenweise zu betrachten ist.

**Beispiel P1:** Im ersten Beispiel zur Berechnung parabolischer Differentialgleichungen sei der elliptische Anteil  $L$  in (7.9) gegeben durch

$$\begin{aligned} L = & a_{11}(t, x, y) \frac{\partial^2 u(t, x_1, x_2)}{\partial x_1^2} + a_{22}(t, x, y) \frac{\partial^2 u(t, x_1, x_2)}{\partial x_2^2} + \\ & b_1(t, x, y) \frac{\partial u(t, x_1, x_2)}{\partial x_1} + b_2(t, x, y) \frac{\partial u(t, x_1, x_2)}{\partial x_2} + c(t, x, y) u(t, x_1, x_2) \end{aligned} \quad (7.10)$$

mit

$$\begin{aligned} a_{11}(t, x_1, x_2) &= 3t \\ a_{22}(t, x_1, x_2) &= t^2 x_1 x_2 \\ b_1(t, x_1, x_2) &= t x_1 x_2 \\ b_2(t, x_1, x_2) &= \frac{(x_1 - x_2)^2}{1000}, \end{aligned}$$

wobei  $c(t, x_1, x_2)$  so gewählt sei, daß

$$u(t, x_1, x_2) = \sin(2\pi x_1 - 10t) \sin(2\pi x_2 - 5t)$$

$$u : \mathbb{R}_0^+ \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

mit  $u_0(x_1, x_2) = \sin(2\pi x_1) \sin(2\pi x_2)$  die Lösung der Problemstellung ist. Die Lösung besteht demnach aus zwei in  $x$ -Richtung und in  $y$ -Richtung laufenden Sinuswellen, deren Ausbreitungsgeschwindigkeiten unterschiedlich sind. Dabei ist die Geschwindigkeit der Sinuswelle in  $x$ -Richtung doppelt so schnell wie die in  $y$ -Richtung. Nachfolgend zeigen die Abbildungen für verschiedene Zeiten  $t$  die numerisch berechneten Ergebnisse. Das hierzu verwendete Gitter ist ein reguläres dünnes Gitter vom Level 6 mit einer konstanten Zeitschrittweite von  $10^{-4}$ .

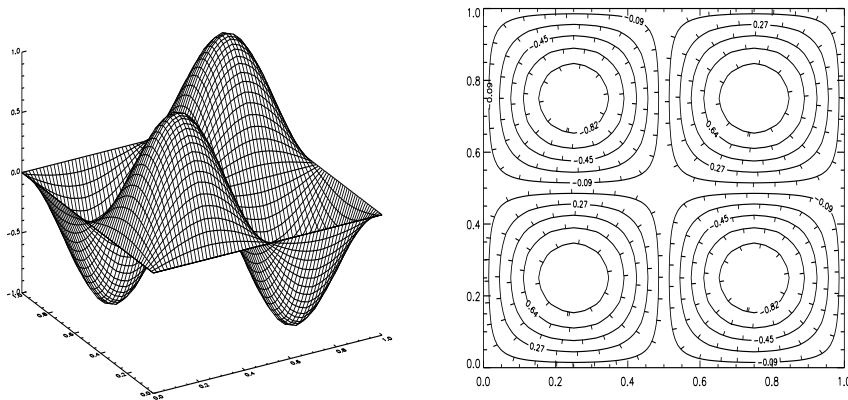


Abbildung 71: Graph und Konturlinien,  $t = 0$  (Problem P1)

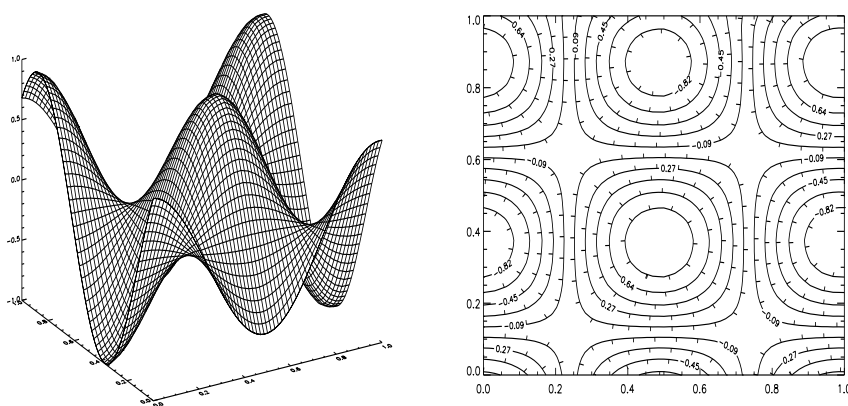
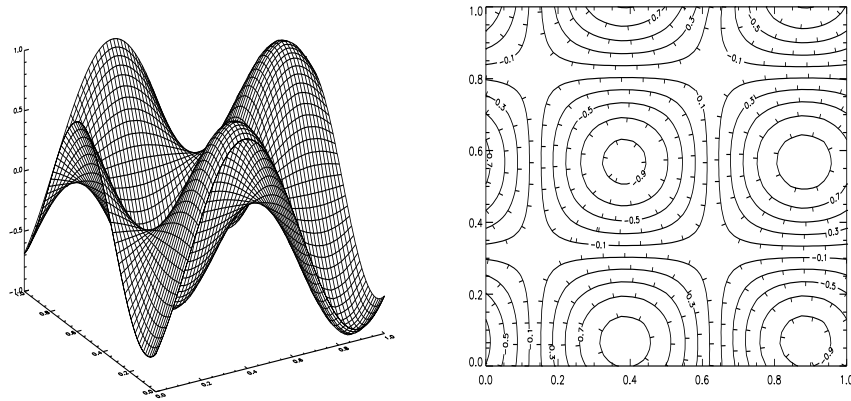


Abbildung 72: Graph und Konturlinien,  $t = 1.5_{-1}$  (Problem P1)

Abbildung 73: Graph und Konturlinien,  $t = 4.0_{-1}$  (Problem P1)

**Beispiel P2:** Im zweiten Beispiel zur Berechnung parabolischer Differentialgleichungen sei der elliptische Anteil  $L$  in (7.9) wie im Beispiel P1, Formel (7.10), gegeben. In diesem Beispiel P2 sind

$$\begin{aligned} a_{11}(t, x_1, x_2) &= \sin(xy) \\ a_{22}(t, x_1, x_2) &= \cos(xy) \\ b_1(t, x_1, x_2) &= e^{\pi xy} \\ b_2(t, x_1, x_2) &= e^{y+xy} \end{aligned}$$

sowie  $c(t, x_1, x_2)$  entsprechend gewählt, daß

$$\begin{aligned} u(t, x_1, x_2) &= x^3 + y^2 + t^2 + e^{xy} \\ u &: \mathbb{R}_0^+ \times \mathbb{R}^2 \rightarrow \mathbb{R} \end{aligned}$$

mit  $u_0(x_1, x_2) = x^3 + y^2 + e^{xy}$  die Lösung der Problemstellung ist. An diesem Beispiel wollen wir den Einfluß der Zeitschrittweite auf die Genauigkeit der zu erzielenden numerischen Lösung bei Verwendung eines festen Gitters untersuchen. Hierzu wählen wir unterschiedliche Zeitschrittweiten  $\Delta t_i$ , die sich in diesem Beispiel vermöge

$$\Delta t_i = 4^{-i} \Delta t_0$$

für  $1 \leq i \leq 9$  mit der Wahl  $\Delta t_0 = 10^{-6}$  ergeben. Mit diesen Zeitschrittweiten wird nun ausgehend von der Zeit  $t = 0$  jeweils ein Zeitschritt auf einem fest gewählten Gitter (hier: reguläres dünnes Gitter, Level 5) gerechnet. Die für jedes  $\Delta t_i$  so erzielten Fehler sind in den beiden Tabellen 43(a) und 43(b) aufgelistet. Dabei gibt die mit  $i$  bezeichnete erste Spalte den Index von  $\Delta t_i$  an und die Spalte, die mit  $q_{\Delta t_i}$  bezeichnet ist, gibt den Quotienten zweier aufeinanderfolgender Zeitschrittweiten an, der in unserem Fall hier stets 4 beträgt.

Tabelle 43: (a) Fehler für Problem P2, reguläres dünnes Gitter

$i$	$q_{\Delta t_i}$	$L_{\infty}^S$	$q_{L_{\infty}^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
0	-	2.81 <sub>-3</sub>	-	-	2.00 <sub>-3</sub>	-	-
1	4.00	7.03 <sub>-4</sub>	4.00	1.00	5.00 <sub>-4</sub>	4.00	1.00
2	4.00	1.76 <sub>-4</sub>	3.99	1.00	1.25 <sub>-4</sub>	4.00	1.00
3	4.00	4.40 <sub>-5</sub>	4.00	1.00	3.13 <sub>-5</sub>	3.99	1.00
4	4.00	1.10 <sub>-5</sub>	4.00	1.00	7.82 <sub>-6</sub>	4.00	1.00
5	4.00	2.75 <sub>-6</sub>	4.00	1.00	1.96 <sub>-6</sub>	3.99	1.00
6	4.00	6.87 <sub>-7</sub>	4.00	1.00	4.89 <sub>-7</sub>	4.01	1.00
7	4.00	1.72 <sub>-7</sub>	3.99	1.00	1.22 <sub>-7</sub>	4.01	1.00
8	4.00	4.29 <sub>-8</sub>	4.01	1.00	3.05 <sub>-8</sub>	4.00	1.00
9	4.00	1.07 <sub>-8</sub>	4.01	1.00	7.64 <sub>-9</sub>	3.99	1.00

Tabelle 43: (b) Fehler für Problem P2, reguläres dünnes Gitter

$i$	$q_{\Delta t_i}$	$L_1^V$	$q_{L_1^V}$	$q_3^V$	$T$	$q_T$	$q_3^S$
0	-	1.78 <sub>-3</sub>	-	-	5.73 <sub>-7</sub>	-	-
1	4.00	4.45 <sub>-4</sub>	4.00	1.00	1.43 <sub>-7</sub>	4.01	1.00
2	4.00	1.11 <sub>-4</sub>	4.01	1.00	3.58 <sub>-8</sub>	3.99	1.00
3	4.00	2.78 <sub>-5</sub>	3.99	1.00	8.95 <sub>-9</sub>	4.00	1.00
4	4.00	6.95 <sub>-6</sub>	4.00	1.00	2.24 <sub>-9</sub>	4.00	1.00
5	4.00	1.74 <sub>-6</sub>	3.99	1.00	5.59 <sub>-10</sub>	4.01	1.00
6	4.00	4.34 <sub>-7</sub>	4.01	1.00	1.40 <sub>-10</sub>	3.99	1.00
7	4.00	1.09 <sub>-7</sub>	3.98	0.99	3.50 <sub>-11</sub>	4.00	1.00
8	4.00	2.71 <sub>-8</sub>	4.02	1.01	8.74 <sub>-12</sub>	4.00	1.00
9	4.00	6.79 <sub>-9</sub>	3.99	1.00	2.19 <sub>-12</sub>	3.99	1.00

Aus den Tabellen erkennt man deutlich, daß die Fehlerordnung bezüglich der Zeit linear ist, da sich alle betrachteten Fehler mit demselben Faktor reduzieren wie sich die Schrittweite in der Zeit verkleinert. Dies belegen auch die Spalten, in denen die entsprechenden Quotienten zu finden sind, die für alle 4 Normen identisch 1 sind.

#### 7.4.2 Nichtlineare Systeme parabolischer Differentialgleichungen

In diesem Abschnitt betrachten wir nichtlineare Systeme parabolischer Differentialgleichungen der Bauart

$$\frac{\partial \vec{u}(t, \mathbf{x})}{\partial t} = L\vec{u}(t, \mathbf{x}) + B \cdot \vec{u}(t, \mathbf{x}) \quad (7.11)$$

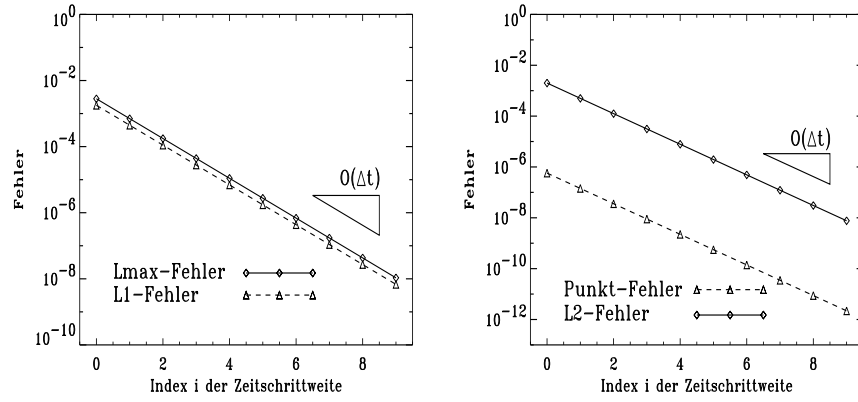


Abbildung 74: Konvergenzverhalten der jeweiligen Fehler für das Problem P2:  $L_\infty^S$  und  $L_1^V$  (links) sowie  $L_2^S$  und der Punktfehler  $T$  (rechts)

mit

$$\begin{aligned}
 B &= B \left( \frac{\partial u_1(t, \mathbf{x})}{\partial x_1}, \dots, \frac{\partial u_1(t, \mathbf{x})}{\partial x_1} \right) \\
 &= \begin{pmatrix} \frac{\partial u_1(t, \mathbf{x})}{\partial x_1} & \frac{\partial u_1(t, \mathbf{x})}{\partial x_2} & \dots & \frac{\partial u_1(t, \mathbf{x})}{\partial x_d} \\ \frac{\partial u_2(t, \mathbf{x})}{\partial x_1} & \frac{\partial u_2(t, \mathbf{x})}{\partial x_2} & \dots & \frac{\partial u_2(t, \mathbf{x})}{\partial x_d} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial u_d(t, \mathbf{x})}{\partial x_1} & \frac{\partial u_d(t, \mathbf{x})}{\partial x_2} & \dots & \frac{\partial u_d(t, \mathbf{x})}{\partial x_d} \end{pmatrix}
 \end{aligned}$$

auf dem Gebiet  $\bar{\Omega}_T^{d+1} = [0, T] \times \bar{\Omega}^d$  mit den Randbedingungen  $\vec{u}(t, \mathbf{x}) = \vec{r}(t, \mathbf{x})$  für  $\mathbf{x} \in \Gamma^d = \partial\bar{\Omega}^d$  sowie der Anfangsbedingung  $\vec{u}(0, \mathbf{x}) = \vec{u}_0(\mathbf{x})$ .

**Beispiel P3:** Wir betrachten im Beispiel P3 das nichtlineare parabolische System

$$\begin{aligned}
 \frac{\partial u(t, x_1, x_2)}{\partial t} &= a_{11}(t, x_1, x_2) \frac{\partial^2 u(t, x_1, x_2)}{\partial x_1^2} + a_{22}(t, x_1, x_2) \frac{\partial^2 u(t, x_1, x_2)}{\partial x_2^2} + \\
 &+ b_1(t, x_1, x_2) \frac{\partial u(t, x_1, x_2)}{\partial x_1} + b_2(t, x_1, x_2) \frac{\partial u(t, x_1, x_2)}{\partial x_2} + \\
 &+ c(t, x_1, x_2) + g_1(t, x_1, x_2) \frac{\partial u(t, x_1, x_2)}{\partial x_1} u(t, x_1, x_2) + \\
 &+ g_2(t, x_1, x_2) \frac{\partial u(t, x_1, x_2)}{\partial x_2} v(t, x_1, x_2)
 \end{aligned}$$

$$\begin{aligned}
\frac{\partial v(t, x_1, x_2)}{\partial t} &= e_{11}(t, x_1, x_2) \frac{\partial^2 v(t, x_1, x_2)}{\partial x_1^2} + e_{22}(t, x_1, x_2) \frac{\partial^2 v(t, x_1, x_2)}{\partial x_2^2} + \\
&+ f_1(t, x_1, x_2) \frac{\partial v(t, x_1, x_2)}{\partial x_1} + f_2(t, x_1, x_2) \frac{\partial v(t, x_1, x_2)}{\partial x_2} + \\
&+ d(t, x_1, x_2) + h_1(t, x_1, x_2) \frac{\partial v(t, x_1, x_2)}{\partial x_1} u(t, x_1, x_2) + \\
&+ h_2(t, x_1, x_2) \frac{\partial v(t, x_1, x_2)}{\partial x_2} v(t, x_1, x_2)
\end{aligned} \tag{7.12}$$

in zwei Raumdimensionen. Dabei seien die Koeffizientenfunktionen wie folgt gewählt:

$$\begin{array}{ll}
a_{11}(t, x_1, x_2) &= \sin(x_1) & e_{11}(t, x_1, x_2) &= \cos(x_1) \\
a_{22}(t, x_1, x_2) &= \sin(x_2) & e_{22}(t, x_1, x_2) &= \cos(x_2) \\
b_1(t, x_1, x_2) &= x_1^2 x_2 & f_1(t, x_1, x_2) &= x_1 x_2^2 \\
b_2(t, x_1, x_2) &= x_1 + x_2 & f_2(t, x_1, x_2) &= x_1 + x_2 \\
g_1(t, x_1, x_2) &= e^{x_1} & h_1(t, x_1, x_2) &= e^{x_1 + x_2} \\
g_2(t, x_1, x_2) &= e^{x_2} & h_2(t, x_1, x_2) &= e^{x_1 x_2}
\end{array}$$

Zudem seien die Funktionen  $c(t, x_1, x_2)$  und  $d(t, x_1, x_2)$  so gewählt, daß

$$\begin{aligned}
u(t, x_1, x_2) &= x_1(1 - x_1)x_2(1 - x_2) \sin(10\pi t) \\
v(t, x_1, x_2) &= (x_1 + \sin(\pi t))(1 + \sin(\pi t) - x_1) \cdot \\
&\quad \cdot (x_2 + 1 - \cos(\pi t))(\cos(\pi t) - x_2)
\end{aligned}$$

die Lösungsfunktionen der angegebenen Problemstellung seien. Die Randbedingungen  $\vec{r}(t, \mathbf{x})$  ergeben sich dann aus der Einschränkung der beiden Lösungsfunktionen auf den Rand  $\Gamma^d$ .

Wie zuvor im Beispiel P2 betrachten wir auf einem fest gewählten Gitter die Abhängigkeit des Fehlers, gemessen in verschiedenen Normen, von der Zeitschrittweite  $\Delta t_i$ , die in diesem Beispiel definiert ist vermöge

$$\Delta t_i = 2^{-i} \Delta t_0 .$$

In den numerischen Versuchen wurde  $\Delta t_0 = 10^{-3}$  gewählt, wobei wie schon im Beispiel P2 das fest gewählte Gitter ein reguläres dünnes Gitter der Leveltiefe 5 ist. In den Tabellen 44 bis 47 sind die erzielten numerischen Ergebnisse zusammengefaßt. Dabei werden in den Tabellen die einzelnen Fehler für die  $u$ -Komponente sowie der  $v$ -Komponente stets gegenübergestellt. Es zeigt sich, daß bei dem verwendeten Beispiels das explizite Eulersche Zeitschrittverfahren von 2. Ordnung konvergent ist, was sich auf das Fehlen entsprechender Terme in der Taylorentwicklung der Lösungsfunktionen zurückführen läßt. Somit stellt sich auch im Fall dünner Gitter wie erwartet dieses Phänomen ein. In den Abbildungen 75 und 76 sind die Resultate der numerischen Berechnungen graphisch zusammengefaßt.

Tabelle 44:  $L_\infty$ -Fehler für Problem P3, reguläres dünnes Gitter

$i$	$q_{\Delta t_i}$	$L_\infty^S(u)$	$q_{L_\infty^S}(u)$	$q_1^S(u)$	$L_\infty^S(v)$	$q_{L_\infty^S}(v)$	$q_1^S(v)$
0	-	2.48 <sub>-5</sub>	-	-	9.15 <sub>-6</sub>	-	-
1	2.00	6.22 <sub>-6</sub>	3.99	1.99	2.29 <sub>-6</sub>	4.00	2.00
2	2.00	1.56 <sub>-6</sub>	3.99	1.99	5.72 <sub>-7</sub>	4.00	2.00
3	2.00	3.89 <sub>-7</sub>	4.01	2.01	1.43 <sub>-7</sub>	4.00	2.00
4	2.00	9.73 <sub>-8</sub>	4.00	2.00	3.57 <sub>-8</sub>	4.01	2.00
5	2.00	2.43 <sub>-8</sub>	4.00	2.00	8.93 <sub>-9</sub>	4.00	2.00
6	2.00	6.08 <sub>-9</sub>	4.00	2.00	2.23 <sub>-9</sub>	4.00	2.00
7	2.00	1.52 <sub>-9</sub>	4.00	2.00	5.58 <sub>-10</sub>	4.00	2.00
8	2.00	3.80 <sub>-10</sub>	4.00	2.00	1.40 <sub>-10</sub>	3.99	1.99
9	2.00	9.50 <sub>-11</sub>	4.00	2.00	3.49 <sub>-11</sub>	4.01	2.01

Tabelle 45:  $L_2$ -Fehler für Problem P3, reguläres dünnes Gitter

$i$	$q_{\Delta t_i}$	$L_2^S(u)$	$q_{L_2^S}(u)$	$q_2^S(u)$	$L_2^S(v)$	$q_{L_2^S}(v)$	$q_2^S(v)$
0	-	1.30 <sub>-5</sub>	-	-	7.17 <sub>-6</sub>	-	-
1	2.00	3.27 <sub>-6</sub>	3.98	1.99	1.79 <sub>-6</sub>	4.01	2.00
2	2.00	8.23 <sub>-7</sub>	3.97	1.99	4.48 <sub>-7</sub>	4.00	2.00
3	2.00	2.06 <sub>-7</sub>	4.00	2.00	1.12 <sub>-7</sub>	4.00	2.00
4	2.00	5.16 <sub>-8</sub>	3.99	2.00	2.80 <sub>-8</sub>	4.00	2.00
5	2.00	1.29 <sub>-8</sub>	4.00	2.00	6.99 <sub>-9</sub>	4.01	2.00
6	2.00	3.23 <sub>-9</sub>	3.99	2.00	1.75 <sub>-9</sub>	3.99	2.00
7	2.00	8.08 <sub>-10</sub>	4.00	2.00	4.37 <sub>-10</sub>	4.00	2.00
8	2.00	2.02 <sub>-10</sub>	4.00	2.00	1.09 <sub>-10</sub>	4.01	2.00
9	2.00	5.08 <sub>-11</sub>	3.98	1.99	2.73 <sub>-11</sub>	3.99	2.00

Tabelle 46:  $L_1$ -Fehler für Problem P3, reguläres dünnes Gitter

$i$	$q_{\Delta t_i}$	$L_1^S(u)$	$q_{L_1^S}(u)$	$q_3^V(u)$	$L_1^S(v)$	$q_{L_1^S}(v)$	$q_3^V(v)$
0	-	1.02 <sub>-5</sub>	-	-	6.73 <sub>-6</sub>	-	-
1	2.00	2.59 <sub>-6</sub>	3.94	1.97	1.68 <sub>-6</sub>	4.01	2.00
2	2.00	6.51 <sub>-7</sub>	3.98	1.99	4.20 <sub>-7</sub>	4.00	2.00
3	2.00	1.63 <sub>-7</sub>	3.99	2.00	1.05 <sub>-7</sub>	4.00	2.00
4	2.00	4.09 <sub>-8</sub>	3.99	1.99	2.62 <sub>-8</sub>	4.01	2.00
5	2.00	1.02 <sub>-8</sub>	4.01	2.00	6.56 <sub>-9</sub>	3.99	2.00
6	2.00	2.56 <sub>-9</sub>	3.98	1.99	1.64 <sub>-9</sub>	4.00	2.00
7	2.00	6.40 <sub>-10</sub>	4.00	2.00	4.10 <sub>-10</sub>	4.00	2.00
8	2.00	1.60 <sub>-10</sub>	4.00	2.00	1.02 <sub>-10</sub>	4.02	2.01
9	2.00	4.00 <sub>-11</sub>	4.00	2.00	2.56 <sub>-11</sub>	3.98	1.99



Tabelle 47: Punkt-Fehler  $T$  für Problem P3, reguläres dünnes Gitter

$i$	$q_{\Delta t_i}$	$T(u)$	$q_T(u)$	$q_3^S(u)$	$T(v)$	$q_T(v)$	$q_3^S(v)$
0	-	$6.16_{-6}$	-	-	$7.52_{-6}$	-	-
1	2.00	$1.60_{-6}$	3.85	1.93	$1.88_{-6}$	4.00	2.00
2	2.00	$4.08_{-7}$	3.92	1.96	$4.69_{-7}$	4.01	2.00
3	2.00	$1.03_{-7}$	3.96	1.98	$1.17_{-7}$	4.01	2.00
4	2.00	$2.59_{-8}$	3.98	1.99	$2.93_{-8}$	3.99	2.00
5	2.00	$6.49_{-9}$	3.99	2.00	$7.32_{-9}$	4.00	2.00
6	2.00	$1.62_{-9}$	4.01	2.00	$1.83_{-9}$	4.00	2.00
7	2.00	$4.06_{-10}$	3.99	2.00	$4.57_{-10}$	4.00	2.00
8	2.00	$1.02_{-10}$	4.98	1.99	$1.14_{-10}$	4.01	2.00
9	2.00	$2.54_{-11}$	4.02	2.01	$2.86_{-11}$	3.99	1.99

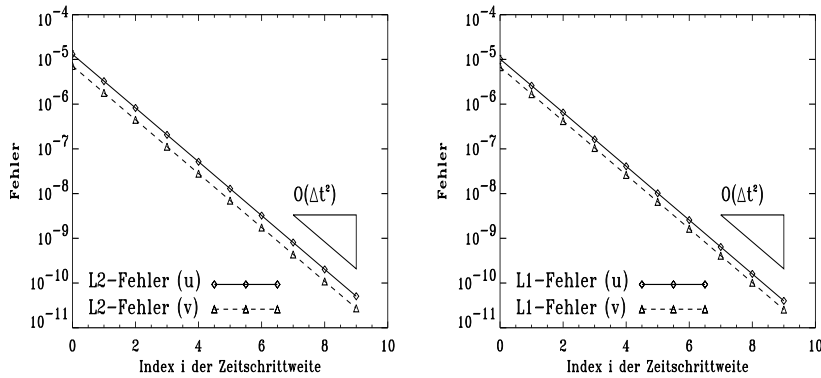


Abbildung 75: Konvergenzverhalten beim Problem P3:  $L_2^S$  (links) sowie  $L_1^V$  (rechts), jeweils beide Komponenten  $u$  und  $v$  in einer Graphik

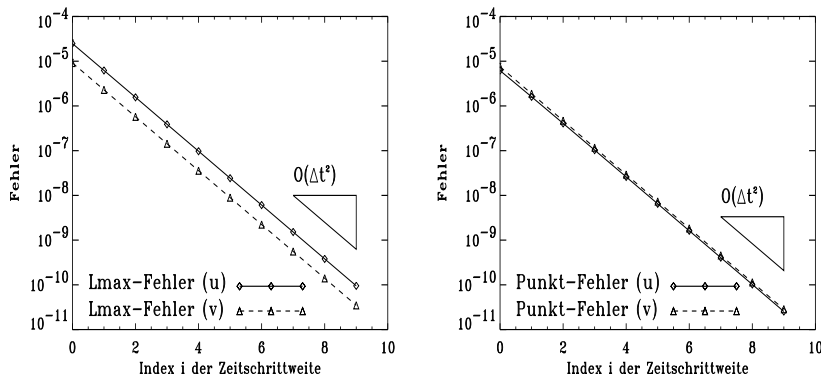


Abbildung 76: Konvergenzverhalten beim Problem P3:  $L_\infty^S$  (links) sowie Punktfehler  $T$  (rechts), jeweils beide Komponenten  $u$  und  $v$  in einer Graphik

Neben dem Einfluß der Zeitschrittweite interessieren wir uns für den zeitlichen Verlauf der Fehler auf einem festen Gitter zusammen mit einer festen Zeitschrittweite. Als Beispiel hierzu wählen wir ein reguläres dünnes Gitter auf dem Level 5 mit der Zeitschrittweite  $\Delta t = 10^{-4}$ . Die Fehlerkurven

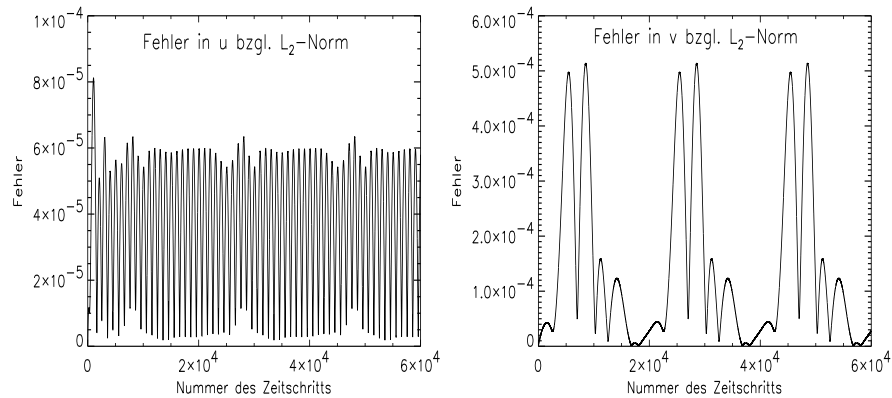
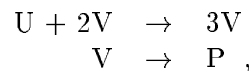


Abbildung 77: (periodischer) zeitlicher Verlauf der  $L_2$ -Norm des Fehlers der  $u$ -Komponente (links) und der  $v$ -Komponente (rechts)

bzgl. der  $L_2$ -Norm über den zeitlichen Verlauf sind in der Abbildung 77 zu sehen, die entsprechenden Kurven für die Maximumsnorm weisen ein nahezu identisches Verhalten auf. Deutlich kann man in der genannten Abbildung erkennen, daß sich der Fehler periodisch mit der Periode der jeweiligen Lösungskomponente verhält.

### 7.4.3 Gray Scott Modell für Reaktion und Diffusion chemischer Spezies

Das Gray Scott Modell [AACHNS, LMOS93, Pea93] basiert auf den beiden chemischen Reaktionen



wobei U, V und P chemische Spezies darstellen. Die mathematischen Gleichungen, mit denen das Verhalten der drei Komponenten untereinander simuliert werden kann, stellen ein nichtlineares System parabolischer Differentialgleichungen dar und lauten [AACHNS, GSS, LMOS93, Pea93, Wil]

$$\frac{\partial u}{\partial t} = D_u \Delta u - uv^2 + f(1 - u) \quad (7.13)$$

$$\frac{\partial v}{\partial t} = D_v \Delta v + uv^2 - (f + k)v \quad (7.14)$$

Dabei bezeichnen  $u$  und  $v$  die Konzentrationen der Spezies U und V,  $D_u$  und  $D_v$  sind die jeweiligen Diffusionsraten,  $k$  ist die Umwandlungsrate von

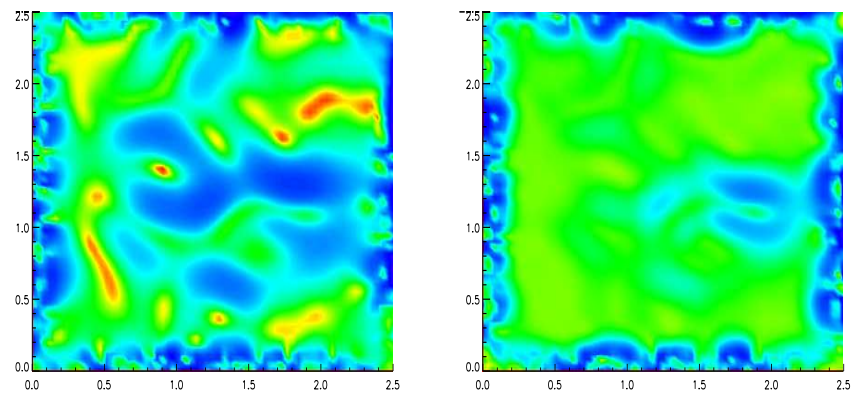
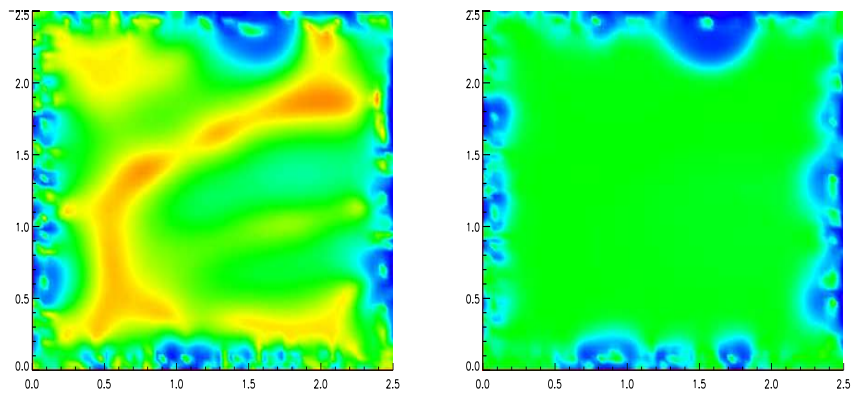
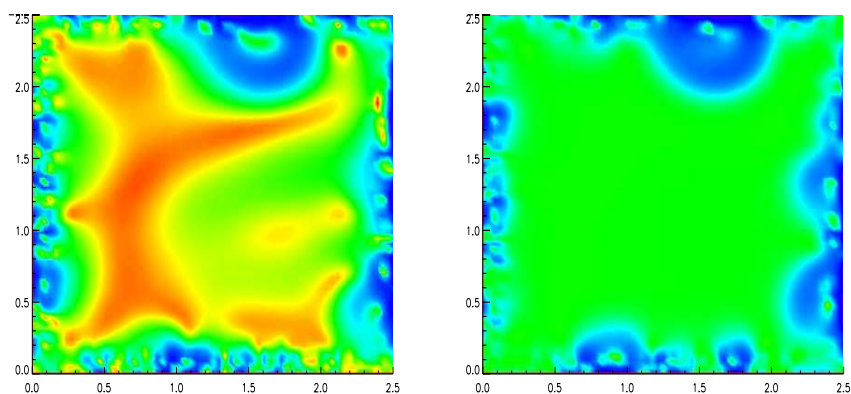
V nach P, und  $f$  repräsentiert die Rate des Prozesses, der U dem System zuführt und gleichzeitig V und P abführt.

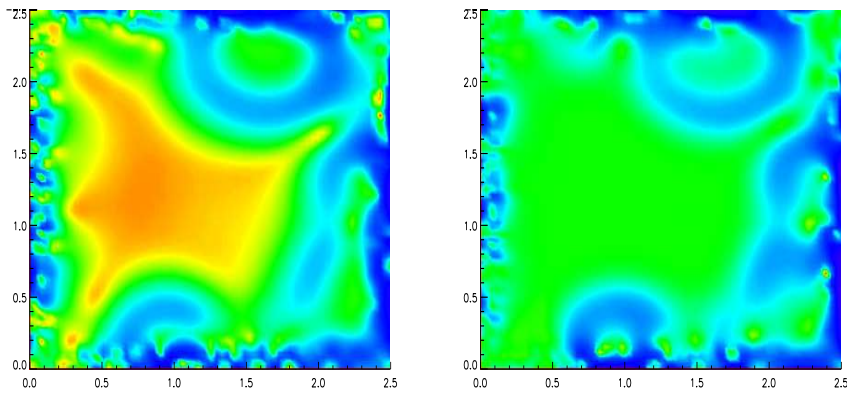
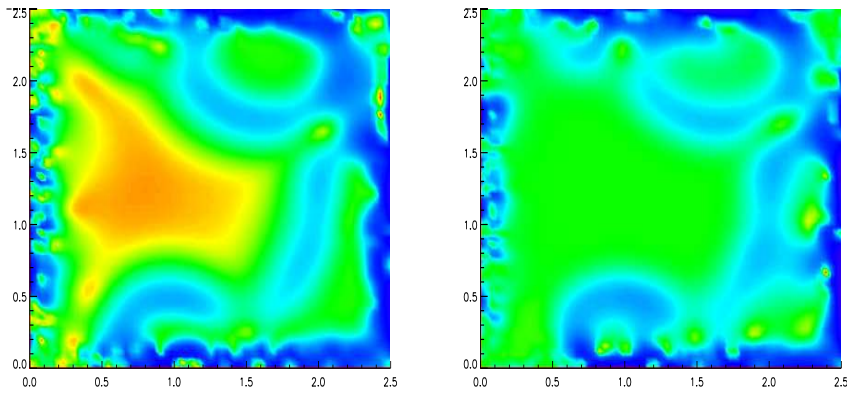
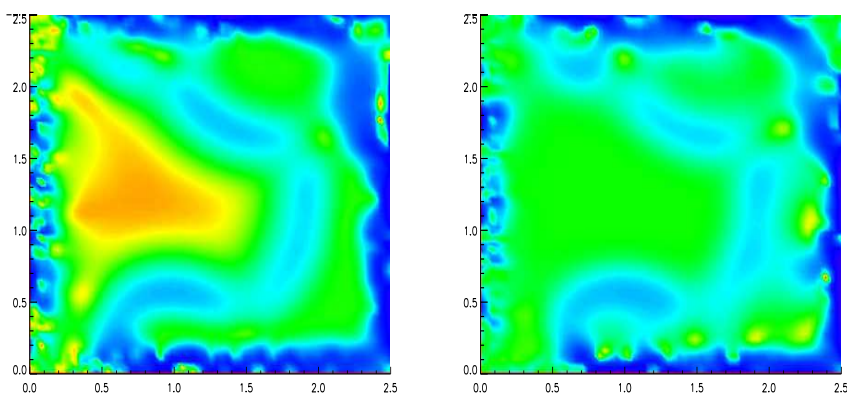
Das System (7.13), (7.14) besitzt für alle Parameterkombinationen  $(f, k) \in \mathbb{R}^2$  die stabile Lösung  $u = 1, v = 0$ , die somit einen stationären Zustand des obigen Systems beschreibt. Das Gray Scott Modell zur Simulation spezieller chemischer Vorgänge führt jedoch, wenn man den stationären Zustand einer beliebigen, "kleinen" Störung unterwirft, für einen großen Bereich von Parametertupeln  $(f, k)$  zu morphologischen Strukturen, die einer Vielzahl unterschiedlich ausgebildeter Konzentrationen entsprechen. Einen Überblick über die zu erwartenden Muster kann man etwa in [AACHNS, GSS, Wil] finden. Die am häufigsten auftretenden Muster sind dabei "fingerförmige" sowie "punktförmige" Verteilungen bzw. Konzentrationen. Insbesondere befindet sich unter [GSS] eine JAVA-Version eines Gray Scott Simulators, der die Eingabe der Parameter  $f$  und  $k$  sowie der Diffusionsraten  $D_u$  und  $D_v$  erlaubt und bei der die zeitliche Entwicklung der Konzentrationen graphisch verfolgt werden kann.

Kleine Störungen einer konstanten Ausgangssituation führen im allgemeinen bei diesem Modell dazu, daß sich diese Störungen in  $u$  und  $v$  zunächst "kreisförmig" ausbreiten und sich im Laufe der Zeit zu den angegebenen morphologischen Strukturen ausbilden. Nachfolgend finden sich die Ergebnisse von drei Simulationen auf dünnen Gitter, wobei zum Vergleich beim dritten Beispiel (Beispiel GS3) auch die Ergebnisse auf einem vollen Gitter berechnet wurden. Die drei verschiedenen Simulationen zeigen dabei drei verschiedene Reaktionsverhalten, die in Abhängigkeit der einzustellenden Parameter  $D_u, D_v, f$  und  $k$  erhalten werden können: die Konzentrationen, die im Gray Scott Modell berechnet werden, gehen für bestimmte Prozeßparameter in einen konstanten, stationären Zustand über (vgl. Beispiel GS2). Unterwirft man den stationären Zustand  $u = 1, v = 0$  einer Störung, etwa in einem quadratischen Teilgebiet in der Mitte des Gesamtgebiets, wie dies auch im Gray Scott Simulator [GSS] eingestellt werden kann, so breitet sich diese Störung im Laufe der Zeit über das gesamte Gebiet aus (Beispiel GS3). In Beispiel GS1 betrachten wir jedoch zunächst den Fall, daß sich fingerförmige Strukturen ausbilden.

**Beispiel GS1:** Die Prozeßparameter sind in diesem Beispiel  $f = 2.5 \cdot 10^{-2}$ ,  $k = 5 \cdot 10^{-2}$ ,  $D_u = 2 \cdot 10^{-5}$  sowie  $D_v = 10^{-5}$  mit  $\Delta t = 10^{-1}$ . Zum Start der Simulation wird eine Randomverteilung (in der Programmiersprache C wurde hier `drand48()` verwendet) der Werte zwischen 0 und 1 auf dem Grundgebiet  $\bar{\Omega} = [0, 2.5]^2$  (vgl. [Pea93]) angenommen.

Dabei bildet sich anfangs eine hohe Konzentration (gelbe/rote Farbe) von U aus, wohingegen die Konzentration von V (verglichen mit der von U) relativ gering ist. Im zeitlichen Verlauf der Simulation erhöht sich die Konzentration von U, bevor sie mit der Erhöhung der Konzentration von V fällt.

Abbildung 78: Konzentrationen  $u$  und  $v$ ,  $t = 40$  (Beispiel GS1)Abbildung 79: Konzentrationen  $u$  und  $v$ ,  $t = 80$  (Beispiel GS1)Abbildung 80: Konzentrationen  $u$  und  $v$ ,  $t = 120$  (Beispiel GS1)

Abbildung 81: Konzentrationen  $u$  und  $v$ ,  $t = 180$  (Beispiel GS1)Abbildung 82: Konzentrationen  $u$  und  $v$ ,  $t = 200$  (Beispiel GS1)Abbildung 83: Konzentrationen  $u$  und  $v$ ,  $t = 220$  (Beispiel GS1)

**Beispiel GS2:** Die Prozeßparameter sind in diesem Beispiel  $f = 5 \cdot 10^{-2}$ ,  $k = 6.25 \cdot 10^{-2}$ ,  $D_u = 10^{-2}$  sowie  $D_v = 5 \cdot 10^{-3}$  mit  $\Delta t = 10^{-2}$ . Zum Start der Simulation wird wiederum eine Randomverteilung der Werte auf dem Grundgebiet  $\bar{\Omega} = [0, 2.5]^2$  (vgl. [Pea93]) angenommen.

Zunächst bilden sich fingerartige Gebilde aus, die sich im Fall der  $u$ -Konzentration nach kurzer Zeit vereinigen und schließlich verschwinden, wobei eine homogene Konzentrationsverteilung erreicht wird. Bei der  $v$ -Konzentration erkennt man, daß dort, wo die Konzentration  $u$  hoch ist, die Konzentration  $v$  gering ist. Läßt man den Prozeß weiterlaufen, so geht auch  $v$  in einen konstanten, stationären Zustand über.

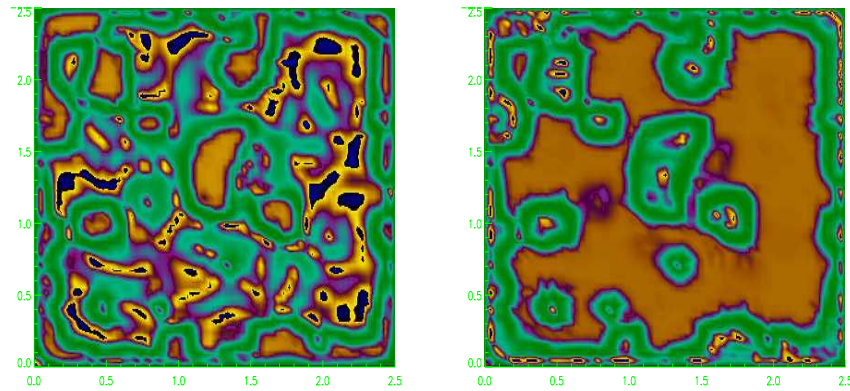


Abbildung 84: Konzentrationen  $u$  und  $v$ ,  $t = 4$  (Beispiel GS2)

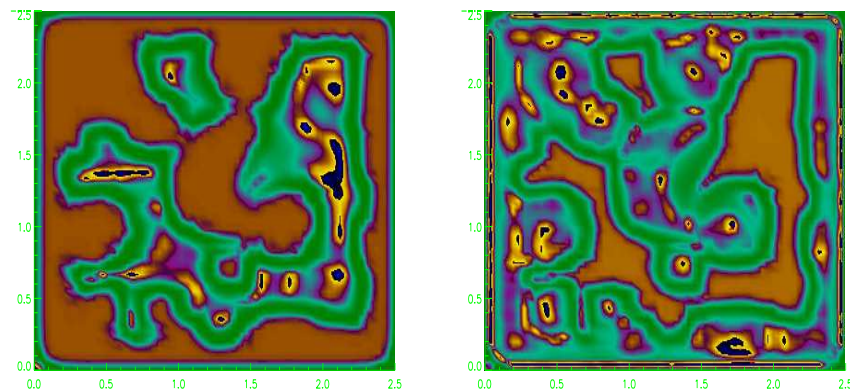
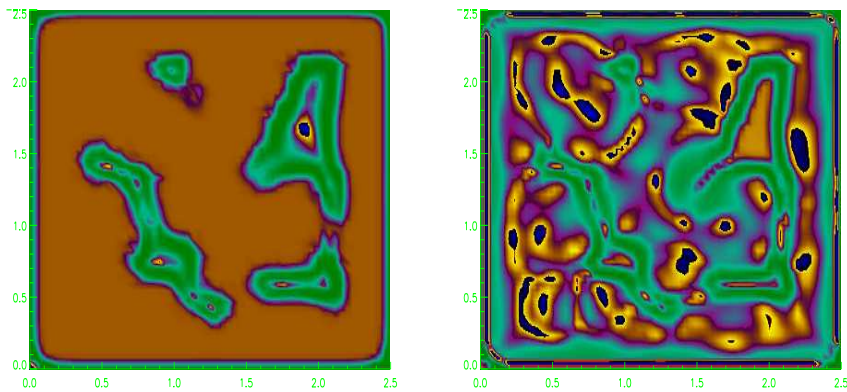
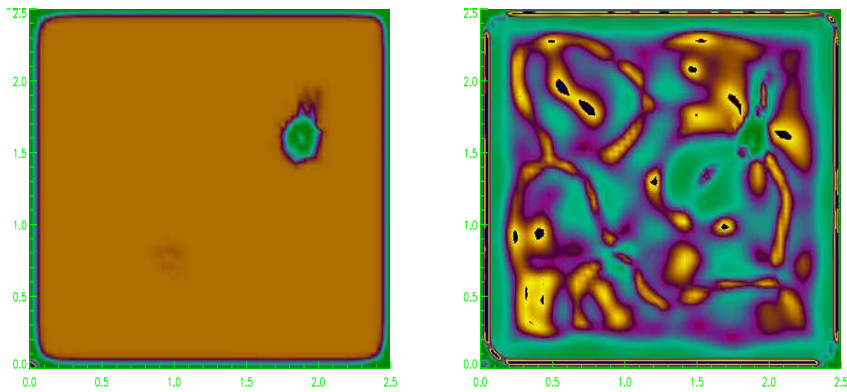
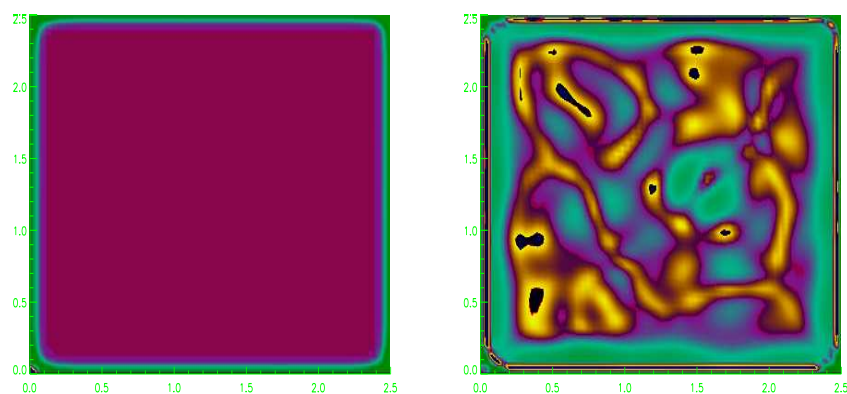


Abbildung 85: Konzentrationen  $u$  und  $v$ ,  $t = 8$  (Beispiel GS2)

**Beispiel GS3:** Die Prozeßparameter sind hier  $f = 10^{-2}$ ,  $k = 7 \cdot 10^{-2}$ ,  $D_u = 3 \cdot 10^{-2}$  sowie  $D_v = 10^{-2}$  mit der Zeitschrittweite  $\Delta t = 10^{-3}$ . Das Gebiet ist wiederum  $\bar{\Omega} = [0, 2.5]^2$ . Zum Start der Simulation wird der stationäre Zustand  $u = 1$ ,  $v = 0$  einer Randomverteilung derart unterworfen, daß

Abbildung 86: Konzentrationen  $u$  und  $v$ ,  $t = 12$  (Beispiel GS2)Abbildung 87: Konzentrationen  $u$  und  $v$ ,  $t = 16$  (Beispiel GS2)Abbildung 88: Konzentrationen  $u$  und  $v$ ,  $t = 20$  (Beispiel GS2)

die Werte in einem quadratischen Teilgebiet in der Mitte des Gesamtgebiets (vgl. [GSS, LMOS93, Pea93]) sowohl für  $u$  als auch für  $v$  gestört werden. In den nachfolgenden Abbildungen 89 bis 92, bei denen jeweils links das Simulationsergebnis auf einem vollen Gitter und rechts die entsprechende Simulation auf einem dünnen Gitter zum gleichen Zeitpunkt gezeigt sind, erkennt man, daß sich qualitativ dasselbe Verhalten einstellt. Dabei wurde für die Simulation ein reguläres dünnen Gitter der Leveltiefe 12 bzw. ein volles Gitter der Leveltiefe 6 verwendet. Zu Zwecken der Ausgabe wurde das dünne Gitter, wie bereits bei den vorangegangenen beiden Beispielen GS1 und GS2, auf ein volles Gitter interpoliert. In diesem Beispiel muß dabei auf ein volles Gitter interpoliert werden, das vollständig in dem betrachteten dünnen Gitter enthalten ist. Würde man hier zu Zwecken der Ausgabe das dünne Gitter der Leveltiefe 12 etwa auf ein volles Gitter der Leveltiefe 12 interpolieren, so würde sich in der graphischen Ausgabe die Struktur des dünnen Gitters widerspiegeln. Somit könnte man der Versuchung erliegen, dünne Gitter aufgrund des Effekts in der Visualisierung als ungeeignet für diese Anwendungen zu erachten. Daher muß (nicht nur hier, sondern generell) bei der Ausgabe von Berechnungen auf dünnen Gittern, bei denen die Struktur des dünnen Gitters sichtbar wird, darauf geachtet werden, ob die Dünngitterstrukturen nur wegen Interpolationen auftreten.

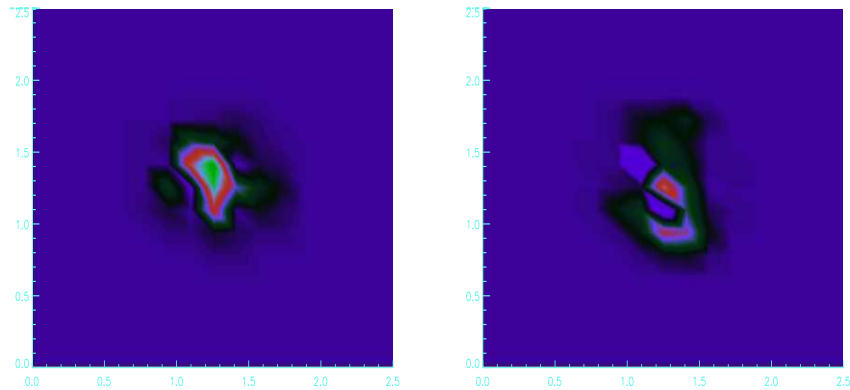
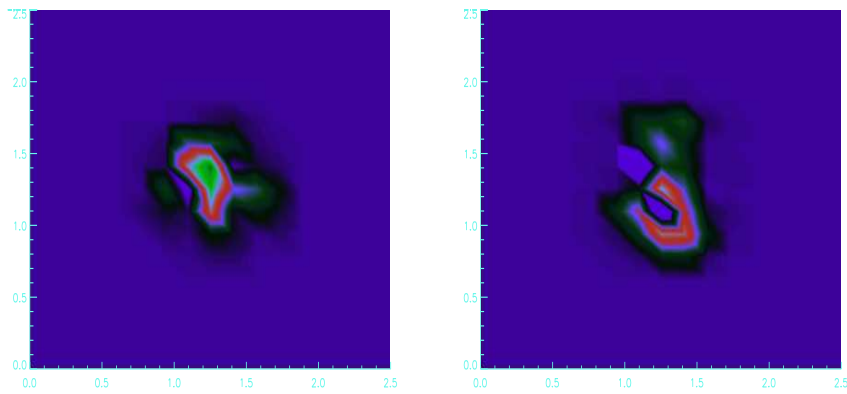
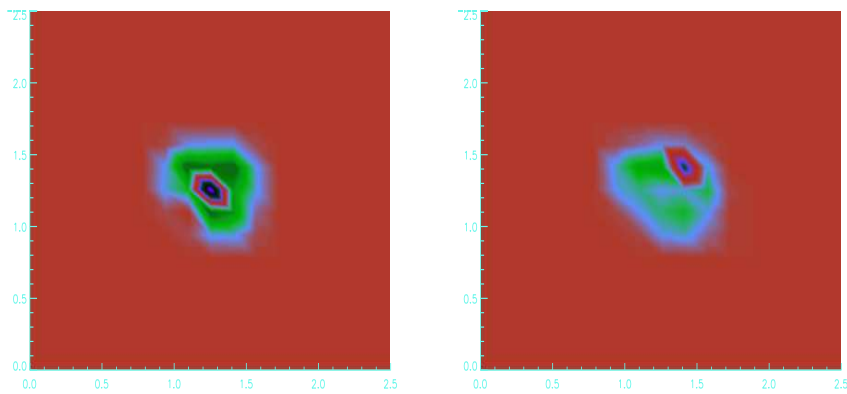
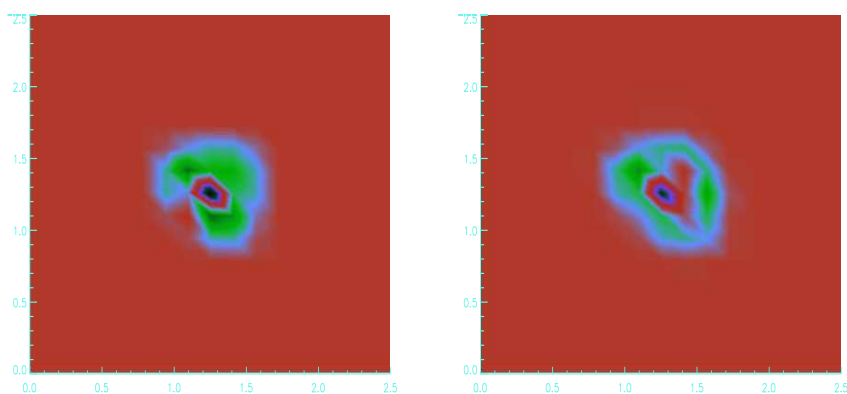


Abbildung 89: Vergleich der Konzentration  $u$ ,  $t = 2 \cdot 10^{-1}$  (Beispiel GS3)



Abbildung 90: Vergleich der Konzentration  $u$ ,  $t = 3 \cdot 10^{-1}$  (Beispiel GS3)Abbildung 91: Vergleich der Konzentration  $v$ ,  $t = 2 \cdot 10^{-1}$  (Beispiel GS3)Abbildung 92: Vergleich der Konzentration  $v$ ,  $t = 3 \cdot 10^{-1}$  (Beispiel GS3)

## 7.5 Navier Stokes Gleichungen

Die Strömung eines Fluids in einem Gebiet  $\Omega \subset \mathbb{R}^n$  im Laufe der Zeit  $t \in \mathbb{R}^+$  wird beschrieben mit Hilfe der Größen

- $\vec{u} : \Omega \times [0, T_{ende}] \rightarrow \mathbb{R}^n$  (Geschwindigkeiten),
- $p : \Omega \times [0, T_{ende}] \rightarrow \mathbb{R}$  (Druck) sowie
- $\rho : \Omega \times [0, T_{ende}] \rightarrow \mathbb{R}$  (Dichte).

Bei der Betrachtung von laminaren Strömungen inkompressibler Fluide gilt für die Dichte insbesondere  $\rho(\vec{x}, t) = const$ , so daß sich die Strömung im stationären Zustand durch das System von partiellen Differentialgleichungen

$$(\vec{u} \cdot \nabla) \vec{u} + \nabla p = \frac{1}{Re} \Delta \vec{u} + \vec{f}(\vec{u}) \quad (7.15)$$

$$\operatorname{div}(\vec{u}) = 0 \quad (7.16)$$

und im instationären Zustand durch das System

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} + \nabla p = \frac{1}{Re} \Delta \vec{u} + \vec{f}(\vec{u}) \quad (7.17)$$

$$\operatorname{div}(\vec{u}) = 0 \quad (7.18)$$

beschreiben läßt, das aus den Impulsgleichungen (7.15) bzw. (7.17) und der Kontinuitätsgleichung (7.16) bzw. (7.18) besteht und das man als Navier Stokes Gleichungen bezeichnet. Der Druck ist hierbei nur bis auf eine Konstante bestimmt, und  $\vec{f}(\vec{u})$  stellt äußere Kräfte (wie etwa die (Erd-)Anziehungskraft) dar, denen das physikalische System unterliegt.

Betrachtet man instationäre, temperaturabhängige Strömungen, so ergeben sich die entsprechenden Gleichungen zu

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u} \cdot \nabla) \vec{u} + \nabla p = \frac{1}{Re} \Delta \vec{u} + \vec{f}(\vec{u}) + (1 - \beta T) \vec{g} \quad (7.19)$$

$$\operatorname{div}(\vec{u}) = 0 \quad (7.20)$$

$$\frac{\partial T}{\partial t} + \frac{\partial(Tu)}{\partial x} + \frac{\partial(Tv)}{\partial y} + \frac{\partial(Tw)}{\partial z} = \frac{1}{RePr} \Delta T + f(T) \quad , \quad (7.21)$$

wenn man die Energiegleichung (7.21) zu den bisherigen Gleichungen (7.17) und (7.18) hinzufügt und die Kopplung von Temperatur und Geschwindigkeiten in den abgeänderten Impulsgleichungen (7.19) berücksichtigt. Dabei wird die Temperatur  $T$  mittels der Boussinesq-Approximation an die Geschwindigkeiten gekoppelt,  $Re$  ist die Reynoldszahl und  $Pr$  die Prandtlzahl. Eine ausführliche Herleitung dieser Gleichungen (sowohl im temperaturabhängigen Fall als auch im Fall ohne Energieberücksichtigung), der

Gültigkeit der Boussinesq–Approximation sowie eine Erklärung der verwendeten Konstanten findet man etwa in [GrDoNe98]. Wir stellen nachfolgend zwei unterschiedliche Lösungsverfahren vor, die mit Hilfe der Methode der Finiten Differenzen auf dünnen Gittern implementiert wurden.

Das erste Verfahren ist hierbei die “*Methode der künstlichen Kompressibilität*” (vgl. etwa [Cho67, PeTa90, Fle91, FePe96]), die auf die stationären Gleichungen (7.15) und (7.16) angewendet wird. Die zweite Methode zur Lösung der Navier Stokes Gleichungen ist eine *Projektionsmethode* [Gre87, PeTa90, Fle91, GrDoNe98, FePe96], bei der die Berechnung des Drucks mit Hilfe eines Poisson–Problems (7.2) gelöst wird, für das Dirichlet–Randbedingungen (und keine Neumann–Randbedingungen) in jedem Zeitschritt ermittelt werden.

### 7.5.1 Die Methode der künstlichen Kompressibilität

Die Methode der künstlichen Kompressibilität, die auf die stationären Gleichungen (7.15) und (7.16) angewendet wird, ist eine pseudo–stationäre Methode. Hier werden künstlich Zeitableitungen eingeführt und von den stationären Gleichungen auf die instationären Gleichungen (7.17) und (7.18) übergegangen. Zudem wird der Druck als zeitabhängige Größe interpretiert und die Kontinuitätsgleichung vermöge

$$\frac{\partial p}{\partial t} + c^2 \cdot \operatorname{div} u = 0 \quad (7.22)$$

modifiziert. Für den Grenzfall  $t \rightarrow \infty$ , für den die stationäre Lösung des Problems angenommen wird, folgt somit

$$\frac{\partial p}{\partial t} \equiv 0$$

und die Gleichungen (7.22) und (7.16) fallen zusammen. Die Gleichung (7.22) ist nur für den Grenzfall  $t \rightarrow \infty$  physikalisch sinnvoll. Gleichung (7.22) gibt im übrigen diesem Verfahren, das aufgrund der Ähnlichkeit der angesprochenen Gleichung und der entsprechenden Gleichung im kompressiblen Fall so genannt wird (vgl. [Fle91], die Bezeichnung stammt von J. A. Chorin [Cho67]), seinen Namen. Der Parameter  $c$  kann als künstliche Schallgeschwindigkeit interpretiert werden und es gilt

$$p = c^2 \rho \ .$$

Wie aus den vorhergehenden Gleichungen zu entnehmen ist, taucht im inkompressiblen Fall die Dichte jedoch nicht explizit in den Berechnungen auf und die beiden Parameter  $c$  und  $\Delta t$  (Zeitschrittweite bei der Zeitdiskretisierung) haben die Rolle eines Relaxationsparameters, wobei die Wahl von  $\Delta t$  nach bestimmten Stabilitätskriterien zu erfolgen hat und abhängig von der

Wahl von  $c$  ist. Für ein explizites Verfahren (explizites Eulerverfahren) wird nach [StKu77, Fle91] die Wahl

$$c^2 \leq \frac{1}{\Delta t}$$

empfohlen. Eine genauere Analyse für die Wahl der beteiligten Parameter findet sich in den angegebenen Quellen.

Aus der Literatur ist dieses Verfahren als verlässliches, aber äußerst langsam konvergierendes Verfahren bekannt. Es ist daher sinnvoll, anstelle eines feinsten Gitters, auf dem man die Lösung berechnen will, die Lösung unter Zuhilfenahme mehrerer gröberer Gitter zu ermitteln. Neben der Berechnung der Navier Stokes Gleichungen auf einem festen Gitter verwenden wir nachfolgend eine geschachtelte Iteration, wobei die Methode der künstlichen Kompressibilität als Glätter verwendet wird. Der Einsatz eines echten Mehrgitter-Verfahrens würde natürlich eine noch geringere Anzahl von Iterationen mit sich bringen, muß aber an dieser Stelle unterbleiben. Die Diskretisierung der Zeitableitung geschieht wie bereits im Abschnitt über parabolische Differentialgleichungen mittels eines expliziten Eulerschrittes. Die Berechnung der Ortsableitungen erfolgt über die zuvor eingeführten Finiten Differenzen Operatoren auf dünnen Gittern. Somit erfolgt die Diskretisierung der Methode der künstlichen Kompressibilität in zwei Dimensionen wie folgt:

$$\begin{aligned} u^{n+1}(x_1, x_2) &= u^n(x_1, x_2) + \Delta t \cdot \left( \Delta^S u^n(x_1, x_2) + f_1(x_1, x_2) - \right. \\ &\quad \left. - \mathbf{D}_1^S((u^n(x_1, x_2))^2) - \mathbf{D}_2^S(u^n(x_1, x_2) \cdot v^n(x_1, x_2)) - \right. \\ &\quad \left. - \mathbf{D}_1^S(p^n(x_1, x_2)) \right) \\ v^{n+1}(x_1, x_2) &= v^n(x_1, x_2) + \Delta t \cdot \left( \Delta^S v^n(x_1, x_2) + f_2(x_1, x_2) - \right. \\ &\quad \left. - \mathbf{D}_1^S(u^n(x_1, x_2) \cdot v^n(x_1, x_2)) - \mathbf{D}_2^S((v^n(x_1, x_2))^2) - \right. \\ &\quad \left. - \mathbf{D}_2^S(p^n(x_1, x_2)) \right) \\ p^{n+1}(x_1, x_2) &= p^n(x_1, x_2) - \\ &\quad c^2 \cdot \Delta t \cdot \left( \mathbf{D}_1^S(u^{n+1}(x_1, x_2)) + \mathbf{D}_2^S(v^{n+1}(x_1, x_2)) \right) \end{aligned}$$

Analog erfolgt die Diskretisierung der Methode der künstlichen Kompressibilität in drei Dimensionen vermöge

$$\begin{aligned} u^{n+1}(\mathbf{x}) &= u^n(\mathbf{x}) + \Delta t \cdot \left( \Delta^S u^n(\mathbf{x}) + f_1(\mathbf{x}) - \mathbf{D}_1^S((u^n(\mathbf{x}))^2) \right. \\ &\quad \left. - \mathbf{D}_2^S(u^n(\mathbf{x}) \cdot v^n(\mathbf{x})) - \mathbf{D}_3^S(u^n(\mathbf{x}) \cdot w^n(\mathbf{x})) - \mathbf{D}_1^S(p^n(\mathbf{x})) \right) \end{aligned}$$

$$\begin{aligned}
v^{n+1}(\mathbf{x}) &= v^n(\mathbf{x}) + \Delta t \cdot \left( \Delta^S v^n(\mathbf{x}) + f_2(\mathbf{x}) - \mathbf{D}_1^S(u^n(\mathbf{x}) \cdot v^n(\mathbf{x})) \right. \\
&\quad \left. - \mathbf{D}_2^S((v^n(\mathbf{x}))^2) - \mathbf{D}_3^S(v^n(\mathbf{x}) \cdot w^n(\mathbf{x})) - \mathbf{D}_2^S(p^n(\mathbf{x})) \right) \\
w^{n+1}(\mathbf{x}) &= w^n(\mathbf{x}) + \Delta t \cdot \left( \Delta^S w^n(\mathbf{x}) + f_3(\mathbf{x}) - \mathbf{D}_1^S(u^n(\mathbf{x}) \cdot w^n(\mathbf{x})) \right. \\
&\quad \left. - \mathbf{D}_2^S(v^n(\mathbf{x}) \cdot w^n(\mathbf{x})) - \mathbf{D}_3^S((w^n(\mathbf{x}))^2) - \mathbf{D}_3^S(p^n(\mathbf{x})) \right) \\
p^{n+1}(\mathbf{x}) &= p^n(\mathbf{x}) - c^2 \cdot \Delta t \cdot \left( \mathbf{D}_1^S(u^n(\mathbf{x})) + \mathbf{D}_2^S(v^n(\mathbf{x})) + \mathbf{D}_3^S(w^n(\mathbf{x})) \right)
\end{aligned}$$

mit  $\mathbf{x} = (x_1, x_2, x_3)$ .

**Beispiel N1:** Als erstes einfaches Testbeispiel dient das in [Cho67] angegebene Beispiel einer parabolischen Strömung durch einen Kanal. Die Lösung  $u(\mathbf{x}) = [u(x_1, x_2), v(x_1, x_2)]$  des Problems auf dem Gebiet  $\bar{\Omega} = [0, 10] \times [0, 1]$  in zwei Raumdimensionen ist hierbei gegeben durch

$$\begin{aligned}
u(x_1, x_2) &= 4x_2(1 - x_2) \\
v(x_1, x_2) &= 0 \\
p(x_1, x_2) &= C - x \quad ,
\end{aligned}$$

wobei  $C$  eine beliebige Konstante ist. Abbildung 93 zeigt die Lösung des Problems. Der Start der Iterationen erfolgt hierbei mit einer nicht divergenzfreen Startlösung. In den nachfolgenden Graphiken der Abbildung 94 sind die Fehlerverhalten für die zwei verschiedenen Parameter  $c = 100$  und  $c = 10$  enthalten. Es zeigt sich hierbei, daß sich für die kleinere Wahl des Relaxationsparameters  $c$  bereits nach einer wesentlich geringeren Zahl von Iterationen eine höhere Genauigkeit in der Lösung einstellt als für den größeren Parameter von  $c$  (beachte hierbei die unterschiedliche Skalierung in Abbildung 94). Damit vergleichbar ist das Verhalten im Fall der geschachtelten Iteration (vgl. hierzu Abbildung 95), wo ebenfalls bei der Wahl des kleineren Parameters  $c$  das Konvergenzverhalten wesentlich besser ist. Zudem erkennt man, daß durch die Anwendung der geschachtelten Iteration die Anzahl nötiger Iterationen zum Erreichen einer vorgegebenen Genauigkeit wesentlich geringer ist als bei Verwendung eines einzelnen Gitters. An dieser Stelle sei nochmals darauf hingewiesen, daß ein Mehrgitter-Verfahren dieselben Genauigkeiten in einem Bruchteil der hier noch nötigen Iterationen liefern könnte.

**Beispiel N2:** Das zweite Beispiel ist die Nischenströmung (lid driven cavity). Hier wird ein Band mit einer vorgegebenen Geschwindigkeit über einem mit einem Fluid gefüllten Behältnis gezogen, wodurch die Flüssigkeit in Bewegung versetzt wird. Für eine genaue Beschreibung des Problems sei unter vielen etwa auf [GGS82, GrDoNe98] verwiesen. Die Nischenströmung wurde für moderate Reynoldszahlen  $Re \leq 100$  gerechnet, das Ergebnis für  $Re = 50$  befindet sich zusammen mit einem adaptiven dünnen Gitter in der Abbildung 96.

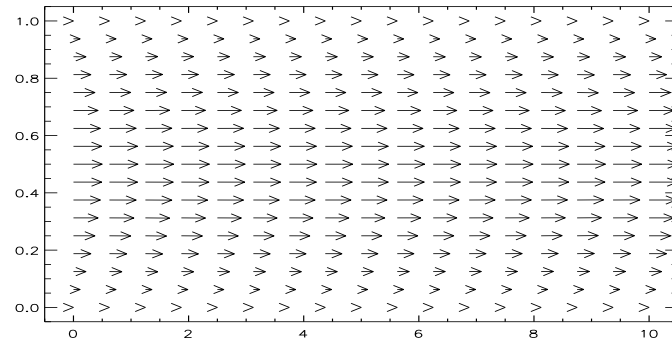


Abbildung 93: Lösung des Problems N1

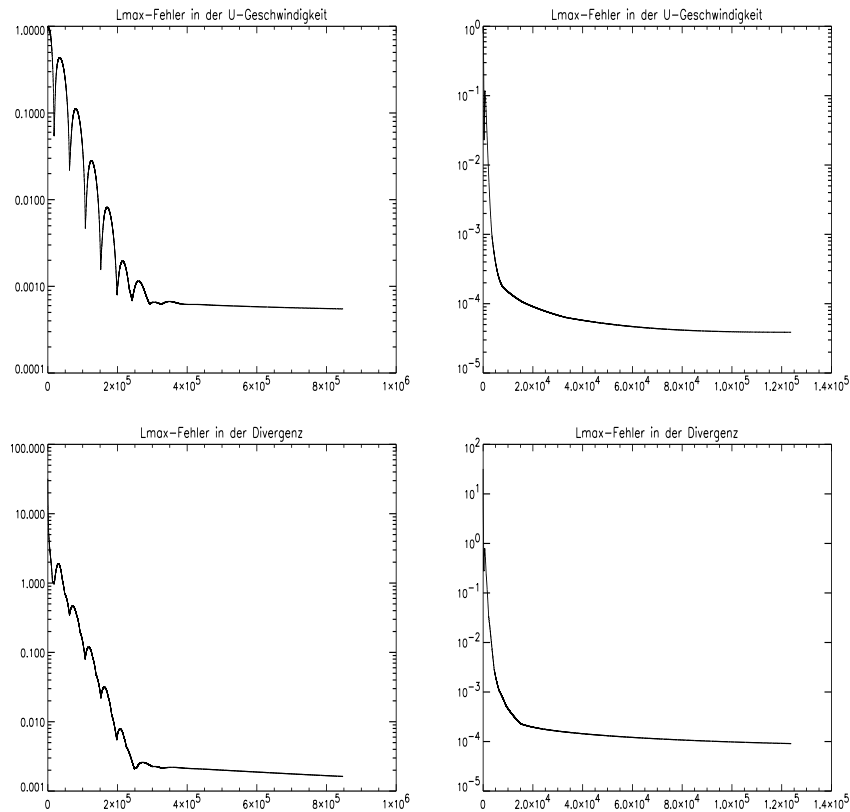
links:  $c=100$ rechts:  $c=10$ 

Abbildung 94: Fehler für das Problem N1 bei der Methode der künstlichen Kompressibilität als Löser (berechnet auf einem festen Gitter): Maximums-Norm in der  $u_1$ -Geschwindigkeit (oben) und Maximums-Norm in der Divergenz (unten)

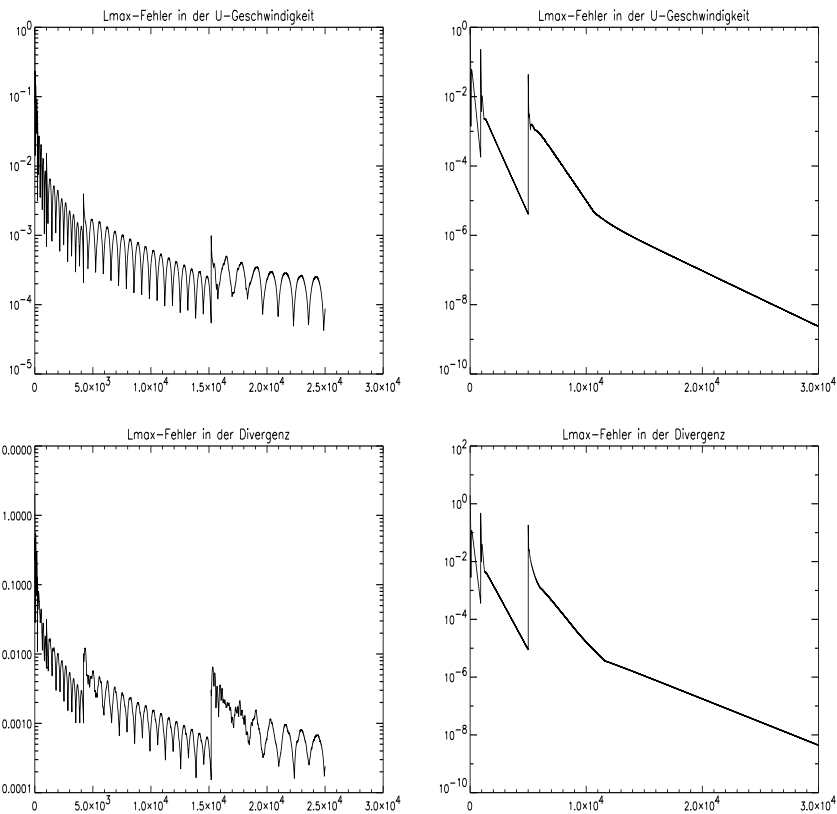
links:  $c=100$ rechts:  $c=10$ 

Abbildung 95: Fehler für das Problem N1 bei der Methode der künstlichen Kompressibilität als Löser (berechnet mittels geschachtelter Iteration, wobei die Methode der künstlichen Kompressibilität als Glätter verwendet wird): Maximums-Norm in der  $u_1$ -Geschwindigkeit (oben) und Maximums-Norm in der Divergenz (unten)

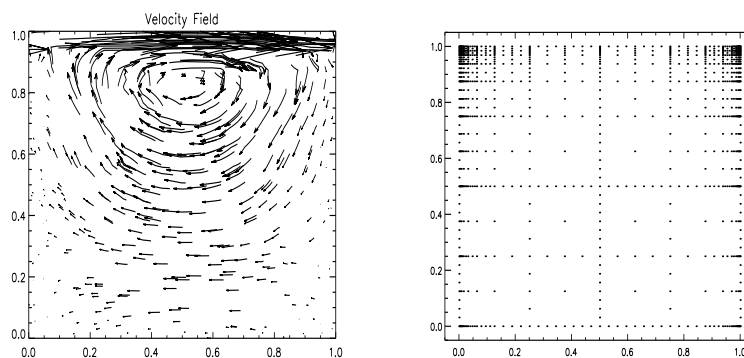


Abbildung 96: Nischenströmung mit  $Re = 50$  und adaptives dünnes Gitter (Beispiel N2)

**Beispiel N3:** Wir betrachten hier ein zweidimensionales Beispiel [Pro97], bei dem die Lösung der stationären Navier Stokes Gleichungen bekannt ist. Dabei lauten die Geschwindigkeiten

$$\begin{aligned} u_1(x, y) &= x^2(1-x)^2(2y-6y^2+4y^3) \\ u_2(x, y) &= -y^2(1-y)^2(2x-6x^2+4x^3) , \end{aligned}$$

und der Druck ergibt sich zu

$$p(x, y) = \left(x^2 + y^2 - \frac{2}{3}\right) .$$

In den Tabellen 48 bis 50 sind die Ergebnisse für den Fehler im Druck, dem Fehler in der Divergenzfreiheit des Geschwindigkeitsfeldes  $\vec{u}$  (vgl. (7.16)) sowie dem Fehler in der ersten Geschwindigkeitskomponente  $u_1(x, y)$  auf sukzessiven Leveln regulärer dünner Gitter angegeben. Dabei sind die Fehler in der zweiten Geschwindigkeitskomponente  $u_2(x, y)$  wegen  $u_1(x, y) = -u_2(y, x)$  bezüglich der Maximumsnorm und der  $L_2$ -Norm identisch zu den Fehlern in  $u_1(x, y)$ . Aufgrund der glatten Lösung sowohl in den Geschwindigkeiten als auch im Druck sind die für dünne Gitter typischen Konvergenzraten von  $\mathcal{O}(h^2 \log(h^{-1}))$  zu beobachten. Die Abbildungen 97 und 98 zeigen das Konvergenzverhalten, in Abbildung 98 (rechts) ist zudem das Strömungsfeld für das Beispiel N3 enthalten.

Tabelle 48: Fehler für Problem N3, hier: Druck  $p$

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	5.60 <sub>-2</sub>	-	-	3.96 <sub>-2</sub>	-	-
3	49	2.33	3.49 <sub>-2</sub>	1.60	0.69	2.05 <sub>-2</sub>	1.93	0.83
4	113	2.31	1.62 <sub>-2</sub>	2.14	0.92	7.97 <sub>-3</sub>	2.57	1.11
5	257	2.27	6.32 <sub>-3</sub>	2.56	1.13	2.73 <sub>-3</sub>	2.92	1.29
6	577	2.25	2.21 <sub>-3</sub>	2.86	1.27	8.60 <sub>-4</sub>	3.17	1.41
7	1281	2.22	7.29 <sub>-4</sub>	3.03	1.37	2.57 <sub>-4</sub>	3.35	1.51
8	2817	2.20	2.25 <sub>-4</sub>	3.24	1.47	7.43 <sub>-5</sub>	3.46	1.57

Tabelle 49: Fehler für Problem N3, hier:  $\nabla \cdot \vec{u}$

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	5.63 <sub>-2</sub>	-	-	3.98 <sub>-2</sub>	-	-
3	49	2.33	3.47 <sub>-2</sub>	1.62	0.70	2.02 <sub>-2</sub>	1.97	0.85
4	113	2.31	1.51 <sub>-2</sub>	2.30	0.99	7.57 <sub>-3</sub>	2.67	1.16
5	257	2.27	5.93 <sub>-3</sub>	2.55	1.21	2.54 <sub>-3</sub>	2.98	1.31
6	577	2.25	2.05 <sub>-3</sub>	2.89	1.29	7.96 <sub>-4</sub>	3.19	1.42
7	1281	2.22	6.63 <sub>-4</sub>	3.09	1.39	2.36 <sub>-4</sub>	3.37	1.52
8	2817	2.20	2.04 <sub>-4</sub>	3.25	1.48	6.79 <sub>-5</sub>	3.48	1.58



Tabelle 50: Fehler für Problem N3, hier: Geschwindigkeit  $u_1(x, u)$

$l$	$P$	$q_P$	$L_\infty^S$	$q_{L_\infty^S}$	$q_1^S$	$L_2^S$	$q_{L_2^S}$	$q_2^S$
2	21	-	$2.34_{-3}$	-	-	$1.17_{-3}$	-	-
3	49	2.33	$1.21_{-3}$	1.93	0.83	$4.90_{-4}$	2.39	1.02
4	113	2.31	$4.13_{-4}$	2.93	1.27	$1.81_{-4}$	2.71	1.17
5	257	2.27	$1.24_{-4}$	3.33	1.47	$5.77_{-5}$	3.14	1.38
6	577	2.25	$3.61_{-5}$	3.43	1.53	$1.72_{-5}$	3.35	1.49
7	1281	2.22	$1.03_{-5}$	3.50	1.58	$4.97_{-6}$	3.46	1.56
8	2817	2.20	$2.92_{-6}$	3.53	1.60	$1.41_{-6}$	3.52	1.60

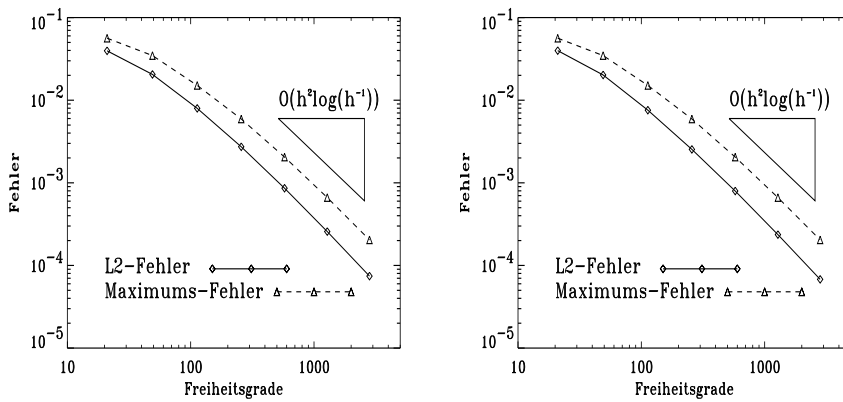


Abbildung 97: Konvergenzverhalten des Drucks (links) und der Divergenzfreiheit (rechts) bzgl. der Maximumsnorm und der  $L_2$ -Norm

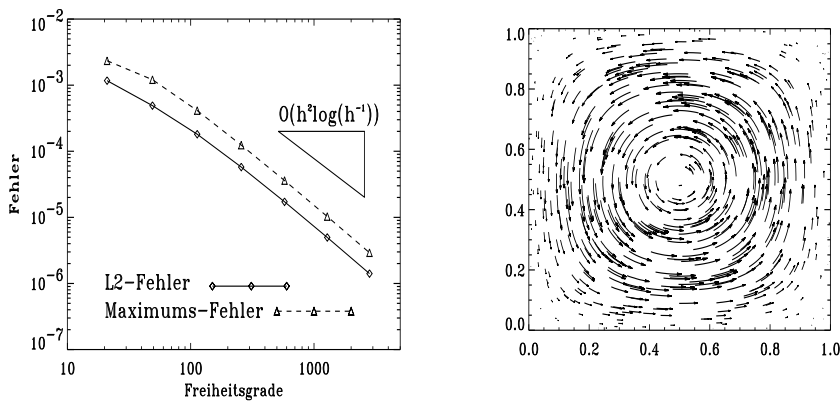


Abbildung 98: Konvergenzverhalten der Geschwindigkeit  $u_1(x, y)$  bzgl. der Maximumsnorm und der  $L_2$ -Norm (links) sowie das Strömungsfeld des Problems N3 (rechts)

### 7.5.2 Ein Projektionsverfahren zur Lösung der Navier Stokes Gleichungen auf dünnen Gittern

Projektionsmethoden zur Lösung der Navier Stokes Gleichungen wurden unabhängig von A. J. Chorin [Cho68] und R. Temam [Tem69] eingeführt. Eine explizite Version einer Projektionsmethode von M. Fortin et. al. [FPT97], die zur Berechnung der Beispiele N4 und N5, die später genauer beschrieben werden, verwendet werden soll, wird anschließend vorgestellt.

Die explizite Methode von M. Fortin et. al. ist eine *fractional step method* mit erster Ordnung bezüglich der Zeitdiskretisierung. Im ersten Schritt zur Berechnung der Gleichungen (7.15) und (7.16) wird hierbei eine vorläufige, nicht divergenzfreie Geschwindigkeit  $u^*$  vermöge

$$\frac{u^* - u^n}{\Delta t} + (\bar{u} \cdot \nabla) \bar{u} = \frac{1}{Re} \Delta \bar{u} + \vec{f}(\bar{u}) \quad , \quad (7.23)$$

ermittelt, was genau der Impulsgleichung unter Vernachlässigung des Drucks entspricht. Dabei wird der Superscript  $n$  hier für die Werte zum Zeitpunkt  $0 \leq t_n \leq T_{ende}$  benutzt. An dieser Stelle sei angemerkt, daß hier lediglich die Zeitdiskretisierung von Interesse ist, und die Ortsdiskretisierung an dieser Stelle auf beliebige Art und Weise vorgenommen werden kann. Der zweite Schritt besteht nun darin, die Geschwindigkeit  $u^*$  mit Hilfe der Gleichungen

$$\frac{u^{n+1} - u^*}{\Delta t} + \nabla p^{n+1} = 0 \quad (7.24)$$

und (7.16) zu korrigieren. Dazu berechnet man die Divergenz von (7.24) und benutzt die Gleichung (7.16), die die Divergenzfreiheit der neuen Iterierten  $u^{n+1}$  fordert. Man erhält somit das Poisson–Problem

$$\Delta p^{n+1} = \frac{1}{\Delta t} \operatorname{div}(u^*) \quad (7.25)$$

zur Berechnung des Drucks  $p^{n+1}$ . Als Randbedingungen für das Problem (7.25) werden in der Literatur in der Regel Neumann–Randbedingungen vorgeschlagen, vgl. etwa [Gre87, PeTa90, Fle91, FePe96], wobei die Aufzählung möglicher Zitate und Quellen hier alles andere als vollständig ist. Zudem finden sich auch Quellen (etwa [Gre87, Sid95]), die die Berechnung von Dirichlet–Randbedingungen vermöge einer Integration über den Druckgradienten am Rand des Gebiets vorschlagen. Wir wollen an dieser Stelle den letztgenannten Weg zur Berechnung des Problems (7.25) gehen.

Für die Diskretisierung der Navier Stokes Gleichungen finden sich in der Literatur zwei grundlegend verschiedene Gittervarianten: Gitter, bei denen in jedem Gitterpunkt die Informationen bzgl. aller unbekanntten Größen gespeichert werden (collocated grids) und Gitter, bei denen unterschiedliche Informationen an unterschiedlichen Gitterpunkten sitzen (staggered grids).

Im Fall dünner Gitter ist bei Verwendung der stückweisen  $d$ -linearen Basis der letztgenannte Fall ungeeignet, wohingegen der Fall der *collocated grids* auch hier verwendet werden kann<sup>16</sup>. Der hier verwendete Algorithmus zur Berechnung der Navier Stokes Gleichungen (7.19)–(7.21), der ein *collocated grid* verwendet und der an die Ausführungen in [Gre87] angelehnt ist, wird nachfolgend unter Berücksichtigung obiger Einführung beschrieben. Dabei wird die Berechnung der Temperatur zum neuen Zeitschritt  $t_n$  vor der Berechnung der nicht divergenzfreien Geschwindigkeit zum selben Zeitschritt durchgeführt. Die Berechnung der Gleichungen (7.17) und (7.18), bei denen keine Temperaturberechnung berücksichtigt wird, erfolgt analog zur nachfolgend vorgestellten Vorgehensweise, wenn alle temperaturabhängigen Größen vernachlässigt werden.

- Initialisierung:

alle notwendigen Daten zum Start der Iteration wie etwa Gitteraufbau, Kennzeichnung der Knoten (Dirichlet-Knoten, innerer Knoten etc.), Einlesen von Dirichlet-Daten usw. werden hier bereitgestellt; Iteriere nun über die Zeitintervalle  $t_n = n \cdot \Delta t$

- Berechnung der Temperatur zum neuen Zeitschritt  $t_{n+1}$ :

$$T^{n+1} = T^n + \Delta t \cdot \left( \frac{1}{RePr} \Delta T^n - \frac{\partial(T^n u^n)}{\partial x} - \frac{\partial(T^n v^n)}{\partial y} - \frac{\partial(T^n w^n)}{\partial z} + f^{n+1}(T^n) \right)$$

- Berechnung intermediärer Geschwindigkeiten  $\bar{u}_{div}^{n+1}$ , die nicht divergenzfrei sind:

$$u_{div}^{n+1} = u^n + \Delta t \cdot \left( \frac{1}{Re} \Delta u^n - \frac{\partial(u^n)^2}{\partial x} - \frac{\partial(u^n v^n)}{\partial y} - \frac{\partial(u^n w^n)}{\partial z} + f^{n+1}(u) + (1 - \beta T^{n+1}) g_u \right)$$

$$v_{div}^{n+1} = v^n + \Delta t \cdot \left( \frac{1}{Re} \Delta v^n - \frac{\partial(v^n u^n)}{\partial x} - \frac{\partial(v^n)^2}{\partial y} - \frac{\partial(v^n w^n)}{\partial z} + f^{n+1}(v) + (1 - \beta T^{n+1}) g_v \right)$$

<sup>16</sup>Dabei sei an dieser Stelle darauf hingewiesen, daß bei Verwendung von *collocated grids* grundsätzlich Stabilitätsprobleme auftreten können.

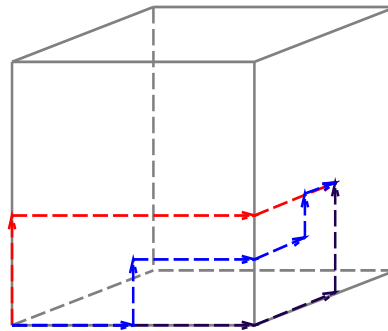
$$w_{div}^{n+1} = w^n + \Delta t \cdot \left( \frac{1}{Re} \Delta w^n - \frac{\partial(w^n u^n)}{\partial x} - \frac{\partial(w^n v^n)}{\partial y} - \frac{\partial(w^n)^2}{\partial z} + f^{n+1}(w) + (1 - \beta T^{n+1}) g_w \right)$$

- Berechnung des Drucks:

Wir lösen die Impulsgleichungen nach dem Gradienten  $\nabla p$  des Drucks  $p$  auf und erhalten somit

$$\nabla p^{n+1} = -\frac{\partial \bar{u}_{div}^{n+1}}{\partial t} - (\bar{u}_{div}^{n+1} \cdot \nabla) \bar{u}_{div}^{n+1} + \frac{1}{Re} \Delta \bar{u}_{div}^{n+1} + \bar{f}^{n+1}(\bar{u}) + (1 - \beta T^{n+1}) \bar{g}$$

Wendet man auf diese Gleichung den Divergenzoperator  $\nabla \cdot$  an, so ergibt sich wegen  $\nabla \cdot (\nabla) = \Delta$  eine Poisson-Gleichung für die Berechnung des Drucks. Durch Integration des tangentialen Anteils des Druckgradienten auf dem Rand  $\Gamma$  des Grundgebiets  $\Omega$  erhalten wir nun Dirichlet-Randbedingungen für die Poisson-Gleichung. Dabei ist es in drei Dimensionen aus theoretischer Sicht unerheblich, welcher Tangentialanteil z.B. auf einer Seite eines Würfels genommen wird, wie in Abbildung 99 zu sehen ist. Integration entlang jedem der ange-



— — —> Integrationswege für die Integration des Gradienten von  $p$

Abbildung 99: Randintegration über  $\nabla p$

zeigten Wege liefert bei (kontinuierlicher) Rechnung den exakten Integralwert bzw. Dirichlet-Randwert. Aus numerischer Sicht ergibt sich das Problem, daß sich die zweidimensionale Integration auf einer Seite der Oberfläche ohne großen rechnerischen Aufwand realisieren lassen soll, und demnach die zu verwendenden Integrationsformeln nur auf den Punkten des zweidimensionalen dünnen Gitters auf der Oberfläche des Gebiets operieren sollen. Somit ist derjenige Integrationsweg aus numerischer Sicht zu bevorzugen, der ausgehend vom Startpunkt des Integrationswegs die größtmögliche Anzahl von Punkten eines dünnen

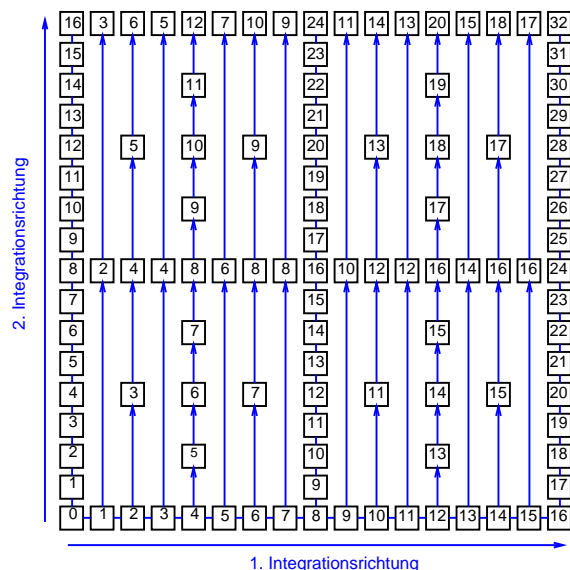


Abbildung 100: Randintegration: Integration zuerst in  $x$ -Richtung, dann in  $y$ -Richtung; die Zahlen geben die Anzahl verwendeter Gitterpunkte zur Integration an

Gitters benutzt. Für die Randintegration in zwei Raumdimensionen zeigen die beiden Abbildungen 100 und 101 zwei unterschiedliche Strategien für die Wahl von Integrationswegen, die beide für sich allein

betrachtet nicht “optimal” sind, wohingegen eine Kopplung beider Vorgehensweisen (man nimmt jeweils denjenigen Fall, bei dem mehr Gitterpunkte zur Gewinnung des Integrationswertes benötigt werden), wie sie in Abbildung 102 zu sehen ist, eine bessere Berechnung der Dirichlet-Daten liefert. Auf einem regulären dünnen Gitter in drei Raumdimensionen, bei dem die Oberflächen aus regulären dünnen Gittern in zwei Raumdimensionen bestehen, ist diese Vorgehensweise einfach realisierbar. Im Falle adaptiver dünner Gitter läuft die Suche nach dem Integrationsweg mit einer maximalen Anzahl an Gitterpunkten zur Integration auf ein Optimierungsproblem für (fast) jeden einzelnen Punkt auf dem Rand des dreidimensionalen dünnen Gitters hinaus. Da dies zu aufwendig ist, kann man in der Praxis wie folgt vorgehen: man führt für jede einzelne der sechs Teilflächen der Oberfläche eines Würfels die Integration in  $x/y$ -Richtung und in  $y/x$ -Richtung (vgl. die beiden Abbildungen 100 und 101) durch und nimmt dann den Integralwert, bei dem mehr Gitterpunkte zur Berechnung verwendet worden sind. Diese Vorgehensweise liefert in der Praxis erheblich bessere numerische Ergebnisse anstatt einer Vorgehensweise nach Abbildung 100 oder 101. So zeigt sich, daß bei der vorgeschlagenen Vorgehensweise bei

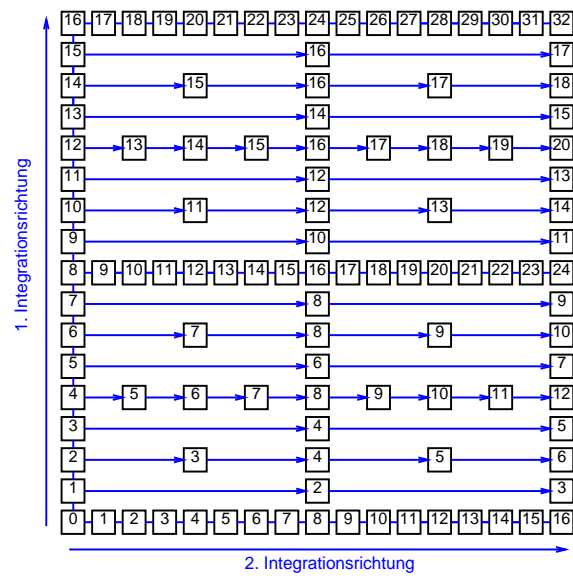


Abbildung 101: Randintegration: Integration zuerst in  $y$ -Richtung, dann in  $x$ -Richtung; die Zahlen geben die Anzahl verwendeter Gitterpunkte zur Integration an

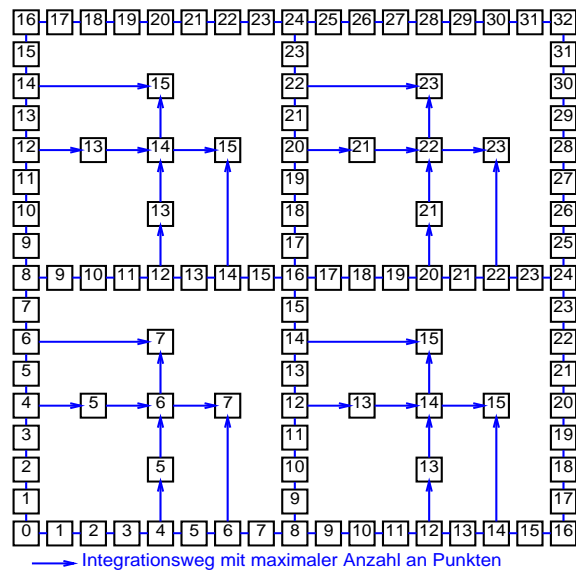


Abbildung 102: Randintegration: Integrationsweg mit maximaler Anzahl an Integrationspunkten; die Zahlen geben die Anzahl verwendeter Gitterpunkte zur Integration an

Lösungen, die symmetrisch bezüglich einer Achse sind, auch die numerischen Lösungen diese Symmetrie aufweisen. Verwendet man jedoch eine der beiden einfachen Integrationsweisen, so zeigen die numerischen Lösungen oft nach bereits wenigen Iterationen Unsymmetrien in der numerischen Lösung, die allein durch die (schlechtere) Randintegration verursacht werden.

- Korrektur der Geschwindigkeiten:

$$\begin{aligned} u^{n+1} &= u_{div}^{n+1} + \Delta t \cdot \frac{\partial p}{\partial x} \\ v^{n+1} &= v_{div}^{n+1} + \Delta t \cdot \frac{\partial p}{\partial y} \\ w^{n+1} &= w_{div}^{n+1} + \Delta t \cdot \frac{\partial p}{\partial z} \end{aligned}$$

Die Diskretisierung der vorgestellten Formeln mittels der Finiten Differenzen Operatoren auf dünnen Gittern führt damit auf folgendes Schema:

- Diskretisierung der Temperatur-Gleichung

$$\begin{aligned} T^{n+1} &= T^n + \Delta t \cdot \left( \frac{1}{RePr} \Delta^S T^n - \mathbf{D}_1^S(T^n u^n) - \mathbf{D}_2^S(T^n v^n) \right. \\ &\quad \left. \mathbf{D}_3^S(T^n w^n) + f^{n+1}(T^n) \right) \end{aligned}$$

- Berechnung der nicht divergenzfreien Geschwindigkeiten

$$\begin{aligned} u_{div}^{n+1} &= u^n + \Delta t \cdot \left( \frac{1}{Re} \Delta^S u^n - \mathbf{D}_1^S((u^n)^2) - \mathbf{D}_2^S((u^n \cdot v^n)) - \right. \\ &\quad \left. \mathbf{D}_3^S(u^n \cdot w^n) + f^{n+1}(u) + (1 - \beta T^{n+1}) g_u \right) \\ v_{div}^{n+1} &= v^n + \Delta t \cdot \left( \frac{1}{Re} \Delta^S v^n - \mathbf{D}_1^S(v^n \cdot u^n) - \mathbf{D}_2^S((v^n)^2) - \right. \\ &\quad \left. \mathbf{D}_3^S(v^n \cdot w^n) + f^{n+1}(v) + (1 - \beta T^{n+1}) g_v \right) \\ w_{div}^{n+1} &= w^n + \Delta t \cdot \left( \frac{1}{Re} \Delta^S w^n - \mathbf{D}_1^S(w^n \cdot u^n) - \mathbf{D}_2^S(w^n \cdot v^n) - \right. \\ &\quad \left. \mathbf{D}_3^S((w^n)^2) + f^{n+1}(w) + (1 - \beta T^{n+1}) g_w \right) \end{aligned}$$

- Diskretisierung der Poisson-Gleichung für den Druck

$$\begin{aligned} \Delta^S p^{n+1} &= \nabla^S \cdot \left( - \frac{\vec{u}_{div}^{n+1} - \vec{u}^n}{\Delta t} - (\vec{u}_{div}^{n+1} \cdot \nabla^S) \vec{u}_{div}^{n+1} + \right. \\ &\quad \left. \frac{1}{Re} \Delta^S \vec{u}_{div}^{n+1} + \vec{f}^{n+1}(\vec{u}) + (1 - \beta T^{n+1}) \vec{g} \right), \end{aligned}$$

wobei  $\nabla^S = (\mathbf{D}_1^S, \mathbf{D}_2^S, \mathbf{D}_3^S)^T$ ,  $\nabla^S \cdot = \mathbf{D}_1^S + \mathbf{D}_2^S + \mathbf{D}_3^S$  und  $\Delta^S \vec{u}_{div}^{n+1} \in \mathbb{R}^3$  komponentenweise zu verstehen ist.

- Korrektur der Geschwindigkeiten:

$$\begin{aligned} u^{n+1} &= u_{div}^{n+1} + \Delta t \cdot \mathbf{D}_1^S(p) \\ v^{n+1} &= v_{div}^{n+1} + \Delta t \cdot \mathbf{D}_2^S(p) \\ w^{n+1} &= w_{div}^{n+1} + \Delta t \cdot \mathbf{D}_3^S(p) \end{aligned}$$

Zur Berechnung der jeweiligen Iterierten verwenden wir neben der soeben angegebenen Diskretisierung die Algebra für Operatoren auf dünnen Gitter aus Abschnitt 3.2, um die zusätzlich noch notwendigen Operationen wie Addition, Skalarmultiplikation oder Multiplikation von Funktionen zu realisieren.

**Beispiel N4:** Als erstes Beispiel zum Test der Projektionsmethode mittels finiter Differenzen auf dünnen Gittern betrachten wir das dreidimensionale Analogon von Chorin's parabolischer Strömung durch einen Kanal. Das Grundgebiet ist hier  $\bar{\Omega}^3 = [0, 10] \times [0, 1]^2$  und die Lösung des Problems ist gegeben durch

$$\begin{aligned} u(x, y, z) &= 4y(1-y)z(1-z) \\ v(x, y, z) &= 0 \\ w(x, y, z) &= 0 \\ p(x, y, z) &= C - x, \end{aligned} \tag{7.26}$$

wobei  $C$  wie bereits im Beispiel N1 eine beliebige Konstante darstellt. Das Ergebnis der Berechnung befindet sich in Abbildung 103.

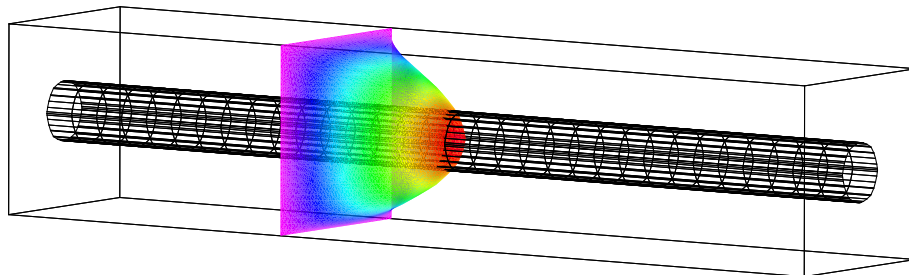


Abbildung 103: Beispiel N4: parabolische Strömung in einem dreidimensionalen Kanal (Strömung des Fluids von links nach rechts)

**Beispiel N5:** Dieses Beispiel ist eine Anwendung aus dem Bereich der chemischen Dampfphasenabscheidung (englisch *CVD*: chemical vapor deposition) und entstammt Arbeiten aus dem BMBF-Projekt<sup>17</sup> PAR-CVD

<sup>17</sup>BMBF-Projekt PAR-CVD, FKZ 01 IR 502 E 2, beteiligte Projektpartner: Friedrich-Alexander-Universität Erlangen-Nürnberg (Lehrstuhl Prof. Dr. Dr. h.c. Durst), Institut für Angewandte Mathematik der Universität Bonn (Abteilung Prof. Dr. Griebel), Institut für Informatik der TU München (Lehrstuhl Prof. Dr. Bode), AIXTRON AG (Aachen)



[CVD97]. In diesem Verbundprojekt erfolgt die Entwicklung mathematischer Modelle und effizienter numerischer Methoden zur Untersuchung und Prozeßoptimierung des Halbleiterwachstums von  $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ,  $\text{Ga}_x\text{In}_{1-x}\text{P}$  und  $\text{In}_x\text{Ga}_{1-x}\text{N}$  durch metallorganische Gasphasenepitaxie (engl. MOCVD: metalorganic vapor phase epitaxy). MOCVD ist ein Verarbeitungsschritt in der Halbleiterprozeßtechnik, bei der einkristalline Schichten hoher Qualität von III-V-Verbindungshalbleitern mit metallorganischen Verbindungen als Ausgangsstoffe erzeugt werden. Das Einsatzgebiet dieser Wachstumstechnik liegt auf der Fertigung von elektronischen und optoelektronischen Bauelementen, die auf III-V-Verbindungshalbleitern bestehen. CVD-Reaktoren, die von der Firma AIXTRON AG hergestellt werden und die Grundlage der Modellbildung und der numerischen Simulationen innerhalb dieses Projekts sind, sind lineare Horizontalreaktoren (vgl. Abbildung 104) sowie radialsymmetrische horizontale Mehrscheibenreaktoren mit einer charakteristischen, planetenartigen Drehbewegung der Wafer um die Reaktormitte.

Aus strömungstechnischer Sicht liegt bei der MOCVD eine laminare Strömung eines chemisch reagierenden Gasmisches mit stark gekoppelter natürlicher und erzwungener Konvektion bei kleinen Reynoldszahlen (ca. 1 – 100), kleinen Machzahlen (ca.  $10^{-2}$ ), mittleren bis hohen Rayleighzahlen ( $10^2 - 10^5$ ) und mit hohen Temperaturgradienten vor. Die mathematischen Gleichungen zur Formulierung des beschriebenen Systems finden sich etwa in [Dau95, Kad96, CVD97].

In diesem Beispiel interessieren uns insbesondere der Einfluß des rotierenden Suszeptors auf die Temperaturverteilung sowie auf die Geschwindigkeiten. Dazu reicht es aus, wenn wir zur numerischen Simulation der ablaufenden Prozesse die für dieses System vereinfachten Gleichungen (7.19) – (7.21) betrachten. Bei der gegebenen Problemstellung wird in einem CVD-Reaktor ein rotierender Suszeptor, auf dem sich während eines echten Versuchs die chemischen Bestandteile ablagern sollen, von unten durch (z.B.) Infrarotlampen stark erhitzt. Neben starken Temperaturgradienten entsteht somit aus mathematischer Sicht das Problem, daß am Rand des Suszeptors Sprünge in den Geschwindigkeiten auftreten. Dies macht den Einsatz adaptiver dünner Gitter sinnvoll und auch notwendig, um die Rechenzeiten gering zu halten und eine möglichst gute Approximation der numerischen Lösung an die tatsächlichen Verhältnisse im Reaktor zu erhalten. Das Fluid (in realen Anwendungen hat man es mit einem Gemisch mehrerer chemischer Spezies zu tun, das neben den zuvor genannten chemischen Stoffen insbesondere aus einem Trägergas besteht) wird in der Simulation mit einem parabolischen Anströmprofil in den CVD-Reaktor eingeleitet.

Es ist zu erwarten, daß die Rotation sowohl auf die Geschwindigkeiten als auch auf die Temperatur erheblichen Einfluß hat. Ohne die Rotation des Suszeptors würde das Fluid lediglich durch den Wärmefluß von seiner Bahn

abgelenkt werden, durch die Rotation des Suszeptors erhält das Fluid selbst einen gewissen Drehimpuls, der in seiner Stärke von der Geschwindigkeit der Rotation des Suszeptors, aber auch von der Einströmgeschwindigkeit des Fluids selbst abhängt.

Dieses Beispiel wurde auf einem adaptiven dünnen Gitter mit maximaler Leveltiefe von 10 gerechnet, wobei insgesamt 27947 Gitterpunkte verwendet wurden. Zur Generierung des Gitters wurde der Suszeptor vor Beginn der Berechnung adaptiv aufgelöst und die maximale Leveltiefe auf 10 beschränkt. Dieselbe Rechnung auf einem vollen Gitter der Leveltiefe 10 würde hingegen  $(2^{10})^3 \approx 1.07 \cdot 10^9$  Gitterpunkte benötigen, knapp 40000 mal mehr Gitterpunkte als das verwendete adaptive dünne Gitter.

Der Einfluß der Rotation ist in den Abbildungen 106 bis 108 zu erkennen. In den Abbildungen 106 und 107 sind die Temperaturverteilungen in einer Höhe von 10% und 20% der Gesamthöhe über dem Suszeptor gezeigt, wobei man deutlich erkennt, daß sich in der Temperatur aufgrund der durch die Rotation verursachten (zusätzlichen) Bewegung des Fluids keine Symmetrie einstellt. Den Einfluß auf die Geschwindigkeit (hier die Geschwindigkeit in Hauptstromrichtung) zeigt die Abbildung 108. Dabei erkennt man insbesondere, daß durch die Rotation des Suszeptors das Fluid bis zu einer gewissen Höhe im Reaktor in negative Hauptstromrichtung transportiert wird. Dabei ist klar, daß die Rotationsgeschwindigkeit des Suszeptors ebenso wie die Einströmgeschwindigkeit des Fluids und die am Boden anliegende Temperatur einen erheblichen Einfluß auf die Strömung im Reaktor haben.

Simulationen auf vollen Gittern benötigen wegen der erheblich höheren Anzahl an Gitterpunkten im Vergleich zu dünnen Gittern wesentlich mehr Hauptspeicher. Bei bisherigen Simulationen auf vollen Gittern war aufgrund der Hauptspeicherbedingten kleineren Auflösungsgenauigkeit der Effekt der Rotation bislang nicht deutlich in den Vordergrund getreten. Neuere Ergebnisse auf feineren Leveln auf Großrechnern zeigen, daß der Effekt der Rotation mit feiner werdender Auflösung des Suszeptors deutlich erkennbar wird.

Zusammen mit der Optimierung der in CVD-Reaktoren ablaufenden chemischen Vorgängen sowie dem auftretenden Strahlungstransport sind Untersuchungen der Geschwindigkeitsprofile und Temperaturverteilungen aktuelle Forschungsgebiete, um eine möglichst optimale Ausbeute der eingesetzten Chemikalien zu erreichen und somit den Kostenfaktor in der Massenproduktion zu senken sowie die ökologische Belastung durch Abfallprodukte zu verringern. Dünne Gitter können hier durch die erheblich geringere Anzahl an Gitterpunkten im Vergleich zu einem vollen Gitter (vgl. vorhergehende Diskussion weiter oben) bei annähernd gleicher Genauigkeit helfen, Simulationsergebnisse schneller zu erhalten und somit den realen Prozeß schneller und kostengünstiger zu optimieren.

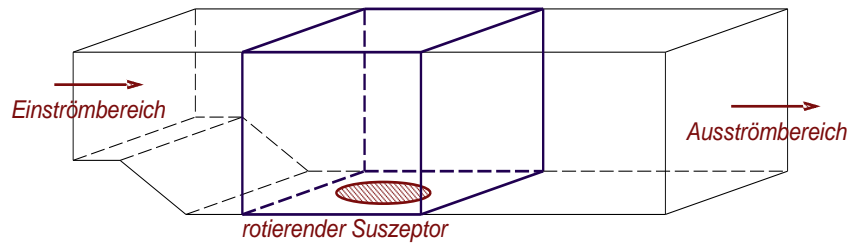


Abbildung 104: Geometrie des Reaktors und Grundgebiet der Simulation (markierte Box)

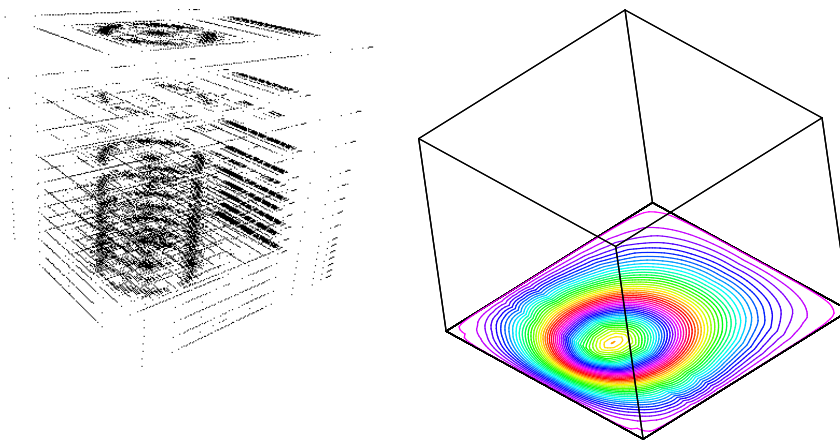


Abbildung 105: adaptives dünnes Gitter der Simulation und Temperaturverteilung am Boden des Grundgebiets

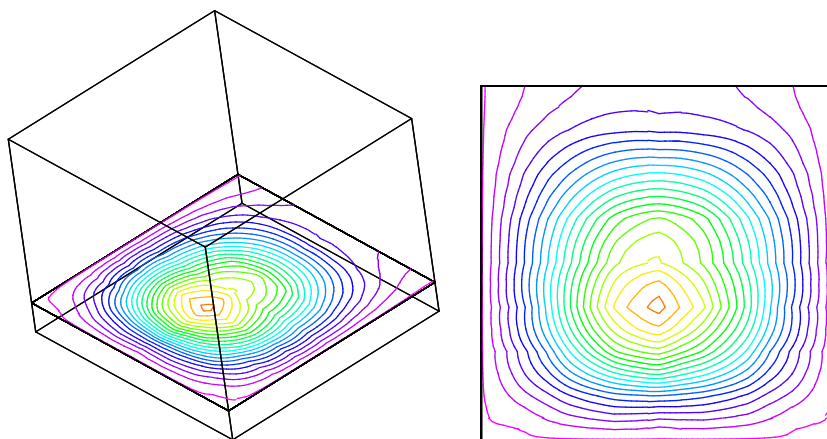


Abbildung 106: Temperaturverteilung 10% der Gesamthöhe des Reaktors über dem Suszeptor

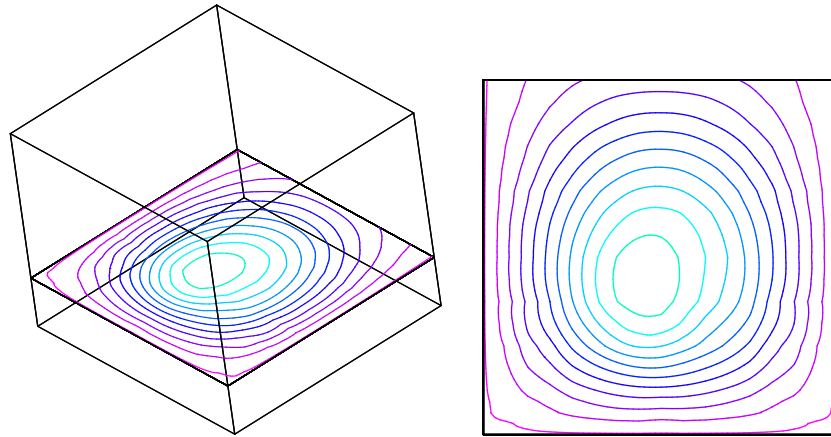


Abbildung 107: Temperaturverteilung 20% der Gesamthöhe des Reaktors über dem Suszeptor

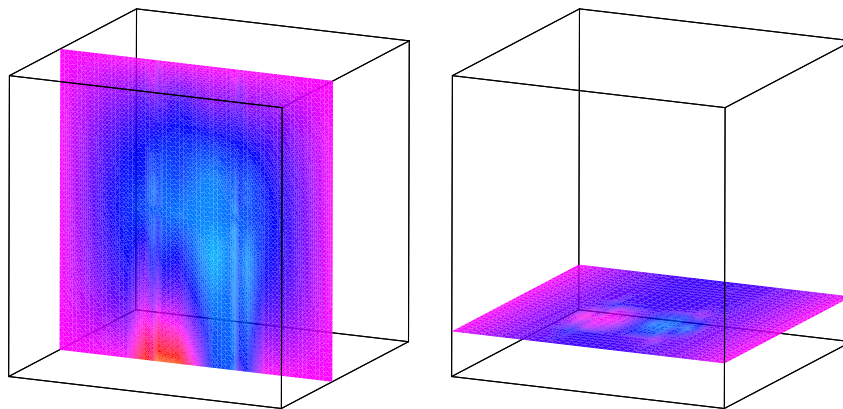


Abbildung 108: Geschwindigkeitskomponente  $u$  in Hauptstromrichtung für  $x = 0.40$  sowie für  $z = 0.20$

## 8 Abschließende Bemerkungen

In dieser Arbeit wurde die Methode der Finiten Differenzen auf dünnen Gittern eingeführt. Die vorgestellte Konstruktion Finiter Differenzen Operatoren auf dünnen Gitter führt unter Beachtung spezieller Glattheitsvoraussetzungen auf ein konsistentes und konvergentes Diskretisierungsverfahren auf dünnen Gittern. Theoretische Aussagen und numerische Berechnungen belegen, daß die vorgestellte Methode zur Berechnung partieller Differentialgleichungen bezüglich der Konvergenz vergleichbar ist mit der Methode der Finiten Elemente auf dünnen Gittern [Zen90, Gri91a, BaZe92, Bun92, Bal94, BDZ96, Pf96, BuDo97, Dor97, Bun98]. Berücksichtigt man die notwendigen Glattheitsvoraussetzungen, so ist im Vergleich zu vollen Gittern das Verhältnis von Aufwand zu Genauigkeit für die Berechnung einer Lösung einer partiellen Differentialgleichung mittels Finiter Differenzen bei dünnen Gittern erheblich besser, so daß dünne Gitter hier vollen Gittern überlegen sind.

Basierend auf dem Konzept einer Algebra für Funktionen auf dünnen Gittern wurde die Diskretisierung elliptischer und parabolischer Differentialgleichungen mittels der Methode der Finiten Differenzen realisiert. Die Verwendung eines Jacobi-Vorkonditionierer in einem iterativen Löser wie dem hier verwendeten BiCGSTAB gewährleistet eine schnelle und effiziente Lösung der auftretenden linearen Gleichungssysteme. Durch die Wahl der Datenstruktur Hash-Tabelle, der Parallelisierung der Operatoren sowie dem Einsatz von Gebietszerlegungstechniken und Gebietstransformationen steht somit auf dünnen Gittern eine einfach programmierbare Methode zur Lösung von PDE's zur Verfügung. Dabei erlaubt gerade die Behandlung allgemeiner Gebiete den Schritt von akademischen Testbeispielen hin zu "real world"-Problemen. Durch das Bereitstellen dieser Komponenten steht durch diese Arbeit nun auf dem Gebiet der dünnen Gittern neben der Methode der Finiten Elemente auch die Methode der Finiten Differenzen zur Diskretisierung zur Verfügung.

Aus anwendungsorientierter Sicht ist der Einsatz sowohl der Methode der Finiten Elemente als auch der Methode der Finiten Differenzen auf dünnen Gittern bei realitätsnahen Problemen voranzutreiben. Ein erstes Beispiel für den Einsatz der Methode der Finiten Differenzen auf dünnen Gittern ist die Simulation von CVD-Prozessen, für die in dieser Arbeit erstmals Ergebnisse auf dünnen Gittern berechnet worden sind. Speziell in diesem Beispiel zeigt sich, daß adaptive Gitter unerlässlich sind, wenn realistische Problemstellungen angegangen werden. Daher müssen effiziente und zuverlässige Fehlerschätzer zur selbst-adaptiven Steuerung bei Aufgabenstellungen mit Sprüngen oder Singularitäten einer Lösung entwickelt werden. Das in dieser Arbeit vorgestellte Konzept zur selbst-adaptiven Steuerung zeigt, daß es un-

ter anderem für das gewählte Beispiel L2 (vgl. Abschnitt 7.2.1) bis auf die Maximumnorm befriedigende Ergebnisse liefert. Eine theoretische Untermauerung dieses Vorgehens ist von großem Interesse, beispielsweise ein Beweis, daß es sich hierbei um einen Fehlerindikator in einer bestimmten Norm handelt. Bislang wurde die Methode der Finiten Differenzen auf dünnen Gittern lediglich für Probleme mit glatten Koeffizientenfunktionen eingesetzt. Da beispielsweise bei der Simulation von Strömungen in porösen Medien (Anwendungsbeispiel: Erdöllagerstätten) springende Diffusionskoeffizienten auftreten, ist es erforderlich, geeignete Finite Differenzen Operatoren auf dünnen Gittern auch für solche Koeffizientenfunktionen mit Sprüngen zu entwickeln. Die Navier Stokes Gleichungen wurden im Rahmen dieser Arbeit lediglich für kleine Reynoldszahlen ( $Re \leq 100$ ) numerisch behandelt. Für Untersuchungen zur Aerodynamik von Flugzeugen oder Fahrzeugen (um nur ein Beispiel zu nennen) ist jedoch die Entwicklung numerisch stabiler Algorithmen auch für hohe Reynoldszahlen für die Methode der Finiten Differenzen auf dünnen Gittern notwendig. Bisherige Verfahren benötigen für realitätsnahe Simulationen eine große Zahl von Gitterpunkten, so daß dünne Gitter hier helfen könnten, den erforderlichen Speicherplatz und gleichzeitig die notwendige Rechenzeit für diese Anwendungen zu senken.

Aus theoretischer Sicht gilt es, die Matrix des diskreten Laplace-Operators in seiner Darstellung mittels Finiter Differenzen auf dünnen Gittern weiter zu untersuchen. So beruhen einige in dieser Arbeit getroffene Aussagen, beispielsweise Stabilitätsaussagen, derzeit noch auf numerischen Berechnungen. Diese sollten theoretisch abgesichert werden, auch wenn die numerischen Ergebnisse nur geringe Zweifel an der Richtigkeit der getroffenen Aussagen aufkommen lassen. Theoretische Untersuchungen werden jedoch durch die notwendige Hierarchisierung beziehungsweise Enthierarchisierung im Aufbau der Finite Differenzen Operatoren auf dünnen Gittern erschwert. Hierdurch ist es – um nur ein Beispiel zu nennen – bislang nicht gelungen, die Eigenwerte und Eigenfunktionen in einer geschlossenen Form anzugeben, wie dies etwa auf vollen Gittern getan werden kann. Bei Anwendungen aus dem Bereich der parabolischen Differentialgleichungen muß analog zur Situation auf vollen Gittern eine CFL-Bedingung eingehalten werden. In vielen praktischen Versuchen hat sich dabei herausgestellt, daß es oftmals ausreicht, bei einer Verfeinerung der Ortsschrittweite (Übergang eines regulären dünnen Gitters vom Level  $l$  auf ein reguläres dünnes Gitter vom Level  $l+1$ ) die Zeitschrittweite mit dem Faktor  $\sqrt{2}$  zu skalieren. Dies ist jedoch lediglich eine Beobachtung und kann zum derzeitigen Zeitpunkt nicht theoretisch nachgewiesen werden.

Zusammenfassend erweist sich die Methode der Finiten Differenzen auf dünnen Gittern als eine Diskretisierungstechnik, die bei vergleichsweise geringem Speicherbedarf und einfacher Programmierung Ergebnisse gleicher Qualität – bei Problemstellungen mit glatten Daten – liefert wie etwa Verfahren auf

vollen Gittern. Die oben angesprochenen Fragen müssen im Rahmen dieser Arbeit leider ungeklärt bleiben. Diese sollen jedoch in künftigen Arbeiten ausführlich behandelt werden.

## A Beweise zu Sätzen vorhergehender Kapitel

In diesem Abschnitt beweisen wir diejenigen Lemmata und Sätze dieser Arbeit, deren Beweise eine gewisse Länge überschreiten und somit dem Lesefluß der Arbeit hinderlich sind. Nach den betroffenen Sätzen sind innerhalb der Kapitel Hinweise auf den jeweiligen Beweis hier im Anhang gegeben. Nachfolgend wird die Notation  $\mathbf{e}_i$  für Einheitsvektoren bezüglich der Raumdimension  $i$  gebraucht. Gelegentlich findet auch die Notation  $\mathbf{e}_x$  Verwendung, wenn der Einheitsvektor bezüglich der Koordinatenrichtung  $x$  betrachtet wird.

### A.1 Beweis zu Lemma 4.2

**Beweis:** Es genügt, den Beweis im zweidimensionalen Fall auf dem Einheitsquadrat  $\bar{\Omega} = [0, 1]^2$  für den Stern (4.5) zu betrachten, für die übrigen Sterne und für höhere Dimensionen überträgt sich der Beweis analog.

Einige der kritischen Punkte sind  $\tilde{\mathbf{x}}_{\mathbf{l},i} = (x_{L,i}, y_{1,1})$  und  $\hat{\mathbf{x}}_{\mathbf{l},i} = (x_{1,1}, y_{L,i})$ , wobei  $L = |\mathbf{l}|_\infty$  der feinste auftretende Level des zugrundeliegenden (regulären) dünnen Gitters sei sowie  $i$  jeweils geeignet gewählt ist. Wendet man den Stern (4.5) in  $y$ -Richtung an einem der kritischen Punkte  $\tilde{\mathbf{x}}_{\mathbf{l},i}$  an, so erhält man

$$\delta_{yy}(\tilde{\mathbf{x}}_{\mathbf{l},i}) = \sum_{i=1}^{\infty} \frac{2h_y^{2(i-1)}}{(2i)!} \frac{\partial^{2i} u(\tilde{\mathbf{x}}_{\mathbf{l},i})}{\partial y^{2i}} \quad (9.1)$$

mit der (lokalen) Maschenweite  $h_y = \frac{1}{2}$  als Diskretisierung für die zweite Ableitung in  $y$ -Richtung. Analoges gilt für die entsprechende Formel in  $x$ -Richtung an den Punkten  $\hat{\mathbf{x}}_{\mathbf{l},i}$ . Ersetzt man dieses  $h_y$  in der angegebenen Formel durch  $\frac{1}{2}$ , so erhält man

$$\begin{aligned} \delta_{yy}(\tilde{\mathbf{x}}_{\mathbf{l},i}) &= \sum_{i=1}^{\infty} \frac{2 \left(\frac{1}{2}\right)^{2(i-1)}}{(2i)!} \frac{\partial^{2i} u(\tilde{\mathbf{x}}_{\mathbf{l},i})}{\partial y^{2i}} \\ &= \frac{\partial^2 u(\tilde{\mathbf{x}}_{\mathbf{l},i})}{\partial y^2} + \frac{1}{48} \cdot \frac{\partial^4 u(\tilde{\mathbf{x}}_{\mathbf{l},i})}{\partial y^4} + \sum_{i=3}^{\infty} \frac{2 \left(\frac{1}{2}\right)^{2(i-1)}}{(2i)!} \frac{\partial^{2i} u(\tilde{\mathbf{x}}_{\mathbf{l},i})}{\partial y^{2i}} . \end{aligned}$$

Damit existieren mindestens  $2^{L-1}$  Gitterpunkte in einem dünnen Gitter in zwei Dimensionen, die bezüglich des gewählten Differenzen-Sterns für die Berechnung des Laplace-Operators (bzw. der gewichteten Addition zweiter Ableitungen, was einem Diffusions-Operator entspricht) einen Diskretisierungsfehler der Ordnung  $\mathcal{O}(1)$  besitzen. In zwei Dimensionen hat man folglich mindestens  $2 \cdot 2^{L-1}$  Punkte, die für den Laplace-Operator einen Fehler der Größenordnung  $\mathcal{O}(1)$  besitzen. In  $d$  Dimensionen sind es mindestens



$d \cdot 2^{l-1}$ , an denen der Konsistenzfehler für den  $d$ -dimensionalen Laplace-Operator die Ordnung  $\mathcal{O}(1)$  besitzt. Folglich ist die Diskretisierung nicht konsistent.  $\square$

## A.2 Beweis zu Satz 4.1

**Beweis:** Wir betrachten hier nur den Beweis in zwei Raumdimensionen und verweisen für höhere Dimensionen auf die Sätze 4.5 und 4.7 und deren Beweise. Im  $2D$ -Fall beziehen sich die Hierarchisierung und Enthierarchisierung in der Darstellung der zweiten Ableitung auf dünnen Gittern mittels Finiten Differenzen lediglich auf  $2 - 1 = 1$  Raumdimension, wodurch der Beweis notationstechnisch einfacher ist als für beliebige Dimensionen. Die Schwierigkeiten, die im Beweis für den mehrdimensionalen Fall auftreten, sind rein notationstechnischer Natur, nicht jedoch in der zugrundeliegenden Mathematik selbst begründet. An dieser Stelle sei angemerkt, daß die Punktmenge eines (nicht-adaptiven) dünnen Gitters bezüglich dem Punkt  $\mathbf{x}_{l,i}$  mit  $\mathbf{l} = \mathbf{i} = \{1, \dots, 1\}$ , der stets ein Punkt des dünnen Gitters ist, punktsymmetrisch ist. Folglich kann man den Beweis für beliebige Raumdimensionen  $j$  auf den Beweis der Aussage für den Operator  $\mathbf{D}_{11}^S$  einschränken. Um unnötige Indizes zu vermeiden, seien in dem Beweis für den zweidimensionalen Fall die Identitäten  $x \equiv x_1$  und  $y \equiv x_2$  vereinbart. Zudem reicht es, den Fall  $\bar{\Omega}^2 = [0, 1]^2$  bzw. allgemein  $\bar{\Omega}^d = [0, 1]^d$  zu betrachten.

Der Operator für eine zweite Ableitung auf dünnen Gittern setzt sich aus drei Einzeloperatoren zusammen. In Abbildung 109 kann man sehen, wie die drei Operatoren auf Gitterpunkten wirken. In dieser Abbildung stehen  $x$  und  $y$  für beliebige, aber feste innere Punkte eines dünnen Gitters. Dabei erkennt man, daß durch die Anwendung der ersten beiden Operatoren für innere Punkte ein 9-Punkt-Stern der Bauart

$$\begin{aligned} & \frac{1}{2^{-2l_1}} \left[ \begin{array}{ccc} -\frac{1}{2} & 1 & -\frac{1}{2} \\ 1 & -2 & 1 \\ -\frac{1}{2} & 1 & -\frac{1}{2} \end{array} \right]_{\mathbf{x}_{l,i}, \mathbf{h}_1^{local}} \circ u(\mathbf{x}) & (9.2) \\ & = \frac{1}{2^{-2l_1}} \left( u(x_{l_1, i_1} - 2^{-l_1}, y_{l_2, i_2}) + u(x_{l_1, i_1} + 2^{-l_1}, y_{l_2, i_2}) + \right. \\ & \quad u(x_{l_1, i_1}, y_{l_2, i_2} - h_y) + u(x_{l_1, i_1}, y_{l_2, i_2} + h_y) - \\ & \quad \frac{1}{2} \cdot \left( u(x_{l_1, i_1} - 2^{-l_1}, y_{l_2, i_2} - h_y) + u(x_{l_1, i_1} - 2^{-l_1}, y_{l_2, i_2} + h_y) + \right. \\ & \quad \left. u(x_{l_1, i_1} + 2^{-l_1}, y_{l_2, i_2} - h_y) + u(x_{l_1, i_1} + 2^{-l_1}, y_{l_2, i_2} + h_y) \right) - \\ & \quad \left. 2 \cdot u(x_{l_1, i_1}, y_{l_2, i_2}) \right) \quad , \end{aligned}$$

entsteht, wobei  $h_y = x_{l_2, i_2} - \mathcal{H}_{j_2}^-(x_{l_2, i_2}) \equiv \mathcal{H}_{j_2}^+(x_{l_2, i_2}) - x_{l_2, i_2}$  die (punkt-

abhängige) Schrittweite zum hierarchischen Vater in  $y$ -Richtung ist. Desweiteren sei  $h_x = 2^{-l}$  die (punktabhängige) Schrittweite zum physikalisch nächsten Gitterpunkt in  $x$ -Richtung und  $h$  die feinste im dünnen Gitter auftretende Maschenweite (Randmaschenweite)<sup>18</sup>. Für Randpunkte bezüglich der Koordinatenrichtung  $y$  stellt die Hierarchisierung sowie die Enthierarchisierung in  $y$ -Richtung die Identität dar, womit sich an diesen Gitterpunkten der bekannte eindimensionale Stern für die Berechnung der zweiten Ableitung einstellt.

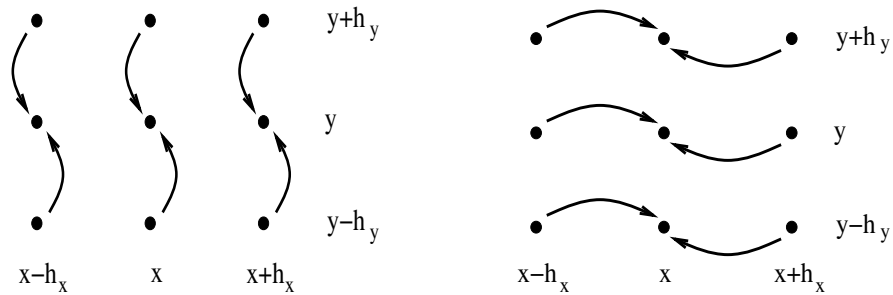


Abbildung 109: Wirkung der Hierarchisierung bzw. Enthierarchisierung in  $y$ -Richtung (links) und der zweiten Ableitung in  $x$ -Richtung (rechts)

Die Beweise der Konsistenzordnungen für Finite Differenzen Operatoren auf dünnen Gittern benötigen wie ihre Analogons im Falle voller Gitter Taylorreihen. Wendet man die Hierarchisierung in  $y$ -Richtung an, die den ersten der drei Operatoren des Operators  $\mathbf{D}_{11}^S$  darstellt, so erhält man folgende Taylorreihen für innere Gitterpunkte:

- an  $\mathbf{x}_{1,i} - h_x \mathbf{e}_x$  :

$$\begin{aligned}
 \tilde{u}(\mathbf{x}_{1,i} - h_x \mathbf{e}_x) &:= u(\mathbf{x}_{1,i} - h_x \mathbf{e}_x) - \frac{1}{2} \left( u(\mathbf{x}_{1,i} - h_x \mathbf{e}_x + h_y \mathbf{e}_y) \right. \\
 &\quad \left. + u(\mathbf{x}_{1,i} - h_x \mathbf{e}_x - h_y \mathbf{e}_y) \right) \\
 &= -\frac{1}{2} \sum_{j=0}^{\infty} \frac{2}{(2(j+1))!} h_y^{2(j+1)} u^{2(j+1),y}(\mathbf{x}_{1,i} - h_x \mathbf{e}_x) \\
 &= -\frac{1}{2} \sum_{j=1}^{\infty} \frac{2}{(2j)!} h_y^{2j} u^{2j,y}(\mathbf{x}_{1,i} - h_x \mathbf{e}_x) \quad (9.3)
 \end{aligned}$$

<sup>18</sup> Legt man ein volles Gitter mit den äquidistanten Maschenweiten  $h_x$  und  $h_y$  zugrunde, stellt der Stern (9.2) einen kompakten 9-Punkt-Stern dar. Im dünnen Gitter, das dem Beweis zugrundeliegt, gilt diese Aussage jedoch nicht.

- an  $\mathbf{x}_{1,i}$  :

$$\begin{aligned}
\tilde{u}(\mathbf{x}_{1,i}) &:= u(\mathbf{x}_{1,i}) - \frac{1}{2} \left( u(\mathbf{x}_{1,i} + h_y \mathbf{e}_y) + u(\mathbf{x}_{1,i} - h_y \mathbf{e}_y) \right) \\
&= -\frac{1}{2} \sum_{j=0}^{\infty} \frac{2}{(2(j+1))!} h_y^{2(j+1)} u^{2(j+1),y}(\mathbf{x}_{1,i}) \\
&= -\frac{1}{2} \sum_{j=1}^{\infty} \frac{2}{(2j)!} h_y^{2j} u^{2j,y}(\mathbf{x}_{1,i}) \tag{9.4}
\end{aligned}$$

- an  $\mathbf{x}_{1,i} + h_x \mathbf{e}_x$  :

$$\begin{aligned}
\tilde{u}(\mathbf{x}_{1,i} + h_x \mathbf{e}_x) &:= u(\mathbf{x}_{1,i} + h_x \mathbf{e}_x) - \frac{1}{2} \left( u(\mathbf{x}_{1,i} + h_x \mathbf{e}_x + h_y \mathbf{e}_y) + \right. \\
&\quad \left. u(\mathbf{x}_{1,i} + h_x \mathbf{e}_x - h_y \mathbf{e}_y) \right) \\
&= -\frac{1}{2} \sum_{j=0}^{\infty} \frac{2}{(2(j+1))!} h_y^{2(j+1)} u^{2(j+1),y}(\mathbf{x}_{1,i} + h_x \mathbf{e}_x) \\
&= -\frac{1}{2} \sum_{j=1}^{\infty} \frac{2}{(2j)!} h_y^{2j} u^{2j,y}(\mathbf{x}_{1,i} + h_x \mathbf{e}_x) \tag{9.5}
\end{aligned}$$

Für Gitterpunkte am Rand mit  $l_2 = 0$  stellt – wie zuvor bereits bemerkt – die Hierarchisierung die Identität dar und somit erhalten wir  $\tilde{u}(\mathbf{x}_{1,i}) = u(\mathbf{x}_{1,i}) = H_1 \circ u(\mathbf{x}_{1,i})$  für alle Gitterpunkte  $\mathbf{x}_{1,i}$  mit  $l_2 = 0$ . Kombiniert man die Resultate (9.3), (9.4) sowie (9.5) der Hierarchisierung durch den Stern für die zweite Ableitung, so erhalten wir

$$\begin{aligned}
v(\mathbf{x}_{1,i}) &= \frac{\tilde{u}(\mathbf{x}_{1,i} - h_x \mathbf{e}_x) - 2\tilde{u}(\mathbf{x}_{1,i}) + \tilde{u}(\mathbf{x}_{1,i} + h_x \mathbf{e}_x)}{h_x^2} \\
&= -\frac{1}{2h_x^2} \left( \sum_{m=0}^{\infty} \sum_{j=1}^{\infty} \frac{(-h_x)^m}{m!} \frac{2}{(2j)!} h_y^{2j} u^{m,x,2j,y}(\mathbf{x}_{1,i}) - \right. \\
&\quad \left. -2 \cdot \sum_{j=1}^{\infty} \frac{2}{(2j)!} h_y^{2j} u^{2j,y}(\mathbf{x}_{1,i}) + \right. \\
&\quad \left. + \sum_{m=0}^{\infty} \sum_{j=1}^{\infty} \frac{h_x^m}{m!} \frac{2}{(2j)!} h_y^{2j} u^{m,x,2j,y}(\mathbf{x}_{1,i}) \right) \\
&= -\frac{1}{2h_x^2} \left( \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-h_x)^m}{m!} \frac{2}{(2j)!} h_y^{2j} u^{m,x,2j,y}(\mathbf{x}_{1,i}) + \right. \\
&\quad \left. + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h_x^m}{m!} \frac{2}{(2j)!} h_y^{2j} u^{m,x,2j,y}(\mathbf{x}_{1,i}) \right)
\end{aligned}$$

$$= -\frac{1}{2} \sum_{m=0}^{\infty} \sum_{j=1}^{\infty} \frac{2h_x^{2m}}{(2(m+1))!} \frac{2h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i})$$

als Ergebnis des Operators  $\mathbf{D}_{11} \circ \mathbf{H}_1$  an dem Dünngitterpunkt  $\mathbf{x}_{1,i}$ . An Randpunkten mit  $l_2 = 0$  erhalten wir das bekannte Ergebnis

$$\begin{aligned} v(\mathbf{x}_{1,i}) &= \sum_{m=0}^{\infty} \frac{2h_x^{2m}}{(2(m+1))!} u^{2(m+1),x}(\mathbf{x}_{1,i}) \\ &= u^{2,x}(\mathbf{x}_{1,i}) + \sum_{m=1}^{\infty} \frac{2h_x^{2m}}{(2(m+1))!} u^{2(m+1),x}(\mathbf{x}_{1,i}) \end{aligned}$$

als Resultat von  $\mathbf{D}_{11} \circ \mathbf{H}_1$ . Die so erhaltenen Summen müssen jetzt nur noch vermöge des Enthierarchisierungsoperators  $\mathbf{E}_1$  miteinander kombiniert werden. Wie bereits im Fall der Hierarchisierung ist die Wirkung der Enthierarchisierung in  $y$ -Richtung für Randpunkte mit  $l_2 = 0$  analog der Wirkung der Identität. Damit ist die Konsistenzordnung zwei für diese Punkte bereits nachgewiesen, da für diese Punkte  $h_x \equiv h$  gilt.

Es reicht folglich aus, an dieser Stelle nur noch innere Gitterpunkte zu betrachten. Hierfür führen wir zu Zwecken der kürzeren Schreibweise die Größen  $\mathcal{F}(\cdot)$  ein. Diese beinhalten nur Terme der Größenordnung  $\mathcal{O}(h^2)$ . Ein Nachweis dieser Ordnung an hierarchischen Vätern eines Punktes  $\mathbf{x}_{1,i}$  ist im Beweis nicht nötig, da der Beweis so aufgebaut ist, daß die Konsistenzordnung an hierarchischen Vaterknoten bereits bewiesen ist, somit auch die Aussage für die  $\mathcal{F}(\cdot)$ -Terme, wie sich im weiteren Verlauf des Beweises herausstellen wird. Es muß also nur der Nachweis  $\mathcal{F}(\mathbf{x}_{1,i}) = \mathcal{O}(h^2)$  für den soeben betrachteten Gitterpunkt erbracht werden. Wendet man nun die Enthierarchisierung in  $y$ -Richtung auf die zuvor berechneten Ergebnisse an, so erhalten wir

$$\begin{aligned} \mathbf{E}_1 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) &= \\ &= \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) + \frac{1}{2} \left( \mathbf{D}_{11} \circ \mathbf{H}_1 \circ \left( u(\mathcal{H}_1^-(\mathbf{x}_{1,i})) + u(\mathcal{H}_1^+(\mathbf{x}_{1,i})) \right) \right) \\ &= - \sum_{m=0}^{\infty} \sum_{j=1}^{\infty} \frac{2 \cdot (2^{l_2-1}h)^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) + \\ &\quad + \frac{1}{2} \left( u^{2,x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \right. \\ &\quad \left. \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} u^{2(m+1),x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) + \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \left( u^{2,x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \right. \\
& \quad \left. \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} u^{2(m+1),x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right) \\
= & - \sum_{m=0}^{\infty} \sum_{j=1}^{\infty} \frac{2 \cdot (2^{l_2-1}h)^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) + \\
& + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{(-h_y)^j}{j!} u^{2,x,j,y}(\mathbf{x}_{1,i}) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right. \\
& \quad \left. + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \sum_{j=0}^{\infty} \frac{(-h_y)^j}{j!} u^{2(m+1),x,j,y}(\mathbf{x}_{1,i}) \right) + \\
& + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{h_y^j}{j!} u^{2,x,j,y}(\mathbf{x}_{1,i}) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right. \\
& \quad \left. + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \sum_{j=0}^{\infty} \frac{h_y^j}{j!} u^{2(m+1),x,j,y}(\mathbf{x}_{1,i}) \right) + \\
= & - \left( \sum_{j=1}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{2,x,2j,y}(\mathbf{x}_{1,i}) + \right. \\
& \quad \left. + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{2 \cdot (2^{l_2-1}h)^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \right) + \\
& + \sum_{j=0}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{2,x,2j,y}(\mathbf{x}_{1,i}) + \frac{1}{2} (\mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i}))) + \\
& + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} u^{2(m+1),x}(\mathbf{x}_{1,i}) + \\
& + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \sum_{j=1}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \\
= & u^{2,x}(\mathbf{x}_{1,i}) + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} u^{2(m+1),x}(\mathbf{x}_{1,i}) + \mathcal{F}(\mathbf{x}_{1,i}) \tag{9.6}
\end{aligned}$$

mit dem Ausdruck

$$\begin{aligned} \mathcal{F}(\mathbf{x}_{\mathbf{l},\mathbf{i}}) &= \tag{9.7} \\ &= \frac{1}{2} (\mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{\mathbf{l},\mathbf{i}})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{\mathbf{l},\mathbf{i}}))) + \\ &\quad + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} \left(1 - (2^{l_2-1})^{2m}\right) u^{2(m+1),x,2j,y}(\mathbf{x}_{\mathbf{l},\mathbf{i}}) \end{aligned}$$

Im Fall  $l_2 = 0$  setzen wir  $\mathcal{F}(\mathbf{x}_{\mathbf{l},\mathbf{i}}) \equiv 0$ . Damit ist die Aussage des Satzes bewiesen, wenn der Term  $\mathcal{F}(\mathbf{x}_{\mathbf{l},\mathbf{i}})$  von der Größenordnung  $\mathcal{O}(h^2)$  ist, da die Summe in (9.6) den führenden Term  $h^2 u^{4,x}(\mathbf{x}_{\mathbf{l},\mathbf{i}})/12$  besitzt.

In einem dünnen Gitter in  $2D$  ist an jedem Gitterpunkt die Identität

$$h_x \cdot h_y = \frac{1}{2} h \tag{9.8}$$

gültig, wobei  $h_x = 2^{-l_1}$  die (lokale) Maschenweite zum nächsten (linken oder rechten) Nachbarknoten in  $x$ -Richtung und  $h_y$  die (lokale) Maschenweite zum (linken oder rechten) hierarchischen Vaterknoten in  $y$ -Richtung darstellt<sup>19</sup>. Sei nun  $l = |\mathbf{l}|_{\infty}$  der maximale auftretende Level des aktuellen dünnen Gitters. Dann folgt hieraus die Identität  $h = 2^{-l}$  und es gilt zudem  $h_y = \frac{1}{2} \cdot h \cdot h_x^{-1} = 2^{(l_1-l)-1}$ . An dieser Stelle sei außerdem nochmals auf die Definition eines dünnen Gitters in verwiesen, die nachfolgend verwendet wird. Wir betrachten nun den Fehlerterm  $\mathcal{F}(\mathbf{x}_{\mathbf{l},\mathbf{i}})$  genauer um herauszufinden, welche Koeffizienten vor welchen Potenzen von  $h$  stehen. Dabei führen wir eine Fallunterscheidung mit vier verschiedenen Fällen durch:

1. Fall: hier betrachten wir Terme der Form

$$\frac{2h^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{\mathbf{l},\mathbf{i}}) \quad , \tag{9.9}$$

die gerade den Anteil des Fehlerterms  $\mathcal{F}(\mathbf{x}_{\mathbf{l},\mathbf{i}})$  ausmachen, bei dem in der Klammer der Faktor 1 genommen wird. Bei diesen Termen ist die  $h$ -Potenz direkt ablesbar.

2. Fall: hier betrachten wir die übrigen Terme

$$- \frac{2h^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} (2^{l_2-1})^{2m} u^{2(m+1),x,2j,y}(\mathbf{x}_{\mathbf{l},\mathbf{i}}) \quad , \tag{9.10}$$

---

<sup>19</sup>An dieser Stelle sei nochmals darauf verwiesen, daß der Beweis ausschließlich die zweite Ableitung bezüglich der  $x$ -Richtung betrachtet und die Bezeichnungen  $h_x$  und  $h_y$  daher in dieser Weise verwendet werden können. Betrachtet man hingegen die zweite Ableitung bezüglich der  $y$ -Richtung, so vertauschen sich die Bedeutungen: dann wäre  $h_x$  die (lokale) Maschenweite bezüglich hierarchischer Vaterknoten in  $x$ -Richtung, und  $h_y$  wäre die (lokale) Maschenweite bezüglich nächster Nachbarn in  $y$ -Richtung.

bei denen sich im Gegensatz zu dem Ausdruck in (9.9) die auftretenden Potenzen bezüglich  $h$  nicht ablesen lassen, womit eine genauere Analyse der Terme erforderlich wird. Dazu unterscheiden wir die 3 Fälle  $j = m$ ,  $j < m$  und  $j > m$ .

- Fall 2a: sei  $j = m$ . Dann gilt

$$\begin{aligned}
& \frac{2h^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} (2^{l_2-1})^{2m} = \\
&= \frac{2}{(2(m+1))!(2j)!} \cdot 2^{-2lm+2l_1j-2lj-2j+2ml_2-2m} \\
&= \frac{2}{(2(m+1))!(2j)!} \cdot \left(2^{2(l_1+l_2-l-1)}\right)^m \cdot \left(2^{-l}\right)^{2m} \cdot 2^{-2m} \\
&= \frac{2}{(2(m+1))!(2j)!} \cdot \underbrace{\left(2^{2(l_1+l_2-l-1)}\right)^m}_{0 \leq \dots \leq 1} \cdot 2^{-2m} \cdot h^{2m} \\
&\leq \frac{2}{(2(m+1))!(2j)!} \cdot 2^{-2m} \cdot h^{2m} \tag{9.11}
\end{aligned}$$

- Fall 2b: sei  $j < m$ . Dann gilt

$$\begin{aligned}
& \frac{2h^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} (2^{l_2-1})^{2m} = \\
&= \frac{2 \cdot 2^{-2m} \cdot (2^{-2l})^j}{(2(m+1))!(2j)!} \cdot \left(2^{2(l_1+l_2-l-1)}\right)^j \cdot \left(2^{-2(l-l_2)}\right)^{m-j} \\
&= \frac{2 \cdot 2^{-2m} \cdot h^{2j}}{(2(m+1))!(2j)!} \cdot \underbrace{\left(2^{2(l_1+l_2-l-1)}\right)^j}_{0 \leq \dots \leq 1} \cdot \underbrace{\left(2^{-2(l-l_2)}\right)^{m-j}}_{0 \leq \dots \leq 1} \\
&\leq \frac{2}{(2(m+1))!(2j)!} \cdot 2^{-2m} \cdot h^{2j} \tag{9.12}
\end{aligned}$$

- Fall 2c: sei  $j > m$ . Dann gilt

$$\begin{aligned}
& \frac{2h^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} (2^{l_2-1})^{2m} = \\
&= \frac{2 \cdot 2^{-2j} \cdot (2^{-2l})^m}{(2(m+1))!(2j)!} \cdot \left(2^{2(l_1+l_2-l-1)}\right)^m \cdot \left(2^{-2(l-l_1)}\right)^{j-m} \\
&= \frac{2 \cdot 2^{-2j} \cdot h^{2m}}{(2(m+1))!(2j)!} \cdot \underbrace{\left(2^{2(l_1+l_2-l-1)}\right)^m}_{0 \leq \dots \leq 1} \cdot \underbrace{\left(2^{-2(l-l_1)}\right)^{j-m}}_{0 \leq \dots \leq 1} \\
&\leq \frac{2}{(2(m+1))!(2j)!} \cdot 2^{-2j} \cdot h^{2m} \tag{9.13}
\end{aligned}$$

Aus der Formel (9.7) ist ersichtlich, daß die Werte der Ableitungen an den jeweiligen Dünngitter-Punkten zu nehmen sind und somit die Abschätzung punktabhängig ist. In den Formeln (9.11), (9.12) und (9.13) wird die Abhängigkeit von der Lage der Gitterpunkte offensichtlich, wenn man die jeweiligen Konstanten bei den Potenzen in  $h$  betrachtet. Um die Fehlerbehandlung jedoch allgemein für alle Punkte durchzuführen, betrachten wir nachfolgend den "worst case" für die Abschätzung, der dann eintritt, wenn  $l_1 + l_2 - l - 1 = 0$  ist und man an geeigneter Stelle jeweils  $l_i = l$  für  $1 \leq i \leq 2$  annimmt. Dabei betrachten wir als "worst case" denjenigen Fall, in dem die auftretenden Terme ordnungstechnisch gesehen jeweils am größten werden.

Folgende Terme tragen zur Konstanten vor der Potenz  $h^2$  bei, wobei die Ableitungsterme vernachlässigt werden. Da wir  $u \in C^\infty(\bar{\Omega})$  vorausgesetzt haben, können wir die auftretenden Ableitungsterme von  $u$  in (9.7) durch  $\|u\|_{C^\infty(\bar{\Omega})}$  abschätzen und brauchen diese somit im weiteren Verlauf nicht mehr berücksichtigen.

- Beitrag aus Fall 1: setzt man  $m = 1$  und addiert dann über  $j$ , so erhält man

$$h^2 \frac{2}{4!} (\cosh(h_y) - 1)$$

Dabei ist zu beachten, daß  $0 \leq h_y \leq 1$  für innere Gitterpunkte gilt und wir somit  $\cosh(h_y) \leq \sqrt[8]{\pi} \leq \frac{\pi}{2}$  abschätzen können. Damit ist der Ausdruck in der Klammer unabhängig von  $h$ .

- Beitrag aus Fall 2a: setzt man  $m = j = 1$ , so erhält man

$$\frac{2}{2!} \frac{h^2}{4} = \frac{h^2}{4}$$

- Beitrag aus Fall 2b: um  $h^2$  zu erhalten, müssen wir  $j = 1$  setzen, was  $m \geq 2$  zur Folge hat. Dann erhalten wir hieraus den Beitrag

$$\frac{2h^2}{2!} \sum_{m=2}^{\infty} \frac{2^{-2m}}{(2(m+1))!} = h^2 \left( \cosh\left(\frac{1}{2}\right) - 1 - \frac{1}{8} - \frac{1}{384} \right) = C \cdot h^2$$

mit  $C \neq C(h)$ .

- Beitrag aus Fall 2c: um  $h^2$  zu erhalten, müssen wir  $m = 1$  setzen, was  $j \geq 2$  zur Folge hat. Dann erhalten wir hieraus den Beitrag

$$\frac{2h^2}{4!} \sum_{j=2}^{\infty} \frac{2^{-2j}}{(2j)!} = \frac{h^2}{4!} \left( \cosh(h_y) - 1 - \frac{h_y^2}{2!} \right) < \frac{\pi}{8} h^2$$



Damit stehen bei allen vier Beiträgen Konstanten vor dem Term  $h^2$ , die unabhängig von  $h$  sind. Diese Vorgehensweise kann man nun für alle auftretenden (geraden) Potenzen von  $h$  durchführen. Listet man wie schon für die Potenz  $h^2$  die einzelnen Beiträge auf, so erhält man für die  $(2k)$ -te Potenz folgende Beiträge:

- Beitrag aus Fall 1: setzt man  $m = k$  und addiert dann über  $j$ , so erhält man

$$h^{2k} \frac{2}{(2(k+1))!} (\cosh(h_y) - 1) < \frac{2\pi h^{2k}}{2 \cdot (2(k+1))!}$$

- Beitrag aus Fall 2a: setzt man  $m = j = k$ , so erhält man

$$\frac{2}{(2(k+1))!} \frac{h^{2k}}{(2k)! 2^{2k}}$$

- Beitrag aus Fall 2b: um  $h^{2k}$  zu erhalten, müssen wir  $j = k$  setzen, was  $m \geq k + 1$  zur Folge hat. Dann erhalten wir hieraus den Beitrag

$$\begin{aligned} \frac{2h^{2k}}{(2k)!} \sum_{m=k+1}^{\infty} \frac{2^{-2m}}{(2(m+1))!} = \\ \frac{2h^{2k}}{(2k)!} \left( \cosh\left(\frac{1}{2}\right) - 1 - \frac{\left(\frac{1}{2}\right)^2}{2!} - \dots - \frac{\left(\frac{1}{2}\right)^{2(k+1)}}{(2(k+1))!} \right) \end{aligned}$$

- Beitrag aus Fall 2c: um  $h^{2k}$  zu erhalten, müssen wir  $m = k$  setzen, was  $j \geq k + 1$  zur Folge hat. Dann erhalten wir hieraus den Beitrag

$$\begin{aligned} \frac{2h^{2k}}{(2(k+1))!} \sum_{j=k+1}^{\infty} \frac{2^{-2j}}{(2j)!} = \\ \frac{2h^{2k}}{(2(k+1))!} \left( \cosh(h_y) - 1 - \frac{h_y^2}{2!} - \dots - \frac{h_y^{2k}}{(2k)!} \right) \end{aligned}$$

Damit sind die Faktoren im “worst case” vor jeder Potenz von  $h$  bestimmt, da in jeweils allen vier Fällen für jede Potenz von  $h$  die beteiligten Faktoren unabhängig von  $h$  sind.

Damit folgt nun insgesamt die Behauptung

$$\|\mathbf{D}_{ii}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^2 u}{\partial x_i^2}\|_{\infty} \leq C \cdot h^2 \|u\|_{C^{\infty}(\bar{\Omega})} \quad (9.14)$$

für die zweite Ableitung bezüglich der Raumdimension  $i$  in zwei Dimensionen ( $1 \leq i \leq 2$ ), was die Aussage des Satzes beweist.

Für entsprechende Beweise in höheren Dimensionen bzw. unter schwächeren Voraussetzungen sei an dieser Stelle auf die Sätze 4.2 und 4.4 sowie 4.5 und 4.7 verwiesen.  $\square$

### A.3 Beweis zu Satz 4.2

**Beweis:** Wir beschränken uns wie zuvor auf den Operator  $\mathbf{D}_{11}^S$ . Aus dem Beweis von Satz 4.1 und [Hac86], Korollar 4.5.4, folgt, daß an den Randpunkten  $\mathbf{x}_{1,i}$  mit  $l_2 = 0$  die Behauptung für Funktionen  $u \in C^4(\bar{\Omega})$  erfüllt ist, was eine schärfere Voraussetzung ist als die im Satz genannte. Wir erhalten für diese Randpunkte den Ausdruck

$$\mathbf{D}_{11}^S u(\mathbf{x}_{1,i}) = u^{2,x}(\mathbf{x}_{1,i}) + h^2 \cdot R(\mathbf{x}_{1,i})$$

mit

$$\begin{aligned} R(\mathbf{x}_{1,i}) &= h^{-4} \left( \int_{x_{1,i_1}}^{x_{1,i_1}+h} u^{4,x}(\xi, y_{l_2, i_2}) \cdot (x_{l_1, i_1} + h - \xi)^3 \frac{1}{3!} d\xi + \right. \\ &\quad \left. \int_{x_{1,i_1}}^{x_{1,i_1}-h} u^{4,x}(\xi, y_{l_2, i_2}) \cdot (x_{l_1, i_1} - h - \xi)^3 \frac{1}{3!} d\xi \right) \\ &= h^{-4} \left( \int_{x_{1,i_1}}^{x_{1,i_1} \pm h} u^{4,x}(\xi, y_{l_2, i_2}) \cdot (x_{l_1, i_1} \pm h - \xi)^3 \frac{1}{3!} d\xi \right) \end{aligned} \quad (9.15)$$

Damit folgt, daß wir am Rand die Abschätzung

$$\|\mathbf{D}_{11}^S \mathbf{R}_{1,i} u - \mathbf{R}_{1,i} \frac{\partial^2 u}{\partial x_1^2}\|_{\infty} \leq \frac{h^2}{12} \|u\|_{C^4(\bar{\Omega})} \quad (9.16)$$

bekommen, bei der wir anstatt  $\|u\|_{C^4(\bar{\Omega})}$  auch die Norm  $\|u\|_{C^{3,1}(\bar{\Omega})}$  nehmen dürfen, wenn wir  $u$  als 3-fach Lipschitz-stetige, differenzierbare Funktion annehmen (vgl. [Hac86], Korollar 4.5.5). Betrachtet man die Wirkung der  $(d-1)$ -dimensionalen Hierarchisierung und der nachfolgenden Ableitung, so kann man die Linearkombination von Funktionswerten (nichts anderes wird durch die Anwendung eines Sterns getan) geeignet zusammenfassen und interpretieren. Dadurch gelangt man zu folgender Interpretation für den aus den ersten beiden Teiloperatoren resultierenden Stern: in einem ersten Schritt berechnet man Finite Differenzen an den Stellen  $\mathbf{x}_{1,i}$ ,  $\mathcal{H}_2^-(\mathbf{x}_{1,i})$  und  $\mathcal{H}_2^+(\mathbf{x}_{1,i})$  bezüglich der lokalen Maschenweite an  $\mathbf{x}_{1,i}$  und transportiert vermöge des Hierarchisierungsoperators alle Ergebnisse an die Stelle  $\mathbf{x}_{1,i}$ . Dabei sei an dieser Stelle angemerkt, daß die lokale Maschenweite in Ableitungsrichtung an den Stellen  $\mathbf{x}_{1,i}$ ,  $\mathcal{H}_2^-(\mathbf{x}_{1,i})$  und  $\mathcal{H}_2^+(\mathbf{x}_{1,i})$  unterschiedlich sind und hier keine Vertauschung von Ableitungsoperator und  $(d-1)$ -dimensionaler Hierarchisierung stattfindet. An den hierarchischen Vaterknoten  $\mathcal{H}_2^+(\mathbf{x}_{1,i})$

und  $\mathcal{H}_2^-(\mathbf{x}_{1,i})$  erhalten wir demnach zunächst die Werte

$$\begin{aligned} & u^{2,x}(\mathcal{H}_2^\pm(\mathbf{x}_{1,i})) + h_x^2 \cdot R(\mathcal{H}_2^\pm(\mathbf{x}_{1,i})) = \\ & = u^{2,x}(\mathcal{H}_2^\pm(\mathbf{x}_{1,i})) + \\ & \quad h_x^2 \left[ h_x^{-4} \left( \int_{x_{l_1,i_1}}^{x_{l_1,i_1} \pm h_x} u^{4,x}(\xi, y_{l_2,i_2} \pm h_y) \cdot (x_{l_1,i_1} \pm h_x - \xi)^3 \frac{1}{3!} d\xi \right) \right] \end{aligned}$$

Dabei bedeuten wie im Beweis zu Satz 4.1  $h_x$  die (lokale) Maschenweite zum linken (bzw. rechten) Nachbarknoten und  $h_y$  bedeutet die (lokale) Maschenweite zu einem hierarchischen Vaterknoten an dem Punkt  $\mathbf{x}_{1,i}$ . An dem Punkt  $\mathbf{x}_{1,i}$  selbst erhält man

$$u^{2,x}(\mathbf{x}_{1,i}) + h_x^2 \cdot R(\mathbf{x}_{1,i})$$

mit

$$R(\mathbf{x}_{1,i}) = h_x^{-4} \left( \int_{x_{l_1,i_1}}^{x_{l_1,i_1} \pm h_x} u^{4,x}(\xi, y_{l_2,i_2}) \cdot (x_{l_1,i_1} \pm h_x - \xi)^3 \frac{1}{3!} d\xi \right)$$

als Ergebnis der Anwendung der Finiten Differenz. Die gewichtete Addition der berechneten Ausdrücke mit den Gewichten des Hierarchisierungs-Operators ergibt dann den Wert an dem Knoten  $\mathbf{x}_{1,i}$  nach Anwendung des ersten Teils  $\mathbf{D}_{11} \circ \mathbf{H}_1$  des Gesamtoperators. Wir entwickeln die Ableitungen in den Integralen an den Stellen  $\mathcal{H}_2^-(\mathbf{x}_{1,i})$  und  $\mathcal{H}_2^+(\mathbf{x}_{1,i})$  nach  $y$  und bekommen somit die Ausdrücke

$$\begin{aligned} & R(\mathcal{H}_2^\pm(\mathbf{x}_{1,i})) = \\ & = h_x^{-4} \left( \int_{x_{l_1,i_1}}^{x_{l_1,i_1} \pm h_x} u^{4,x}(\xi, y_{l_2,i_2} \pm h_y) \cdot (x_{l_1,i_1} \pm h_x - \xi)^3 \frac{1}{3!} d\xi \right) \\ & = h_x^{-4} \left[ \int_{x_{l_1,i_1}}^{x_{l_1,i_1} \pm h_x} \left( u^{4,x}(\xi, y_{l_2,i_2}) \pm h_y u^{4,x,1,y}(\xi, y_{l_2,i_2}) + \right. \right. \\ & \quad \left. \left. h_y^2 \left( h_y^{-2} \int_{y_{l_2,i_2}}^{y_{l_2,i_2} \pm h_y} u^{4,x,2,y}(\xi, \eta) \cdot (y_{l_2,i_2} \pm h_y - \eta)^1 \frac{1}{1!} d\eta \right) \right) \right. \\ & \quad \left. \cdot (x_{l_1,i_1} \pm h_x - \xi)^3 \frac{1}{3!} d\xi \right] \end{aligned} \tag{9.17}$$

als Fehlerterme. Insgesamt erhält man hierdurch

$$\begin{aligned}
\mathbf{D}_{11}^S &= \mathbf{E}_1 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 = \\
&= u^{2,x}(\mathbf{x}_{1,i}) + h_x^2 \left( h_x^{-4} \int_{x_{l_1,i_1}}^{x_{l_1,i_1} \pm h_x} u^{4,x}(\xi, y_{l_2,i_2}) \cdot (x_{l_1,i_1} \pm h_x - \xi)^3 \frac{1}{3!} d\xi \right) \\
&\quad - \frac{1}{2} \left( u^{2,x}(\mathcal{H}_2^\pm(\mathbf{x}_{1,i})) + \right. \\
&\quad \left. + h_x^2 \left( h_x^{-4} \int_{x_{l_1,i_1}}^{x_{l_1,i_1} \pm h_x} u^{4,x}(\xi, y_{l_2,i_2} \pm h_y) \cdot (x_{l_1,i_1} \pm h_x - \xi)^3 \frac{1}{3!} d\xi \right) \right) + \\
&\quad + \frac{1}{2} \left( u^{2,x}(\mathcal{H}_2^\pm(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) \\
&\stackrel{(9.17)}{=} u^{2,x}(\mathbf{x}_{1,i}) + \frac{1}{2} \left( \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) \\
&\quad - \frac{1}{2} h_x^2 h_y^2 \left( \frac{h_x^{-4} h_y^{-2}}{3! 1!} \int_{x_{l_1,i_1}}^{x_{l_1,i_1} \pm h_x} \int_{y_{l_2,i_2}}^{y_{l_2,i_2} \pm h_y} u^{4,x,2,y}(\xi, \eta) \cdot \right. \\
&\quad \left. (x_{l_1,i_1} \pm h_x - \xi)^3 \cdot (y_{l_2,i_2} \pm h_y - \eta) d\eta d\xi \right) \quad (9.18) \\
&\leq u^{2,x}(\mathbf{x}_{1,i}) + \frac{1}{2} \left( \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) + \frac{1}{2} \frac{4h_x^2 h_y^2}{2!4!} \|u\|_{C_{m^{ix}}^{2+2e_1}(\bar{\Omega})} \\
&\stackrel{(9.8)}{=} u^{2,x}(\mathbf{x}_{1,i}) + \frac{1}{2} \left( \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) + \frac{1}{2} \frac{h^2}{2!4!} \|u\|_{C_{m^{ix}}^{2+2e_1}(\bar{\Omega})} \\
&\quad (9.19)
\end{aligned}$$

als Ergebnis der zweiten Ableitung mittels Finiter Differenzen auf dünnen Gittern. Setzt man an den Randpunkten mit  $l_2 = 0$  die Fehlerterme  $\mathcal{F}(\mathbf{x}_{1,i})$  identisch 0, so sind die entsprechenden Terme an inneren Gitterpunkten von der Ordnung  $\mathcal{O}(h^2)$ . Der Term  $\mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i}))$  besteht an inneren Gitterpunkte ebenfalls nur aus Einzeltermen der Form  $c \cdot h^2 \cdot \|u\|_{C_{m^{ix}}^{2+2e_1}(\bar{\Omega})}$ , die mit dem Gewicht  $\frac{1}{2}$  jeweils einen Level tiefer transportiert werden. Betrachtet man sich die Konstante vor dem gesamten zusätzlichen Fehlerterm pro Knoten, so erkennt man wegen

$$\sum_{i=1}^L c \cdot \left(\frac{1}{2}\right)^i \leq \sum_{i=0}^{\infty} c \cdot \left(\frac{1}{2}\right)^i = 2 \cdot c, \quad (9.20)$$

daß sich aufgrund des wiederholten Faktors  $\frac{1}{2}$  kein logarithmischer Fehlerordnungsterm im zusätzlichen Fehler entwickelt. Damit ist die Aussage des Satzes bewiesen.  $\square$

#### A.4 Beweis zu Satz 4.4

**Beweis:** Aus dem Beweis zu Satz 4.2 ist bereits bekannt, daß an den Randpunkten  $\mathbf{x}_{1,i}$  mit  $l_2 = 0$  die Behauptung für Funktionen  $u \in C^4(\bar{\Omega})$  erfüllt ist, was eine schärfere Voraussetzung ist als die im Satz genannte: (A.3) führt zusammen mit (9.15) auf die Abschätzung (9.16) für Randpunkte, die wie im Beweis zu Satz 4.2 angegeben noch weiter verschärft werden kann.

In (9.18) tragen insgesamt vier Integrale zum Fehlerterm bei, wobei jedes einzelne Integral für den Beweis der Aussage von Satz 4.4 zweifach partiell integriert werden muß. Wir führen dies exemplarisch an einem der 4 Integrale

$$I := -\frac{1}{2}h_x^2h_y^2 \left( \frac{h_x^{-4}}{3!} \frac{h_y^{-2}}{1!} \int_{x_{l_1,i_1}}^{x_{l_1,i_1}+h_x} \int_{y_{l_2,i_2}}^{y_{l_2,i_2}+h_y} u^{4,x,2,y}(\xi, \eta) \cdot (x_{l_1,i_1} + h_x - \xi)^3 \cdot (y_{l_2,i_2} + h_y - \eta) d\eta d\xi \right)$$

aus (9.18) durch. Um die Notation einfacher zu halten, führen wir für die nachfolgende Rechnung die Schreibweise

$$f(a, \overline{c-d}) = f(a, c) - f(a, d) \quad (9.21)$$

sowie

$$f(\overline{a-b}, \overline{c-d}) = f(a, \overline{c-d}) - f(b, \overline{c-d})$$

ein. Damit erhalten wir

$$\begin{aligned} I &= -\frac{1}{2}h_x^1h_y^1 \left( \frac{h_x^{-3}}{3!} \frac{h_y^{-1}}{1!} \int_{x_{l_1,i_1}}^{x_{l_1,i_1}+h_x} \int_{x_{l_2,i_2}}^{x_{l_2,i_2}+h_y} u^{4,x,2,y}(\xi, \eta) \cdot (x_{l_1,i_1} + h_x - \xi)^3 \cdot (y_{l_2,i_2} + h_y - \eta) d\eta d\xi \right) \\ &= -\frac{1}{2}h_x^1h_y^1 \left( \frac{h_x^{-3}}{3!} h_y^{-1} \int_{x_{l_1,i_1}}^{x_{l_1,i_1}+h_x} \int_{y_{l_2,i_2}}^{y_{l_2,i_2}+h_y} u^{4,x,1,y}(\xi, \overline{\eta - y_{l_2,i_2}}) \cdot (x_{l_1,i_1} + h_x - \xi)^3 \cdot d\eta d\xi \right) \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2} h_x^1 h_y^1 \left( \frac{h_x^{-3}}{2!} h_y^{-1} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_y} u^{3, x, 1, y}(\xi - x_{l_1, i_1}, \eta - y_{l_2, i_2}) \cdot \right. \\
&\quad \left. \cdot (x_{l_1, i_1} + h_x - \xi)^2 \cdot d\eta d\xi \right) \\
&\stackrel{(9.8)}{\leq} -\frac{h}{4 \cdot \sqrt{2}} \cdot h^{1+\mu+1/2} \cdot \|u\|_{C^{4,1}(\bar{\Omega})}
\end{aligned}$$

Der Exponent  $\mu + 1/2$  mit  $\mu \geq 0$  kommt dabei dadurch zustande, daß wegen (9.8) in obiger Abschätzung mindestens  $h^{1/2}$  stehen muß, punktweise jedoch aufgrund der Lipschitz-Abschätzung auch ein höherer Exponent auftreten kann. Es gilt jedoch die Einschränkung  $0 \leq \mu \leq \frac{1}{2}$ , wobei die obere Grenze für  $\mu$  angenommen wird.

Obige Abschätzung erhält man für alle  $4 = 2^d$  (hier:  $d = 2$ ) Integrale. Somit kann man den Fehlerterm am Punkt  $\mathbf{x}_{1,i}$  angeben zu

$$\mathcal{F}(\mathbf{x}_{1,i}) + C \cdot h^{1+\sigma} \cdot \|u\|_{C_{\text{mix},L}^{1+2 \cdot e_1}(\bar{\Omega})} ,$$

wobei  $\sigma = \mu + \frac{1}{2} \geq \frac{1}{2}$  gilt und somit die Aussage des Satzes unter Verwendung der Diskussion der Konstanten des Terms  $h^2$  im Beweis zu Satz 4.2 bewiesen ist.  $\square$

## A.5 Beweis zu Satz 4.5

**Beweis:** Im Beweis zu Satz 4.2 kann man erkennen, daß durch die Enthierarchisierung ein Restglied mit dem führenden Term  $h_x^2 h_y^2$  entsteht, der die Konsistenzordnung 2 des Operators liefert. Wir wollen in diesem Beweis wiederum nur die Ableitung in  $x_1$ -Richtung betrachten. Wie zuvor bedeuten nachfolgend  $h_x$  die lokale Schrittweite zum nächsten Nachbarn in  $x_1$ -Richtung sowie  $h_{y_j} \equiv h_{x_j}$  für  $2 \leq j \leq d$  die Schrittweiten zu den hierarchischen Vaterknoten in die entsprechenden Raumdimensionen. Um Mißverständnissen vorzubeugen und um den Beweis in Analogie zur Notation im Beweis von Satz 4.2 zu führen, bezeichnet  $x$  (ohne Indizierung) in diesem Beweis die Richtung, in der die Finite Differenz angewendet werden soll, und  $y_j = x_j$  für  $2 \leq j \leq d$  durchläuft alle übrigen Richtungen.

Wir betrachten hier den Enthierarchisierungoperator  $\mathbf{E}_1$  als Tensorproduktoperator (vgl. Definition 2.1) und wenden die einzelnen Komponenten hintereinander an. Dabei sei an dieser Stelle nochmals darauf hingewiesen, daß die Reihenfolge der Einzeloperatoren bei der  $d$ -dimensionalen Hierarchisierung bzw. Enthierarchisierung unerheblich ist und die Einzeloperatoren un-

tereinander vertauschbar sind. Damit ergibt sich durch sukzessive Anwendung der Operatoren  $\mathbf{E}^j$  für  $2 \leq j \leq d$  schließlich

$$\begin{aligned}
\mathbf{E}^d \circ \dots \circ \mathbf{E}^2 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 &= \\
&= u^{2,x}(\mathbf{x}_{1,\mathbf{i}}) + \mathcal{F}_{1,\dots,d}(\mathbf{x}_{1,\mathbf{i}}) - \\
&\quad - h_x^2 h_{y_2}^2 \dots h_{y_d}^2 \left( \frac{h_x^{-4}}{3!} \frac{h_{y_2}^{-2}}{1!} \dots \frac{h_{y_j}^{-2}}{1!} \int_{x_{l_1,i_1} \pm h_x}^{x_{l_1,i_1} \pm h_x} \int_{y_{l_2,i_2} \pm h_{y_2}}^{y_{l_2,i_2} \pm h_{y_2}} \dots \int_{y_{l_d,i_d} \pm h_{y_d}}^{y_{l_d,i_d} \pm h_{y_d}} \right. \\
&\quad \left. \left( -\frac{1}{2} \right)^{d-1} u^{4,x,2,y_2,\dots,2,y_j}(\xi, \eta_2, \dots, \eta_j, y_{j+1}, \dots, y_d) \cdot (x_{l_1,i_1} \pm h_x - \xi)^3 \cdot \right. \\
&\quad \left. \cdot (y_{l_2,i_2} \pm h_{y_2} - \eta_2) \cdot \dots \cdot (y_{l_d,i_d} \pm h_{y_d} - \eta_d) d\eta_d \dots d\eta_2 d\xi \right) \quad (9.22)
\end{aligned}$$

Analog zu Formel (9.8) erhält man für reguläre dünne Gitter in  $d$  Raumdimensionen die Identität

$$h_x \cdot \prod_{j=2}^d h_{y_j} = \left( \frac{1}{2} \right)^{d-1} h. \quad (9.23)$$

Berechnung der insgesamt  $2^d$  Integrale in (9.22) unter Verwendung des Mittelwertsatzes der Integralrechnung ergibt nun unter Beachtung der Identität (9.23)

$$\begin{aligned}
\mathbf{E}_1 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 &= \\
&= u^{2,x}(\mathbf{x}_{1,\mathbf{i}}) + \mathcal{F}_{1,\dots,d}(\mathbf{x}_{1,\mathbf{i}}) + h^2 \frac{2^d \cdot 2^{1-d}}{4!(2!)^{d-1}} \left( \frac{1}{2} \right)^{2(d-1)} \|u\|_{C_{mix}^{2+2 \cdot e_1}(\bar{\Omega})} \\
&= u^{2,x}(\mathbf{x}_{1,\mathbf{i}}) + \mathcal{F}_{1,\dots,d}(\mathbf{x}_{1,\mathbf{i}}) + h^2 \frac{2^{2-d}}{4!2^{2(d-1)}} \|u\|_{C_{mix}^{2+2 \cdot e_1}(\bar{\Omega})},
\end{aligned}$$

was die Aussage des Satzes beweist. Dabei ist wie zuvor  $\mathcal{F}_{1,\dots,d}(\mathbf{x}_{1,\mathbf{i}})$  ein Fehlerterm, von dem bereits bekannt ist, daß er von der Ordnung  $\mathcal{O}(h^2)$  ist. Eine analoge Diskussion wie in (9.20) enthüllt, daß kein logarithmischer Faktor durch die Summation von Fehlertermen entsteht. Somit ist die Aussage des Satzes bewiesen.  $\square$

## A.6 Beweis zu Satz 4.7

**Beweis:** Der Beweis läuft prinzipiell wie der Beweis von Satz 4.4. Dabei ergibt sich die erste Behauptung wiederum sofort aus einer partiellen Integration in Ableitungsrichtung (vgl. [Hac86], Analogon zu vollen Gittern).

Für den Beweis der zweiten Aussage verwenden wir wie zuvor im Beweis zu Satz 4.4 die partielle Integration, um die Glattheitsvoraussetzung von

$C_{mix}^{2+2 \cdot \mathbf{e}_1}(\bar{\Omega})$ , wie sie in Satz 4.5 gefordert wird, auf  $C_{mix,L}^{1+2 \cdot \mathbf{e}_1}(\bar{\Omega})$  zu senken. Dabei muß im  $d$ -dimensionalen Fall die partielle Integration  $d$ -mal angewendet werden. Um die Notation übersichtlich zu halten, setzen wir nachfolgend in Analogie zu (9.21)

$$u(\xi, \overline{\eta_2 - y_2}, y_3, \dots, y_d) := u(\xi, \eta_2, y_3, \dots, y_d) - u(\xi, y_2, y_3, \dots, y_d) \quad (9.24)$$

Entsprechend wird  $u(\overline{\xi - x}, \overline{\eta_2 - y_2}, \dots, \overline{\eta_d - y_d})$  usw. definiert. Wie zuvor im Beweis von Satz 4.4 führen wir die partielle Integration an einem der auftretenden Integrale beispielhaft vor. Dazu führen wir zunächst nur die Integration bezüglich der Koordinatenrichtung  $\eta_d$  durch, wobei wir den Vorfaktor unberücksichtigt lassen:

$$\begin{aligned} & \frac{h_x^{-3}}{3!} \frac{h_{y_2}^{-1}}{1!} \dots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \dots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} u^{4, x, 2, y_2, \dots, 2, y_d}(\xi, \eta_2, \dots, \eta_d) \cdot \\ & (x_{l_1, i_1} + h_x - \xi)^3 \cdot (x_{l_2, i_2} + h_{y_2} - \eta_2) \dots (x_{l_d, i_d} + h_{y_d} - \eta_d) d\xi d\eta_2 \dots d\eta_d = \\ & = \frac{h_x^{-3}}{3!} \frac{h_{y_2}^{-1}}{1!} \dots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \dots \int_{y_{l_{d-1}, i_{d-1}}}^{y_{l_{d-1}, i_{d-1}} + h_{y_{d-1}}} \\ & \left[ u^{4, x, 2, y_2, \dots, 2, y_{d-1}, 1, y_d}(\xi, \eta_2, \dots, \eta_d) \cdot (x_{l_1, i_1} + h_x - \xi)^3 \cdot \right. \\ & \left. \cdot (x_{l_2, i_2} + h_{y_2} - \eta_2) \dots (x_{l_d, i_d} + h_{y_d} - \eta_d) \right]_{x_{l_d, i_d}}^{x_{l_d, i_d} + h_{y_d}} d\xi d\eta_2 \dots d\eta_{d-1} - \\ & - \frac{h_x^{-3}}{3!} \frac{h_{y_2}^{-1}}{1!} \dots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \dots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} \\ & \left( u^{4, x, 2, y_2, \dots, 2, y_{d-1}, 1, y_d}(\xi, \eta_2, \dots, \eta_d) \cdot (x_{l_1, i_1} + h_x - \xi)^3 \cdot \right. \\ & \left. \cdot (x_{l_2, i_2} + h_{y_2} - \eta_2) \dots (x_{l_{d-1}, i_{d-1}} + h_{y_{d-1}} - \eta_{d-1}) \cdot (-1) \right) d\xi d\eta_2 \dots d\eta_d \\ & = \frac{h_x^{-3}}{3!} \frac{h_{y_2}^{-1}}{1!} \dots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \dots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} \\ & \left( u^{4, x, 2, y_2, \dots, 2, y_{d-1}, 1, y_d}(\xi, \eta_2, \dots, \overline{\eta_d - y_{l_d, i_d}}) \cdot (x_{l_1, i_1} + h_x - \xi)^3 \cdot \right. \\ & \left. \cdot (x_{l_2, i_2} + h_{y_2} - \eta_2) \dots (x_{l_{d-1}, i_{d-1}} + h_{y_{d-1}} - \eta_{d-1}) \right) d\xi d\eta_2 \dots d\eta_d \quad (9.25) \end{aligned}$$

Hieraus erkennt man, daß man die partiellen Integrationen bezüglich der Raumdimensionen  $\eta_j$  für  $2 \leq j \leq d-1$  analog zu der soeben durchgeführten partiellen Integration bezüglich der Raumdimension  $\eta_d$  durchführen kann.



Man erhält nach diesen insgesamt  $d-1$  partiellen Integrationen die Identität

$$\begin{aligned}
& \frac{h_x^{-3}}{3!} \frac{h_{y_2}^{-1}}{1!} \cdots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \cdots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} u^{4, x, 2, y_2, \dots, 2, y_d}(\xi, \eta_2, \dots, \eta_d) \cdot \\
& (x_{l_1, i_1} + h_x - \xi)^3 \cdot (x_{l_2, i_2} + h_{y_2} - \eta_2) \cdots (x_{l_d, i_d} + h_{y_d} - \eta_d) d\xi d\eta_2 \cdots d\eta_d = \\
& = \frac{h_x^{-3}}{3!} \frac{h_{y_2}^{-1}}{1!} \cdots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \cdots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} \\
& \left( u^{4, x, 1, y_2, \dots, 1, y_{d-1}, 1, y_d}(\xi, \overline{\eta_2 - y_{l_2, i_2}}, \dots, \overline{\eta_d - y_{l_d, i_d}}) \cdot \right. \\
& \left. \cdot (x_{l_1, i_1} + h_x - \xi)^3 \right) \cdot d\xi d\eta_2 \cdots d\eta_d \quad (9.26)
\end{aligned}$$

Wendet man analog wie im Beweis von Satz 4.4 noch partielle Integration in die verbleibende Ableitungsrichtung an, so erhält man aus (9.26) die Formel

$$\begin{aligned}
& \frac{h_x^{-3}}{3!} \frac{h_{y_2}^{-1}}{1!} \cdots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \cdots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} u^{4, x, 2, y_2, \dots, 2, y_d}(\xi, \eta_2, \dots, \eta_d) \cdot \\
& (x_{l_1, i_1} + h_x - \xi)^3 \cdot (x_{l_2, i_2} + h_{y_2} - \eta_2) \cdots (x_{l_d, i_d} + h_{y_d} - \eta_d) d\xi d\eta_2 \cdots d\eta_d = \\
& = -\frac{h_x^{-3}}{2!} \frac{h_{y_2}^{-1}}{1!} \cdots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \cdots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} \\
& \left( u^{3, x, 1, y_2, \dots, 1, y_{d-1}, 1, y_d}(x_{l_1, i_1}, \overline{\eta_2 - y_{l_2, i_2}}, \dots, \overline{\eta_d - y_{l_d, i_d}}) \cdot \right. \\
& \left. \cdot (x_{l_1, i_1} + h_x - \xi)^2 \right) \cdot d\xi d\eta_2 \cdots d\eta_d + \\
& + \frac{h_x^{-3}}{2!} \frac{h_{y_2}^{-1}}{1!} \cdots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \cdots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} \\
& \left( u^{3, x, 1, y_2, \dots, 1, y_{d-1}, 1, y_d}(\xi, \overline{\eta_2 - y_{l_2, i_2}}, \dots, \overline{\eta_d - y_{l_d, i_d}}) \cdot \right. \\
& \left. \cdot (x_{l_1, i_1} + h_x - \xi)^2 \right) \cdot d\xi d\eta_2 \cdots d\eta_d \\
& = \frac{h_x^{-3}}{2!} \frac{h_{y_2}^{-1}}{1!} \cdots \frac{h_{y_d}^{-1}}{1!} \int_{x_{l_1, i_1}}^{x_{l_1, i_1} + h_x} \int_{y_{l_2, i_2}}^{y_{l_2, i_2} + h_{y_2}} \cdots \int_{y_{l_d, i_d}}^{y_{l_d, i_d} + h_{y_d}} \\
& \left( u^{3, x, 1, y_2, \dots, 1, y_{d-1}, 1, y_d}(\xi - x_{l_1, i_1}, \overline{\eta_2 - y_{l_2, i_2}}, \dots, \right. \\
& \left. \overline{\eta_d - y_{l_d, i_d}}) \cdot (x_{l_1, i_1} + h_x - \xi)^2 \right) \cdot d\xi d\eta_2 \cdots d\eta_d
\end{aligned}$$

Die Anzahl an Lipschitzabschätzungen, die bei einem der insgesamt  $2^d$   $d$ -dimensionalen Integrale durchzuführen ist, beträgt dabei  $2^{d-1}$ . Dabei kann man die Terme nach den Regeln der Kombinatorik grundsätzlich so

abschätzen, daß wie im zweidimensionalen Fall die kleinste (lokale) Maschenweite (bezüglich aller Raumdimensionen) als Faktor vor der Norm  $\|u\|_{C_{mix,L}^{1+2\cdot e_1}(\bar{\Omega})}$  zu stehen kommt. Wie zuvor im zweidimensionalen Fall eliminiert die durch (9.23) auftretende negative Potenz von 2 dabei den Faktor  $2^{d-1}$ , der durch die Anzahl von Summanden der Lipschitzabschätzungen für ein Integral entsteht. Zudem gilt für die kleinste (lokale) Maschenweite wegen (9.23) stets

$$\min_{\bar{h} \in \{h_x, h_{y_j}; 2 \leq j \leq d\}} \bar{h} \leq 2^{-d} \cdot h^{\frac{1}{d}},$$

womit sich in Analogie zu Satz 4.4  $\sigma > \frac{1}{d}$  ergibt. Insgesamt erhält man somit den Fehlerterm

$$\mathcal{F}(\mathbf{x}_{1,i}) + C \cdot h^{1+\sigma} \|u\|_{C_{mix,L}^{1+2\cdot e_1}(\bar{\Omega})},$$

wobei wie zuvor für den Fehlerterm  $\mathcal{F}(\mathbf{x}_{1,i})$ , der Fehlerterme von Vaterknoten beinhaltet, klar ist, daß er von der Ordnung  $h^{1+\sigma}$  bzw.  $h^2$  ist. Die identische Diskussion um den mit jedem Level neu zu multiplizierenden Faktor  $\frac{1}{2}$  vor den einzelnen Fehlertermen führt auch hier zur Erkenntnis, daß sich kein logarithmischer Faktor im Fehlerterm einstellt. Somit ist die Aussage des Satzes bewiesen.  $\square$

## A.7 Beweis zu Satz 4.8

**Beweis:** Der Beweis für die drei ersten Behauptungen des Satzes läuft analog zum Beweis des Satzes 4.1. Mit denselben Argumenten wie dort reicht es hierbei aus, lediglich die Ableitungen  $\mathbf{D}_1^{S,\pm}$  und  $\mathbf{D}_1^{S,0}$  in Koordinatenrichtung  $i = 1$  zu betrachten. Wir beschränken uns zudem wiederum auf den Fall  $d = 2$  mit  $\bar{\Omega} = [0, 1]^2$ . Die Erweiterung auf  $d$  Dimensionen läuft dabei wie im Beweis zu Satz 4.1 (Konsistenz der zweiten Ableitung  $\mathbf{D}_i^S$ ).

Nach Anwendung der Hierarchisierung erhalten wir dieselben Ergebnisse wie in (9.3), (9.4) und (9.5). Nach der Anwendung des Sterns für die erste Ableitung bekomme wir folgende Taylorreihen an inneren Gitterpunkten:

$$\begin{aligned} \mathbf{D}_1^{S,+} : v(\bar{\mathbf{x}}_{1,i}) &= \frac{-1}{2h_x} \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h_x^m}{m!} \frac{2h_y^{2j}}{(2j)!} u^{m,x,2j,y}(\bar{\mathbf{x}}_{1,i}) \\ &= - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h_x^{m-1}}{m!} \frac{h_y^{2j}}{(2j)!} u^{m,x,2j,y}(\bar{\mathbf{x}}_{1,i}) \end{aligned} \quad (9.27)$$

$$\begin{aligned}
\mathbf{D}_1^{S,-} : v(\vec{\mathbf{x}}_{1,i}) &= \frac{-1}{2h_x} \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-h_x)^m}{m!} \frac{2h_y^{2j}}{(2j)!} u^{m,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \\
&= - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-h_x)^{m-1}}{m!} \frac{h_y^{2j}}{(2j)!} u^{m,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \quad (9.28)
\end{aligned}$$

$$\begin{aligned}
\mathbf{D}_1^{S,0} : v(\vec{\mathbf{x}}_{1,i}) &= \frac{-1}{2 \cdot (2h_x)} \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{2h_x^{2m-1}}{(2m-1)!} \frac{2h_y^{2j}}{(2j)!} u^{2m-1,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \\
&= - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h_x^{2(m-1)}}{(2m-1)!} \frac{h_y^{2j}}{(2j)!} u^{2m-1,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \quad (9.29)
\end{aligned}$$

An Randpunkten mit  $l_2 = 0$ , an denen die Operatoren  $\mathbf{H}_1$  und  $\mathbf{E}_1$  die Identität darstellen, erhalten wir die von vollen Gittern bekannten Formeln

$$\begin{aligned}
\mathbf{D}_1^{S,+} : v(\vec{\mathbf{x}}_{1,i}) &= \sum_{m=1}^{\infty} \frac{h^{m-1}}{m!} u^{m,x}(\vec{\mathbf{x}}_{1,i}) \\
&= u^{1,x}(\vec{\mathbf{x}}_{1,i}) + \sum_{m=2}^{\infty} \frac{h^{m-1}}{m!} u^{m,x}(\vec{\mathbf{x}}_{1,i}) \\
\mathbf{D}_1^{S,-} : v(\vec{\mathbf{x}}_{1,i}) &= \sum_{m=1}^{\infty} \frac{(-h)^{m-1}}{m!} u^{m,x}(\vec{\mathbf{x}}_{1,i}) \\
&= u^{1,x}(\vec{\mathbf{x}}_{1,i}) + \sum_{m=2}^{\infty} \frac{(-h)^{m-1}}{m!} u^{m,x}(\vec{\mathbf{x}}_{1,i}) \\
\mathbf{D}_1^{S,0} : v(\vec{\mathbf{x}}_{1,i}) &= \sum_{m=1}^{\infty} \frac{h^{2(m-1)}}{(2m-1)!} u^{2m-1,x}(\vec{\mathbf{x}}_{1,i}) \\
&= u^{1,x}(\vec{\mathbf{x}}_{1,i}) + \sum_{m=2}^{\infty} \frac{h^{2(m-1)}}{(2m-1)!} u^{2m-1,x}(\vec{\mathbf{x}}_{1,i})
\end{aligned}$$

Kombiniert man diese Ergebnisse nun mit dem Operator  $\mathbf{E}_1$ , so erhält man den gewünschten Finite Differenzen Operator für erste Ableitungen auf dünnen Gittern. Man rufe sich hierfür nochmals die Identität  $h_x = 2^{l_2-1}h$  ins Gedächtnis. Wir starten mit der Rückwärts-Differenz:

$$\begin{aligned}
\mathbf{D}_1^{S,-} \circ u(\mathbf{x}_{1,i}) &\equiv \mathbf{E}_1 \circ \mathbf{D}_1^- \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) = \\
&= \mathbf{D}_1^- \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) + \frac{1}{2} \left( \mathbf{D}_1^- \circ \mathbf{H}_1 \circ \left( u(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + u(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right) \right) \\
&= - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-h_x)^{m-1}}{m!} \frac{(h_y)^{2j}}{(2j)!} u^{m,x,2j,y}(\mathbf{x}_{1,i}) +
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \left( u^{1,x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \sum_{m=2}^{\infty} \frac{(-h)^{m-1}}{m!} u^{m,x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) \\
& + \frac{1}{2} \left( u^{1,x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \sum_{m=2}^{\infty} \frac{(-h)^{m-1}}{m!} u^{m,x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right) \\
= & - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-h_x)^{m-1}}{m!} \frac{(h_y)^{2j}}{(2j)!} u^{m,x,2j,y}(\mathbf{x}_{1,i}) + \\
& + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{(-h_y)^j}{j!} u^{1,x,j,y}(\mathbf{x}_{1,i}) + \mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \right. \\
& \left. + \sum_{m=2}^{\infty} \sum_{j=0}^{\infty} \frac{(-h)^{m-1}}{m!} \frac{(-h_y)^j}{j!} u^{m,x,j,y}(\mathbf{x}_{1,i}) \right) + \\
& + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{h_y^j}{j!} u^{1,x,j,y}(\mathbf{x}_{1,i}) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \right. \\
& \left. + \sum_{m=2}^{\infty} \sum_{j=0}^{\infty} \frac{(-h)^{m-1}}{m!} \frac{h_y^j}{j!} u^{m,x,j,y}(\mathbf{x}_{1,i}) \right) + \\
= & - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-h_x)^{m-1}}{m!} \frac{(h_y)^{2j}}{(2j)!} u^{m,x,2j,y}(\mathbf{x}_{1,i}) + \\
& \frac{1}{2} (\mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i}))) \\
& + \sum_{j=0}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{1,x,2j,y}(\mathbf{x}_{1,i}) + \sum_{m=2}^{\infty} \sum_{j=0}^{\infty} \frac{(-h)^{m-1}}{m!} \frac{h_y^{2j}}{(2j)!} u^{m,x,2j,y}(\mathbf{x}_{1,i}) \\
= & - \left( \sum_{j=1}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{1,x,2j,y}(\mathbf{x}_{1,i}) + \right. \\
& \left. + \sum_{m=2}^{\infty} \sum_{j=1}^{\infty} \frac{(-2^{l_2-1}h)^{m-1}}{m!} \frac{(h_y)^{2j}}{(2j)!} u^{m,x,2j,y}(\mathbf{x}_{1,i}) \right) + \\
& + u^{1,x}(\mathbf{x}_{1,i}) + \sum_{j=1}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{1,x,2j,y}(\mathbf{x}_{1,i}) + \sum_{m=2}^{\infty} \frac{(-h)^{m-1}}{m!} u^{m,x}(\mathbf{x}_{1,i}) + \\
& + \sum_{m=2}^{\infty} \sum_{j=1}^{\infty} \frac{(-h)^{m-1}}{m!} \frac{(h_y)^{2j}}{(2j)!} u^{m,x,2j,y}(\mathbf{x}_{1,i}) + \\
& + \frac{1}{2} (\mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i})))
\end{aligned}$$

$$\begin{aligned}
&= u^{1,x}(\mathbf{x}_{1,i}) + \sum_{m=2}^{\infty} \frac{(-h)^{m-1}}{m!} u^{m,x}(\mathbf{x}_{1,i}) + \\
&\quad + \frac{1}{2} (\mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i}))) + \\
&\quad + \sum_{m=2}^{\infty} \sum_{j=1}^{\infty} \frac{(-h)^{m-1}}{m!} \frac{(h_y)^{2j}}{(2j)!} \left(1 + (-2^{l_2-1})^{m-1}\right) u^{m,x,2j,y}(\mathbf{x}_{1,i}) \\
&= u^{1,x}(\mathbf{x}_{1,i}) + \sum_{m=2}^{\infty} \frac{(-h)^{m-1}}{m!} u^{m,x}(\mathbf{x}_{1,i}) + \mathcal{F}^-(\mathbf{x}_{1,i}) \tag{9.30}
\end{aligned}$$

Der Beweis für die Vorwärts-Differenz ist hierbei identisch zu dem für die Rückwärts-Differenz, wenn man  $-h_x$  durch  $h_x$  und  $-h$  durch  $h$  substituiert. Zudem erhält man anstelle von  $\mathcal{F}^-(\mathbf{x}_{1,i})$  notationstechnisch den Term  $\mathcal{F}^+(\mathbf{x}_{1,i})$ , der für jeweils beide Differenzen den Fehler beschreibt, den man additiv zu dem Fehler enthält, der im Fall voller Gitter auftritt.

Es bleibt noch zu zeigen, daß  $\mathcal{F}^-(\mathbf{x}_{1,i})$  von erster Ordnung  $\mathcal{O}(h)$  ist, da die Summe in (9.30) von der Ordnung  $\mathcal{O}(h)$  ist. Analog erhält man dann die Aussage für  $\mathcal{F}^+(\mathbf{x}_{1,i})$ . Die ausführliche Formel für  $\mathcal{F}^-(\mathbf{x}_{1,i})$  lautet

$$\begin{aligned}
\mathcal{F}^-(\mathbf{x}_{1,i}) &= \frac{1}{2} (\mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i}))) + \\
&\quad + \sum_{m=2}^{\infty} \sum_{j=1}^{\infty} \frac{(-h)^{m-1}}{m!} \frac{(h_y)^{2j}}{(2j)!} \left(1 + (-2^{l_2-1})^{m-1}\right) u^{m,x,2j,y}(\mathbf{x}_{1,i}) \\
&= \frac{1}{2} (\mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i}))) + \\
&\quad + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(-h)^m}{(m+1)!} \frac{(h_y)^{2j}}{(2j)!} \left(1 + (-2^{l_2-1})^m\right) u^{m+1,x,2j,y}(\mathbf{x}_{1,i}) \tag{9.31}
\end{aligned}$$

Da wie im Beweis der zweiten Ableitung mittels Finiter Differenzen auf dünnen Gittern auch hier bereits bewiesen ist, daß in (9.31) der Fehlerterm  $\frac{1}{2} (\mathcal{F}^-(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}^-(\mathcal{H}_2^+(\mathbf{x}_{1,i})))$  von erster Ordnung  $\mathcal{O}(h)$  ist, bleibt nur noch zu zeigen, daß die Doppelsumme im Ausdruck in (9.31) ebenfalls von erster Ordnung ist. An dieser Stelle sei an den Beweis von Satz 4.1 erinnert, wo wir  $h_x$  und  $h_y$  genauer bestimmt haben. Wie in dem angesprochenen Beweis führen wir auch hier eine Fallunterscheidung mit insgesamt 4 Fällen durch, um den Fehlerterm genauer zu analysieren.

1. Fall: analog zum Vorgehen im Beweis zu Satz 4.1 betrachten wir hier Terme der Form

$$\frac{(-h)^m}{(m+1)!} \frac{(h_y)^{2j}}{(2j)!} u^{m+1,x,2j,y}(\mathbf{x}_{1,i}) \quad , \tag{9.32}$$

die gerade den Anteil des Fehlerterms  $\mathcal{F}^-(\mathbf{x}_{1,i})$  ausmachen, bei dem in der Klammer der Faktor 1 genommen wird. Bei diesen Termen ist die  $h$ -Potenz direkt ablesbar.

2. Fall 2: hier betrachten wir wie im Beweis zuvor die übrigen Terme

$$\frac{(-h)^m}{(m+1)!} \frac{(h_y)^{2j}}{(2j)!} \left(-2^{l_2-1}\right)^m u^{m+1,x,2j,y}(\mathbf{x}_{1,i}) \quad , \quad (9.33)$$

bei denen sich im Gegensatz zu den Termen in (9.32) die auftretenden Potenzen von  $h$  nicht direkt ablesen lassen, womit auch hier eine genauere Analyse der Terme erforderlich wird. Dazu unterscheiden wir die 3 Fälle  $m = 2j$ ,  $2j < m$  und  $2j > m$ :

- Fall 2a: sei  $m = 2j$ . Dann gilt

$$\begin{aligned} & \frac{(-h)^m}{(m+1)!} \frac{h_y^{2j}}{(2j)!} \left(-2^{l_2-1}\right)^m = \\ & = (-1)^m \cdot \frac{1}{(m+1)!(2j)!} \cdot \frac{2^{-lm+2l_1j-2l_2j+ml_2-m}}{(m+1)!(2j)!} \\ & = (-1)^m \cdot \frac{1}{(m+1)!(2j)!} \cdot \left(2^{l_1+l_2-l-1}\right)^m \cdot \left(2^{-l}\right)^m \cdot 2^{-m} \\ & = (-1)^m \cdot \frac{1}{(m+1)!(2j)!} \cdot \underbrace{\left(2^{l_1+l_2-l-1}\right)^m}_{0 \leq \dots \leq 1} \cdot 2^{-m} \cdot h^m \\ & \leq (-1)^m \frac{1}{(m+1)!(2j)!} \cdot 2^{-m} \cdot h^m \\ & = \frac{2^{-2j}}{(m+1)!(2j)!} \cdot h^{2j} \end{aligned}$$

Damit erhalten wir in diesem Fall nur gerade Potenzen von  $h$ .

- Fall 2b: sei  $2j < m$ . Dann gilt

$$\begin{aligned} & \frac{(-h)^m}{(m+1)!} \frac{h_y^{2j}}{(2j)!} \left(-2^{l_2-1}\right)^m = \\ & = (-1)^m \cdot \frac{2^{-m} \cdot (2^{-l})^{2j}}{(m+1)!(2j)!} \cdot \left(2^{l_1+l_2-l-1}\right)^{2j} \cdot \left(2^{l_2-l}\right)^{m-2j} \\ & = (-1)^m \cdot \frac{2^{-m} \cdot h^{2j}}{(m+1)!(2j)!} \cdot \underbrace{\left(2^{l_1+l_2-l-1}\right)^{2j}}_{0 \leq \dots \leq 1} \cdot \underbrace{\left(2^{l_2-l}\right)^{m-2j}}_{0 \leq \dots \leq 1} \\ & \leq (-1)^m \cdot \frac{2^{-2m}}{(m+1)!(2j)!} \cdot h^{2j} \end{aligned}$$

Damit erhalten wir auch in diesem Fall ausschließlich gerade Potenzen von  $h$ .

- Fall 2c: sei  $2j > m$ . Dann gilt

$$\begin{aligned}
& \frac{(-h)^m}{(m+1)!} \frac{h_y^{2j}}{(2j)!} (-2^{l_2-1})^m = \\
& = (-1)^m \cdot \frac{2^{-2j} \cdot (2^{-l})^m}{(m+1)!(2j)!} \cdot \left(2^{l_1+l_2-l-1}\right)^m \cdot \left(2^{l_1-l}\right)^{2j-m} \\
& = (-1)^m \cdot \frac{2^{-2j} \cdot h^m}{(m+1)!(2j)!} \cdot \underbrace{\left(2^{l_1+l_2-l-1}\right)^m}_{0 \leq \dots \leq 1} \cdot \underbrace{\left(2^{l_1-l}\right)^{2j-m}}_{0 \leq \dots \leq 1} \\
& \leq (-1)^m \cdot \frac{2^{-2j}}{(m+1)!(2j)!} \cdot h^m
\end{aligned}$$

Damit stehen in allen vier Fällen Konstanten vor  $h$ , also Größen, die unabhängig von  $h$  sind. Wie zuvor wollen wir nun die Konstanten vor beliebigen Potenzen von  $h$  bestimmen. Dazu unterscheiden wir nach geraden und ungeraden Potenzen und erhalten somit folgende Beiträge:

- ungerade Potenzen  $m = 2k + 1$ :

- Fall 1: Sei  $m = 2k + 1$ . Dann erhalten wir

$$\frac{-h^{2k+1}}{(2(k+1))!} (\cosh(h_y) - 1)$$

- Fall 2c: Sei  $m = 2k + 1$  und  $2j > 2k + 1$ , also  $j \geq k + 1$ . Dann gilt

$$\begin{aligned}
& \frac{h^{2k+1}}{(2(k+1))!} \sum_{j=k+1}^{\infty} \frac{2^{-2j}}{(2j)!} = \\
& \frac{h^{2k+1}}{(2(k+1))!} \left( \cosh\left(\frac{1}{2}\right) - 1 - \frac{\left(\frac{1}{2}\right)^2}{2!} - \dots - \frac{\left(\frac{1}{2}\right)^{2k}}{(2k)!} \right)
\end{aligned}$$

- gerade Potenzen  $m = 2k$ :

- Fall 1: Sei  $m = 2k$ . Dann erhalten wir

$$\frac{h^{2k}}{(2k+1)!} (\cosh(h_y) - 1)$$

- Fall 2a: Sei  $m = 2k = 2j$ , also  $k = j$ . Dann erhalten wir

$$2^{-2k} \cdot \frac{h^{2k}}{(2k)!(2k+1)!}$$

– Fall 2b: Sei  $j = k$  und  $m > 2k$ . Dann gilt

$$\begin{aligned} \frac{h^{2k}}{(2k)!} \sum_{m=2k+1}^{\infty} \frac{2^{-m}}{(m+1)!} &= \\ \frac{2h^{2k}}{(2k)!} \left( \sqrt{e} - 1 - \frac{1}{2} - \dots - \frac{\left(\frac{1}{2}\right)^{2k+1}}{(2k+1)!} \right) \end{aligned}$$

– Fall 2c: Sei  $m = 2k$  und  $j > k$ . Dann gilt

$$\begin{aligned} \frac{h^{2k}}{(2k+1)!} \sum_{j=k+1}^{\infty} \frac{2^{-2j}}{(2j)!} &= \\ \frac{h^{2k}}{(2k+1)!} \left( \cosh\left(\frac{1}{2}\right) - 1 - \frac{\left(\frac{1}{2}\right)^2}{2!} - \dots - \frac{\left(\frac{1}{2}\right)^{2k}}{(2k)!} \right) \end{aligned}$$

Damit sind im “worst case” die Konstanten aller  $h$ -Potenzen bestimmte, da jeder der einzelnen obigen Ausdrücke unabhängig von  $h$  ist. Somit haben wir die Konsistenzordnung  $\mathcal{O}(h)$  sowohl für die Vorwärts- als auch für die Rückwärts-Differenz nachgewiesen und es gilt folglich die Abschätzung

$$\|\mathbf{D}_j^{S,\pm} \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial u}{\partial x_j}\|_{\infty} \leq C \cdot h^1 \|u\|_{C^{\infty}(\bar{\Omega})}$$

Kommen wir nun zum Beweis für die zentrale Differenz  $\mathbf{D}_1^{S,0}$ :

$$\begin{aligned} \mathbf{D}_1^{S,0} \circ u(\mathbf{x}_{1,i}) &\equiv \mathbf{E}_1 \circ \mathbf{D}_1^0 \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) = & (9.34) \\ &= \mathbf{D}_1^0 \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) + \frac{1}{2} \left( \mathbf{D}_1^0 \circ \mathbf{H}_1 \circ \left( u(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + u(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right) \right) \\ &= - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h_x^{2(m-1)}}{(2m-1)!} \frac{h_y^{2j}}{(2j)!} u^{2m-1,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \\ &\quad + \frac{1}{2} \left( u^{1,x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \sum_{m=1}^{\infty} \frac{h^{2m}}{(2m+1)!} u^{2m+1,x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right. \\ &\quad \left. + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) + \\ &\quad + \frac{1}{2} \left( u^{1,x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \sum_{m=1}^{\infty} \frac{h^{2m}}{(2m+1)!} u^{2m+1,x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right. \\ &\quad \left. + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right) \end{aligned}$$



$$\begin{aligned}
&= - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h_x^{2(m-1)}}{(2m-1)!} \frac{h_y^{2j}}{(2j)!} u^{2m-1,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \\
&\quad + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{(-h_y)^j}{j!} u^{1,x,j,y}(\vec{\mathbf{x}}_{1,i}) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \right. \\
&\quad \quad \left. + \sum_{m=1}^{\infty} \sum_{j=0}^{\infty} \frac{h^{2m}}{(2m+1)!} \frac{(-h_y)^j}{j!} u^{2m+1,x,j,y}(\mathbf{x}_{1,i}) + \right) \\
&\quad + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{h_y^j}{j!} u^{1,x,j,y}(\vec{\mathbf{x}}_{1,i}) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \right. \\
&\quad \quad \left. + \sum_{m=1}^{\infty} \sum_{j=0}^{\infty} \frac{h^{2m}}{(2m+1)!} \frac{h_y^j}{j!} u^{2m+1,x,j,y}(\mathbf{x}_{1,i}) + \right) \\
&= - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{(2^{l_2-1}h)^{2m}}{(2m+1)!} \frac{h_y^{2j}}{(2j)!} u^{2m+1,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \\
&\quad + u^{1,x}(\vec{\mathbf{x}}_{1,i}) + \sum_{j=1}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{1,x,2j,y}(\vec{\mathbf{x}}_{1,i}) - \sum_{j=1}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{1,x,2j,y}(\vec{\mathbf{x}}_{1,i}) \\
&\quad + \sum_{m=1}^{\infty} \frac{h^{2m}}{(2m+1)!} u^{2m+1,x}(\mathbf{x}_{1,i}) + \frac{1}{2} \left( \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right) \\
&\quad + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h^{2m}}{(2m+1)!} \frac{h_y^{2j}}{(2j)!} u^{2m+1,x,2j,y}(\mathbf{x}_{1,i}) \\
&= u^{1,x}(\vec{\mathbf{x}}_{1,i}) + \sum_{m=1}^{\infty} \frac{h^{2m}}{(2m+1)!} u^{2m+1,x}(\mathbf{x}_{1,i}) \\
&\quad + \frac{1}{2} \left( \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right) + \\
&\quad + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{h^{2m}}{(2m+1)!} \frac{h_y^{2j}}{(2j)!} \left( 1 - (2^{l_2-1})^{2m} \right) u^{2m+1,x,2j,y}(\mathbf{x}_{1,i}) \\
&= u^{1,x}(\vec{\mathbf{x}}_{1,i}) + \sum_{m=1}^{\infty} \frac{h^{2m}}{(2m+1)!} u^{2m+1,x}(\mathbf{x}_{1,i}) + \mathcal{F}(\mathbf{x}_{1,i}) \tag{9.35}
\end{aligned}$$

Der ‘‘zusätzliche Fehlerterm’’  $\mathcal{F}(\mathbf{x}_{1,i})$  in (9.35) ist dabei annahernd derselbe wie schon im Fall der zweiten Ableitung. Unterschiede bestehen lediglich in einem konstanten Faktor 2 und unterschiedlichen Ableitungen, mit denen die Potenzen der Maschenweiten zu multiplizieren sind. Entscheidend fur die Ordnung dieses Terms ist jedoch ausschlielich der Beitrag der Summanden der Doppelsumme, der sich mit den Potenzen der Schrittweiten beschaftigt. Da diese mit denen im Fall der zweiten Ableitung auf dunnen Gittern iden-

tisch sind, ist die Konsistenzordnung zwei auch für die zentrale Differenz auf dünnen Gittern bewiesen und es gilt die Abschätzung

$$\|\mathbf{D}_j^{S,0} \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial u}{\partial x_j}\|_\infty \leq C \cdot h^2 \|u\|_{C^\infty(\bar{\Omega})} .$$

Dabei werden wie zuvor sämtliche auftretenden Ableitungen über den Ausdruck  $\|u\|_{C^\infty(\bar{\Omega})}$  abgeschätzt.

*Beweis von (4.30):*

Die eindimensionalen Gitterlinien, bei denen  $d - 1$  Koordinaten fest sind und lediglich eine Koordinate als variabel angesehen wird, stellen eindimensionale, volle Gitter dar. Eingeschränkt auf solche Gitterlinien gilt  $\mathbf{D}_{ii} = \mathbf{D}_i^+ \circ \mathbf{D}_i^- = \mathbf{D}_i^- \circ \mathbf{D}_i^+$ , falls  $i$  diejenige Koordinatenrichtung angibt, die als variabel angesehen wird (vgl. hierzu z.B. [Hac86, GrRo92] u.v.m.). Der Operator  $\mathbf{D}_{ii}^S$  arbeitet lokal auf eindimensionalen, vollen Gittern. Folglich erhalten wir

$$\begin{aligned} \mathbf{D}_{ii}^S &= \mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i \\ &= \mathbf{E}_i \circ \mathbf{D}_i^+ \circ \mathbf{D}_i^- \circ \mathbf{H}_i \\ &= \mathbf{E}_i \circ \mathbf{D}_i^+ \circ \mathbf{H}_i \circ \mathbf{E}_i \circ \mathbf{D}_i^- \circ \mathbf{H}_i \\ &= \mathbf{D}_i^{S,+} \circ \mathbf{D}_i^{S,-} \end{aligned}$$

Zudem haben wir

$$\begin{aligned} \mathbf{D}_{ii}^S &= \mathbf{E}_i \circ \mathbf{D}_{ii} \circ \mathbf{H}_i \\ &= \mathbf{E}_i \circ \mathbf{D}_i^- \circ \mathbf{D}_i^+ \circ \mathbf{H}_i \\ &= \mathbf{E}_i \circ \mathbf{D}_i^- \circ \mathbf{H}_i \circ \mathbf{E}_i \circ \mathbf{D}_i^+ \circ \mathbf{H}_i \\ &= \mathbf{D}_i^{S,-} \circ \mathbf{D}_i^{S,+} , \end{aligned}$$

was die Behauptung beweist.

*Beweis von (4.32):*

Addiert man die beiden Ergebnisse für Vorwärts- und Rückwärts-Differenz auf dünnen Gittern, so erhält man nach kurzer Rechnung die entsprechende Aussage. Kürzer und mit weniger Rechenaufwand verbunden ist der Nachweis, wenn man sich mittels

$$\begin{aligned} \mathbf{D}_i^{S,0} &= \mathbf{E}_i \circ \mathbf{D}_i^0 \circ \mathbf{H}_i \\ &= \mathbf{E}_i \circ \left( \frac{1}{2} \cdot \mathbf{D}_i^- + \frac{1}{2} \cdot \mathbf{D}_i^+ \right) \circ \mathbf{H}_i \\ &= \frac{1}{2} \cdot \mathbf{E}_i \circ \mathbf{D}_i^+ \circ \mathbf{H}_i + \frac{1}{2} \cdot \mathbf{E}_i \circ \mathbf{D}_i^- \circ \mathbf{H}_i \\ &= \frac{1}{2} \cdot \mathbf{D}_i^{S,+} + \frac{1}{2} \cdot \mathbf{D}_i^{S,-} \end{aligned}$$

behilflich ist und sich damit wie zuvor auf lokal eindimensionale, volle Gitter zurückgezogen hat. Damit ist die zentrale Differenz auf dünnen Gittern genauso wie auf vollen Gittern die gewichtete Addition von Vorwärts- und Rückwärts-Differenz.  $\square$

## A.8 Beweis zu Satz 4.10

**Beweis:** In einer Raumdimension ist nichts zu zeigen. Somit beschränken wir uns zunächst auf den Beweis der Behauptung in zwei Raumdimensionen und setzen  $h_x = 2^{-l_1}$  und  $h_y = 2^{(l_1-l_2)-1}$ . Wie bei den Beweisen zuvor reicht es aus, lediglich die Aussage für den Fall  $j = 1$  zu untersuchen. Wie wir schon im Beweis von Satz 4.1 gesehen haben, erhalten wir für Gitterpunkte am Rand mit  $l_2 = 0$  die Aussage aus der Tatsache, daß für diese Punkte die Operatoren  $\mathbf{H}_1$  und  $\mathbf{E}_1$  identisch mit der Identität sind. Aus notations-technischen Gründen formulieren wir dieses Ergebnis nachfolgend mit den Bezeichnungen aus Definition 4.2:

$$\begin{aligned} \mathbf{E}_1 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) &= \\ &= u^{2,x}(\mathbf{x}_{1,i}) + \sum_{m=1}^{\infty} \frac{2h_x^{2m}}{(2(m+1))!} u^{2(m+1),x}(\mathbf{x}_{1,i}) \\ &= \frac{1}{h_x^2} \left( u(\mathbf{x}_{1,i} - h_x \cdot \mathbf{e}_x) - 2u(\mathbf{x}_{1,i}) + u(\mathbf{x}_{1,i} + h_x \cdot \mathbf{e}_x) \right) \\ &= \frac{1}{h_x^2} \left( u(\mathcal{N}_1^-(\mathbf{x}_{1,i})) - 2u(\mathbf{x}_{1,i}) + u(\mathcal{N}_1^+(\mathbf{x}_{1,i})) \right) \end{aligned}$$

An einem inneren Gitterpunkt  $\mathbf{x}_{1,i}$  erhält man nach Anwendung des Operators  $\mathbf{H}_1$  den Term

$$\mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) = u(\mathbf{x}_{1,i}) - \frac{1}{2} \left( u(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + u(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right)$$

Analoge Ergebnisse erhält man an den Gitterpunkten  $\mathcal{N}_1^-(\mathbf{x}_{1,i}) = \mathbf{x}_{1,i} - h_x \cdot \mathbf{e}_x$  und  $\mathcal{N}_1^+(\mathbf{x}_{1,i}) = \mathbf{x}_{1,i} + h_x \cdot \mathbf{e}_x$ . Die Kombination dieser Ausdrücke mittels des Finiten Differenzen Sterns für die zweite Ableitung führt auf die Formel

$$\begin{aligned} \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u(\mathbf{x}_{1,i}) &= \\ &= \frac{1}{h_x^2} \left[ \left( u(\mathcal{N}_1^-(\mathbf{x}_{1,i})) - 2u(\mathbf{x}_{1,i}) + u(\mathcal{N}_1^+(\mathbf{x}_{1,i})) \right) \right. \\ &\quad \left. \frac{u(\mathcal{H}_2^-(\mathcal{N}_1^-(\mathbf{x}_{1,i}))) - 2u(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + u(\mathcal{H}_2^-(\mathcal{N}_1^+(\mathbf{x}_{1,i})))}{2} \right. \\ &\quad \left. \frac{u(\mathcal{H}_2^+(\mathcal{N}_1^-(\mathbf{x}_{1,i}))) - 2u(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + u(\mathcal{H}_2^+(\mathcal{N}_1^+(\mathbf{x}_{1,i})))}{2} \right] \end{aligned}$$

Für den weiteren Verlauf des Beweises stützen wir uns auf die Induktion nach dem Level bzgl. der  $y$ -Richtung. Dabei gehen wir davon aus, daß die Aussage des Satzes bereits für einen Level  $l_2$  bewiesen ist. Dies können wir an dieser Stelle tun, da wir bereits festgestellt haben, daß die Aussage für  $l_2 = 0$  Gültigkeit besitzt. Die Induktionsbehauptung ist nun, daß die Behauptung auch auf dem Level  $l_2 \geq 1$  gilt, falls sie bereits auf dem Level  $l_2 - 1$  richtig war.

Nach Induktionsvoraussetzung haben wir an den Gitterpunkten  $\mathcal{H}_2^-(\mathbf{x}_{1,i})$  und  $\mathcal{H}_2^+(\mathbf{x}_{1,i})$  die Behauptung bereits bewiesen und dort die Terme

$$\begin{aligned} \mathbf{E}_1 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u\left(\mathcal{H}_2^-(\mathbf{x}_{1,i})\right) &= \\ &= \frac{1}{h_x^2} \left( u\left(\mathcal{H}_2^-\left(\mathcal{N}_1^-(\mathbf{x}_{1,i})\right)\right) - 2u\left(\mathcal{H}_2^-(\mathbf{x}_{1,i})\right) + u\left(\mathcal{H}_2^-\left(\mathcal{N}_1^+(\mathbf{x}_{1,i})\right)\right) \right) \end{aligned}$$

sowie

$$\begin{aligned} \mathbf{E}_1 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u\left(\mathcal{H}_2^+(\mathbf{x}_{1,i})\right) &= \\ &= \frac{1}{h_x^2} \left( u\left(\mathcal{H}_2^+\left(\mathcal{N}_1^-(\mathbf{x}_{1,i})\right)\right) - 2u\left(\mathcal{H}_2^+(\mathbf{x}_{1,i})\right) + u\left(\mathcal{H}_2^+\left(\mathcal{N}_1^+(\mathbf{x}_{1,i})\right)\right) \right) , \end{aligned}$$

errechnet. Wendet man nun den Operator  $\mathbf{E}_1$  auf die Teilergebnisse an, so erhält man den Term

$$\begin{aligned} \mathbf{E}_1 \circ \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u\left(\mathbf{x}_{1,i}\right) &= \\ &= \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u\left(\mathbf{x}_{1,i}\right) + \\ &\quad + \frac{1}{2} \left( \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u\left(\mathcal{H}_2^-(\mathbf{x}_{1,i})\right) + \mathbf{D}_{11} \circ \mathbf{H}_1 \circ u\left(\mathcal{H}_2^+(\mathbf{x}_{1,i})\right) \right) \\ &= \frac{1}{h_x^2} \left( u\left(\mathcal{N}_1^-(\mathbf{x}_{1,i})\right) - 2u\left(\mathbf{x}_{1,i}\right) + u\left(\mathcal{N}_1^+(\mathbf{x}_{1,i})\right) \right) \end{aligned}$$

als Ergebnis des Finiten Differenzen Operators für dünne Gitter. Da dies gerade das Ergebnis für die gewöhnliche Finite Differenz auf vollen Gittern darstellt, ist die Aussage des Satzes für  $d = 2$  bewiesen.

Nun können wir zum Beweis der Aussage in  $d$  Dimensionen übergehen. Wie zuvor beschränken wir uns auf den Nachweis für nur eine zweite Ableitung, hier  $\mathbf{D}_{i_d i_d}^S$ . Hierzu benutzen wir nachfolgend eine Induktion über die Dimensionen, so daß der Beweis nun zwei Induktionen ineinander schachtelt. Betrachten wir zunächst einen Punkt auf dem Rand. Falls  $l_j = 0$  für  $2 \leq j \leq d$  gilt, haben wir den gewöhnlichen Differenzenstern anzuwenden, da die anzuwendenden Hierarchisierungen und Enthierarchisierungen für diese Punkte wie die Identität wirken. Ist  $l_j \neq 0$  für genau ein  $j$  für  $2 \leq j \leq d$ , so fallen wir auf den zweidimensionalen Fall zurück, für den wir bereits die Aussage

bewiesen haben. Analog gelangt man zur entsprechenden Aussage in drei Dimensionen, wenn  $l_j \neq 0$  für genau zwei Indizes  $j$  für  $2 \leq j \leq d$  gilt. Somit kann man sich mittels der Induktion Dimension für Dimension vorarbeiten.

*Induktionsvoraussetzung:* Die Behauptung des Satzes ist für eine Dimension  $\tilde{d} = d - 1$  erfüllt. Die Richtigkeit der Aussage wurde für die Dimension  $\tilde{d} = 2$  bereits gezeigt, für  $\tilde{d} = 1$  ist die Behauptung trivial.

*Induktionsbehauptung:* Die Behauptung des Satzes gilt für die Dimension  $d$ , wenn sie für die Dimension  $d - 1$  gilt.

*Induktionsbeweis:* Der Ausgangspunkt für die Induktion ist die Tatsache, daß die Behauptung für die Dimension  $d - 1$  und alle kleineren Dimensionen Gültigkeit besitzt. Damit ist für einen beliebigen Randpunkt, für den  $l_j = 0$  für mindestens ein  $j$  für  $2 \leq j \leq d$  gilt, im Falle von  $d$  Dimensionen die Behauptung bereits bewiesen, da man die Aussage dann für eine Dimension  $\leq d$  zu beweisen hätte und diese somit nach Induktionsvoraussetzung bereits bewiesen ist. Damit ist für Randpunkte nichts zu beweisen und wir können uns auf innere Gitterpunkte beschränken.

Die Anwendung der Hierarchisierung  $\mathbf{H}_{i_d}$  in alle Koordinatenrichtungen außer  $i_d$  führt für den Gitterpunkt  $\mathbf{x}_{1,i}$  auf

$$\begin{aligned} \mathbf{H}_{i_d} \circ u(\mathbf{x}_{1,i}) &= \\ &= u(\mathbf{x}_{1,i}) - \frac{1}{2} \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \left( u(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i})) + u(\mathcal{H}_{i_1}^+(\mathbf{x}_{1,i})) \right) \\ &\quad + \frac{1}{4} \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \sum_{\substack{i_2=1 \\ i_2 \notin \{i_1, i_d\}}}^d \left( u(\mathcal{H}_{i_2}^-(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i}))) + u(\mathcal{H}_{i_2}^-(\mathcal{H}_{i_1}^+(\mathbf{x}_{1,i}))) + \right. \\ &\quad \left. + u(\mathcal{H}_{i_2}^+(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i}))) + u(\mathcal{H}_{i_2}^+(\mathcal{H}_{i_1}^+(\mathbf{x}_{1,i}))) \right) + \dots + \\ &\quad + \left( -\frac{1}{2} \right)^{d-1} \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \dots \sum_{\substack{i_{d-1}=1 \\ i_{d-1} \notin I_{d-1}}}^d u(\mathcal{H}_{i_{d-1}}^\pm(\mathcal{H}_{i_{d-2}}^\pm(\dots(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i})))\dots))) \end{aligned}$$

mit  $I_{d-1} = \{i_1, \dots, i_{d-2}, i_d\}$  und

$$\begin{aligned} u(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i}))) &= u(\mathcal{H}_{i_2}^-(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i}))) + u(\mathcal{H}_{i_2}^-(\mathcal{H}_{i_1}^+(\mathbf{x}_{1,i}))) + \\ &\quad u(\mathcal{H}_{i_2}^+(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i}))) + u(\mathcal{H}_{i_2}^+(\mathcal{H}_{i_1}^+(\mathbf{x}_{1,i}))) \end{aligned}$$

Analog erhält man die Formeln, wenn mehrere Operatoren für hierarchische Vaterknoten (vgl. Definition 4.2) in einem Ausdruck involviert sind. Analoge Formeln erzielt man an den Gitterpunkten  $\mathbf{x}_{1,i} - h_{i_d} \mathbf{e}_{i_d}$  und  $\mathbf{x}_{1,i} + h_{i_d} \mathbf{e}_{i_d}$ , die bei der Anwendung der Finiten Differenz benötigt werden. Durch die

Anwendung des Finite Differenzen Sterns kombinieren wir die Ausdrücke an den Punkten  $\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i}) = \mathbf{x}_{1,i} - h_{i_d} \mathbf{e}_{i_d}$ ,  $\mathbf{x}_{1,i}$  und  $\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i}) = \mathbf{x}_{1,i} + h_{i_d} \mathbf{e}_{i_d}$  und erhalten somit

$$\begin{aligned}
& \mathbf{D}_{i_d i_d} \mathbf{H}_{i_d} \circ u(\mathbf{x}_{1,i}) = \\
&= \frac{1}{h_{i_d}^2} \left( \mathbf{H}_{i_d} \circ u(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i})) - 2\mathbf{H}_{i_d} \circ u(\mathbf{x}_{1,i}) + \mathbf{H}_{i_d} \circ u(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i})) \right) \\
&= \frac{1}{h_{i_d}^2} \left( u(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i})) - 2u(\mathbf{x}_{1,i}) + u(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i})) \right) - \\
&\quad - \frac{1}{2 \cdot h_{i_d}^2} \left( \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d u(\mathcal{H}_{i_1}^\pm(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i}))) - 2u(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i})) \right. \\
&\quad \quad \left. + u(\mathcal{H}_{i_1}^\pm(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i}))) \right) + \\
&\quad + \frac{1}{4 \cdot h_{i_d}^2} \left( \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \sum_{\substack{i_2=1 \\ i_2 \notin \{i_1, i_d\}}}^d u(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i})))) \right. \\
&\quad \quad \left. - 2u(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i}))) + u(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i})))) \right) + \dots + \\
&\quad + \frac{1}{(-2)^{d-1} \cdot h_{i_d}^2} \left( \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \cdots \sum_{\substack{i_{d-1}=1 \\ i_{d-1} \notin I_{d-1}}}^d \right. \\
&\quad \quad u(\mathcal{H}_{i_{d-1}}^\pm(\mathcal{H}_{i_{d-2}}^\pm(\cdots(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i}))))\cdots))) \\
&\quad \quad - 2u(\mathcal{H}_{i_{d-1}}^\pm(\mathcal{H}_{i_{d-2}}^\pm(\cdots(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i}))))\cdots))) \\
&\quad \quad \left. u(\mathcal{H}_{i_{d-1}}^\pm(\mathcal{H}_{i_{d-2}}^\pm(\cdots(\mathcal{H}_{i_2}^\pm(\mathcal{H}_{i_1}^\pm(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i}))))\cdots))) \right)
\end{aligned}$$

Der Ausgangspunkt war hierbei, daß nach der Induktion über die Level für beliebige, aber feste Dimension  $d$  die Behauptung an allen hierarchischen Vaterknoten bereits als bewiesen vorausgesetzt werden kann. Zusätzlich sind im Fall eines vollen Gitters die Operationen  $\mathcal{N}_{j_1}^\pm(\cdot)$  für die Bestimmung der nächsten Nachbarn in Raumdimension  $j_1$  sowie  $\mathcal{H}_{j_2}^\pm(\cdot)$  für die Ermittlung der Vaterknoten in Raumdimension  $j_2$  vertauschbar, zum Beispiel gilt hier

die Identität

$$\begin{aligned}
\mathbf{D}_{i_d i_d}^S \circ u\left(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i})\right) &= \\
&= \frac{1}{h_{i_d}^2} \left( u\left(\mathcal{N}_{i_d}^-\left(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i})\right)\right) - 2u\left(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i})\right) + u\left(\mathcal{N}_{i_d}^+\left(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i})\right)\right) \right) \\
&= \frac{1}{h_{i_d}^2} \left( u\left(\mathcal{H}_{i_1}^-\left(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i})\right)\right) - 2u\left(\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i})\right) + u\left(\mathcal{H}_{i_1}^-\left(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i})\right)\right) \right)
\end{aligned}$$

für das Ergebnis am Gitterpunkt  $\mathcal{H}_{i_1}^-(\mathbf{x}_{1,i})$ . Ähnliche Formeln ergeben sich für die hierarchischen Vaterknoten in die übrigen Raumdimensionen. Damit haben wir alle Zutaten für den Beweis in  $d$  Raumdimensionen gesammelt und erhalten abschließend

$$\begin{aligned}
\mathbf{E}_{i_d} \circ \mathbf{D}_{i_d i_d} \circ \mathbf{H}_{i_d} \circ u\left(\mathbf{x}_{1,i}\right) &= \\
&= \mathbf{D}_{i_d i_d} \circ \mathbf{H}_{i_d} \circ u\left(\mathbf{x}_{1,i}\right) + \\
&\quad + \frac{1}{2} \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \mathbf{D}_{i_d i_d} \circ \mathbf{H}_{i_d} \circ u\left(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i})\right) - \\
&\quad - \frac{1}{4} \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \sum_{\substack{i_2=1 \\ i_2 \notin \{i_1, i_d\}}}^d \mathbf{D}_{i_d i_d} \circ \mathbf{H}_{i_d} \circ u\left(\mathcal{H}_{i_2}^\pm\left(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i})\right)\right) + \dots - \\
&\quad - \left(\frac{1}{2}\right)^{d-1} \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \dots \sum_{\substack{i_{d-1}=1 \\ i_{d-1} \notin I_{d-1}}}^d \mathbf{D}_{i_d i_d} \circ \mathbf{H}_{i_d} \circ \\
&\quad \circ u\left(\mathcal{H}_{i_{d-1}}^\pm\left(\mathcal{H}_{i_{d-2}}^\pm\left(\dots\left(\mathcal{H}_{i_2}^\pm\left(\mathcal{H}_{i_1}^\pm\left(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i})\right)\right)\right)\dots\right)\right)\right) \\
&= \mathbf{D}_{i_d i_d} \circ \mathbf{H}_{i_d} \circ u(\mathbf{x}_{1,i}) + \frac{1}{2 \cdot h_{i_d}^2} \left( \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d u\left(\mathcal{H}_{i_1}^\pm\left(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i})\right)\right) \right. \\
&\quad \left. - 2u\left(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i})\right) + u\left(\mathcal{H}_{i_1}^\pm\left(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i})\right)\right) \right) - \\
&\quad - \frac{1}{4 \cdot h_{i_d}^2} \left( \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \sum_{\substack{i_2=1 \\ i_2 \notin \{i_1, i_d\}}}^d u\left(\mathcal{H}_{i_2}^\pm\left(\mathcal{H}_{i_1}^\pm\left(\mathcal{N}_{i_d}^-(\mathbf{x}_{1,i})\right)\right)\right) \right. \\
&\quad \left. - 2u\left(\mathcal{H}_{i_2}^\pm\left(\mathcal{H}_{i_1}^\pm(\mathbf{x}_{1,i})\right)\right) + u\left(\mathcal{H}_{i_2}^\pm\left(\mathcal{H}_{i_1}^\pm\left(\mathcal{N}_{i_d}^+(\mathbf{x}_{1,i})\right)\right)\right) \right) + \dots -
\end{aligned}$$

$$\begin{aligned}
& -\frac{1}{(-2)^{d-1} \cdot h_{i_d}^2} \left( \sum_{\substack{i_1=1 \\ i_1 \neq i_d}}^d \cdots \sum_{\substack{i_{d-1}=1 \\ i_{d-1} \notin I_{d-1}}}^d \right. \\
& \quad u \left( \mathcal{H}_{i_{d-1}}^{\pm} \left( \mathcal{H}_{i_{d-2}}^{\pm} \left( \cdots \left( \mathcal{H}_{i_2}^{\pm} \left( \mathcal{H}_{i_1}^{\pm} \left( \mathcal{N}_{i_d}^{-}(\mathbf{x}_{1,i}) \right) \right) \right) \right) \right) \right) - \\
& \quad - 2u \left( \mathcal{H}_{i_{d-1}}^{\pm} \left( \mathcal{H}_{i_{d-2}}^{\pm} \left( \cdots \left( \mathcal{H}_{i_2}^{\pm} \left( \mathcal{H}_{i_1}^{\pm}(\mathbf{x}_{1,i}) \right) \right) \right) \right) \right) \\
& \quad \left. u \left( \mathcal{H}_{i_{d-1}}^{\pm} \left( \mathcal{H}_{i_{d-2}}^{\pm} \left( \cdots \left( \mathcal{H}_{i_2}^{\pm} \left( \mathcal{H}_{i_1}^{\pm} \left( \mathcal{N}_{i_d}^{+}(\mathbf{x}_{1,i}) \right) \right) \right) \right) \right) \right) \right) \\
& = \frac{1}{h_{i_d}^2} \left( u \left( \mathcal{N}_{i_d}^{-}(\mathbf{x}_{1,i}) \right) - 2u(\mathbf{x}_{1,i}) + u \left( \mathcal{N}_{i_d}^{+}(\mathbf{x}_{1,i}) \right) \right)
\end{aligned}$$

Dies ist aber gerade das Ergebnis der gewöhnlichen Finiten Differenz für die zweite Ableitung. Damit ist die Aussage des Satzes auch im Fall von  $d$  Dimensionen bewiesen.  $\square$

## A.9 Beweis zu Satz 4.13

**Beweis:** Wir betrachten den Beweis für Satz 4.13 lediglich für den zweidimensionalen Fall, der analog zum Beweis von Satz 4.1 läuft. Der Beweis für die Erweiterung auf den  $d$ -dimensionalen Fall funktioniert dabei analog wie der zugehörige Beweis für den Operator  $\mathbf{D}_{ii}^S$  der Konsistenzordnung zwei.

Wir beschränken uns in diesem Beweis auf den Operator  $\mathbf{D}_{11,4}^S$ , da die Beweise für beide Operatoren (also für  $\mathbf{D}_{11,4}^S$  und  $\mathbf{D}_{11,6}^S$ ) im wesentlichen identisch sind wie für den in Satz 4.1 bewiesenen Operator  $\mathbf{D}_{11}^S$ .

An Randpunkten  $\mathbf{x}_{1,i}$  mit  $l_2 = 0$ , an denen der Differenzenstern aufgesetzt werden kann, ergibt Taylorreihenentwicklung die Aussage

$$\begin{aligned}
v(\mathbf{x}_{1,i}) &= \sum_{m=0}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \frac{16 - 4 \cdot 2^{2m}}{12} u^{2(m+1),x}(\mathbf{x}_{1,i}) \\
&= u^{2,x}(\mathbf{x}_{1,i}) + \sum_{m=2}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \frac{16 - 4 \cdot 2^{2m}}{12} u^{2(m+1),x}(\mathbf{x}_{1,i})
\end{aligned}$$

und die Aussage des Satzes ist für diese Punkt bewiesen. Um analog wie zum Beweis von Satz 4.1 vorgehen zu können, benötigen wir neben den Werten (9.3), (9.4) und (9.5) noch die Werte



- an  $\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x$  :

$$\begin{aligned}
\tilde{u}(\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x) &:= u(\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x) - \frac{1}{2} \left( u(\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x + h_y \mathbf{e}_y) \right. \\
&\quad \left. + u(\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x - h_y \mathbf{e}_y) \right) \\
&= -\frac{1}{2} \sum_{j=0}^{\infty} \frac{2}{(2(j+1))!} h_y^{2(j+1)} u^{2(j+1),y}(\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x) \\
&= -\frac{1}{2} \sum_{j=1}^{\infty} \frac{2}{(2j)!} h_y^{2j} u^{2j,y}(\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x) \quad (9.36)
\end{aligned}$$

- an  $\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x$  :

$$\begin{aligned}
\tilde{u}(\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x) &:= u(\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x) - \frac{1}{2} \left( u(\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x + h_y \mathbf{e}_y) \right. \\
&\quad \left. + u(\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x - h_y \mathbf{e}_y) \right) \\
&= -\frac{1}{2} \sum_{j=0}^{\infty} \frac{2}{(2(j+1))!} h_y^{2(j+1)} u^{2(j+1),y}(\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x) \\
&= -\frac{1}{2} \sum_{j=1}^{\infty} \frac{2}{(2j)!} h_y^{2j} u^{2j,y}(\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x) \quad (9.37)
\end{aligned}$$

Wendet man nun den Finiten Differenzen Stern (4.35) an und setzt die entsprechenden Werte ein, so erhalten wir

$$\begin{aligned}
v(\mathbf{x}_{1,i}) &= \\
&= \frac{1}{12h_x^2} \left( \tilde{u}(\mathbf{x}_{1,i} - 2h_x \mathbf{e}_x) - 16\tilde{u}(\mathbf{x}_{1,i} - h_x \mathbf{e}_x) + 30\tilde{u}(\mathbf{x}_{1,i}) - \right. \\
&\quad \left. - 16\tilde{u}(\mathbf{x}_{1,i} + h_x \mathbf{e}_x) - \tilde{u}(\mathbf{x}_{1,i} + 2h_x \mathbf{e}_x) \right) \\
&= -\frac{1}{2} \sum_{m=0}^{\infty} \sum_{j=1}^{\infty} \frac{2h_x^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^m}{12} \cdot \frac{2h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \\
&\quad + \frac{1}{2} \left( u^{2,x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \right. \\
&\quad \left. + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} u^{2(m+1),x}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) \right) \\
&\quad + \frac{1}{2} \left( u^{2,x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \right. \\
&\quad \left. + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} u^{2(m+1),x}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) \right)
\end{aligned}$$

$$\begin{aligned}
&= - \sum_{m=0}^{\infty} \sum_{j=1}^{\infty} \frac{2(2^{l_2-1} \cdot h)^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^m}{12} \cdot \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \\
&\quad + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{(-h_y)^j}{j!} u^{2,x,j,y}(\mathbf{x}_{1,i}) + \mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \right. \\
&\quad \left. + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} \sum_{j=0}^{\infty} \frac{(-h_y)^j}{j!} u^{2(m+1),x,j,y}(\mathbf{x}_{1,i}) \right) \\
&\quad + \frac{1}{2} \left( \sum_{j=0}^{\infty} \frac{(h_y)^j}{j!} u^{2,x,j,y}(\mathbf{x}_{1,i}) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i})) + \right. \\
&\quad \left. + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} \sum_{j=0}^{\infty} \frac{(h_y)^j}{j!} u^{2(m+1),x}(\mathbf{x}_{1,i}) \right) \\
&= - \sum_{j=1}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{2,x,2j,y}(\mathbf{x}_{1,i}) \\
&\quad - \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{2(2^{l_2-1} \cdot h)^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^m}{12} \cdot \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \\
&\quad + \sum_{j=0}^{\infty} \frac{h_y^{2j}}{(2j)!} u^{2,x,2j,y}(\mathbf{x}_{1,i}) + \frac{1}{2} (\mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i}))) \\
&\quad + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} u^{2(m+1),x}(\mathbf{x}_{1,i}) \\
&\quad + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} \cdot \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \\
&= u^{2,x}(\mathbf{x}_{1,i}) + \sum_{m=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} u^{2(m+1),x}(\mathbf{x}_{1,i}) \\
&\quad + \frac{1}{2} (\mathcal{F}(\mathcal{H}_2^-(\mathbf{x}_{1,i})) + \mathcal{F}(\mathcal{H}_2^+(\mathbf{x}_{1,i}))) \\
&\quad + \sum_{m=1}^{\infty} \sum_{j=1}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} \cdot \left( 1 - (2^{l_2-1})^{2m} \right) \cdot \\
&\quad \quad \quad \cdot \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \\
&=: u^{2,x}(\mathbf{x}_{1,i}) + \sum_{m=2}^{\infty} \frac{2h^{2m}}{(2(m+1))!} \cdot \frac{16 - 4 \cdot 2^{2m}}{12} u^{2(m+1),x}(\mathbf{x}_{1,i}) \\
&\quad + \mathcal{F}(\mathbf{x}_{1,i}) \quad , \tag{9.38}
\end{aligned}$$

wobei wir den im Vergleich zur Situation auf einem vollen Gitter auftretenden zusätzlichen Fehlerterm  $\mathcal{F}(\mathbf{x}_{1,i})$  in (9.38) definiert haben. Im Fall  $l_2 = 0$

setzen wir für die Randpunkte, bei denen der Finite Differenzen Stern aufgesetzt werden kann,  $\mathcal{F}(\mathbf{x}_{1,i}) = 0$ . Damit ist wie beim Beweis des Satzes 4.1 die Aussage dieses Satzes bewiesen, wenn der Term  $\mathcal{F}(\mathbf{x}_{1,i})$  von der Größenordnung  $\mathcal{O}(h^4)$  ist, da die Summe in (9.38) den führenden Term  $h^4 u^{6,x}(\mathbf{x}_{1,i})/90$  besitzt.

Um die Ordnung  $\mathcal{O}(h^4)$  auch für den zusätzlichen Fehlerterm zu zeigen, unterscheiden wir wie im Beweis zu Satz 4.1 verschiedene Fälle, hier jedoch anstatt 4 nun insgesamt 10 Fälle, die wir nachfolgend auflisten wollen. Der einfacheren Notation wegen definieren wir für diesen Beweis die Größen

$$A = \frac{16}{12}, \quad B = \frac{4 \cdot 2^{2m}}{12}, \quad C = 1, \quad D = (2^{l_2-1})^{2m},$$

die in dem zusätzlichen Fehlerterm als Summanden der beiden Ausdrücke  $(A - B) \cdot (C - D)$  in den Klammern auftreten.

1. Fall: wir betrachten hier Terme der Form

$$\frac{4}{3} \frac{2h^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}), \quad (9.39)$$

die gerade den Anteil des Fehlerterms  $\mathcal{F}(\mathbf{x}_{1,i})$  ausmachen, wenn man in (9.38) die Klammern ausmultipliziert und hier jeweils die ersten beiden Summanden einer jeden Klammer betrachtet (Kombination  $A$  mit  $C$ ). Bei den Termen in (9.39) ist die Fehlerordnung direkt ablesbar. Dabei unterscheiden sich die Terme in (9.39) und (9.9) lediglich um den konstanten Faktor  $4/3$ .

2. Fall: wir betrachten hier Terme der Form

$$-\frac{2}{3} \frac{(2h)^{2m}}{(2(m+1))!} \frac{h_y^{2j}}{(2j)!} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}), \quad (9.40)$$

die die Kombination  $B/C$  beim Ausmultiplizieren der Klammern des Fehlerterms darstellen. Auch bei den Termen in (9.40) ist die Ordnung direkt ablesbar. Dabei unterscheiden sich die Terme in (9.40) und (9.9) lediglich um den konstanten Faktor  $-1/3$  sowie bezüglich der Potenzen von  $2h$  anstatt von  $h$ .

3. Fall (a-d): hier betrachten wir die Kombination  $A/D$ , die zu den Ausdrücken

$$-\frac{8h^{2m}}{3(2(m+1))!} \frac{h_y^{2j}}{(2j)!} (2^{l_2-1})^{2m} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}), \quad (9.41)$$

führt. Wir müssen die Betrachtung dieses Falles aufspalten in die 4 Fälle (3a)  $j = m$ , (3b)  $1 < j < m$ , (3c)  $j > m$  sowie (3d)  $j =$

1. Dabei stimmen die Ergebnisse, die man in den ersten drei Fällen erhält, mit denen in (9.11), (9.12) und (9.13) bis auf eine allen Termen gemeinsame Konstante überein. Betrachtet werden muß hier lediglich der Fall  $j = 1$ . Dabei ist jedoch klar, daß für einen beliebigen, aber fest gewählten Gitterpunkt eines dünnen Gitters der Ausdruck

$$\frac{8(2^{l_2-1})^{2m}}{3(2(m+1))!} \frac{h_y^2}{2!} h^{2m} \quad (9.42)$$

von der Lage des Gitterpunkts abhängt, bezüglich den Potenzen von  $h$  jedoch konstant ist (man betrachte  $h \rightarrow h/2$ ). Somit ist auch hier die  $h$ -Unabhängigkeit der Koeffizienten der Potenzen von  $h$  gegeben. Zudem handelt es sich hier für jede Potenz in  $h$  nur um einen einzigen Term, eine Summation bei der Bestimmung des Koeffizienten vor den  $h$ -Potenzen im Gesamtausdruck ist im Gegensatz zu den ersten drei Fällen also nicht notwendig.

4. Fall (a-d): hier werden die Kombination  $B/D$  behandelt, die zu den Ausdrücken

$$\frac{2(2h)^{2m}}{3(2(m+1))!} \frac{h_y^{2j}}{(2j)!} (2^{l_2-1})^{2m} u^{2(m+1),x,2j,y}(\mathbf{x}_{1,i}) \quad , \quad (9.43)$$

führt. Wie zuvor in Fall 3 werden auch hier die 4 Fälle (4a)  $j = m$ , (4b)  $1 < j < m$ , (4c)  $j > m$  sowie (4d)  $j = 1$  untersucht. Dabei unterscheiden sich alle 4 hier betrachteten Fälle lediglich um einen allen gemeinsamen konstanten Faktor sowie dem Austausch von  $h$  durch  $2h$  im Vergleich mit den 4 Fällen (3a)–(3d). Die Koeffizienten für  $j = 1$  ergeben sich hier zu

$$\frac{2(2^{l_2-1})^{2m}}{3(2(m+1))!} \frac{h_y^2}{2!} (2h)^{2m} \quad . \quad (9.44)$$

Somit ist die  $h$ -Unabhängigkeit der Koeffizienten auch für alle 4 hier angesprochenen Fälle gewährleistet.

Da sich die angegebenen Terme der einzelnen Fälle stets lediglich um eine Konstante oder dem Austausch von  $h$  durch  $2h$  von den entsprechenden Termen im Beweis zu Satz 4.1 unterscheiden, ergeben sich bei der Abschätzung der Koeffizienten ähnliche Größen, die sich ebenfalls nur durch Konstanten oder  $2h$  anstelle von  $h$  von den Resultaten unterscheiden, die am Ende des Beweises zu Satz 4.1 gewonnen worden sind. Folglich sind auch hier die Koeffizienten im “worst case” vor jeder Potenz von  $h$  bestimmt, da in jedem der 10 Fälle die Koeffizienten unabhängig von  $h$  sind.

Damit folgt nun insgesamt die Behauptung

$$\| \mathbf{D}_{i,i,4}^S \mathbf{R}_{1,i} u - \tilde{\mathbf{R}}_{1,i} \frac{\partial^4}{\partial x_i^4} \|_\infty \leq C \cdot h^4 \|u\|_{C^\infty(\bar{\Omega})} \quad (9.45)$$

für die zweite Ableitung bezüglich der Raumdimension  $i$  in zwei Dimensionen ( $1 \leq i \leq 2$ ), was die Aussage des Satzes beweist.  $\square$

## Literatur

- [AACHNS] ABELSON H., ADAMS N. I., COORE D., HANSON C., NAGPAL R., SUSSMAN G. J.: *Gray Scott Model of Reaction Diffusion*,  
<http://www-swiss.ai.mit.edu/~switz/amorphous/GrayScott/index.html>
- [AbEg98] ABRASHIN V. N., EGOROV A. A.: *Iterative Methods of Domain Decomposition in Solving stationary Problems of mathematical Physics*, Proceedings of the Second International Conference “Finite Difference Methods: Theory and Applications” (CFDM98), Vol. 1, pp. 7–11, Minsk, Weißrußland, eds. A. A. Samarskii, 1998
- [AHET91] ARCILLA A., HÄUSER J., EISEMAN P., THOMPSON J.: *Numerical Grid Generation in Computational Fluid Dynamics and Related Fields*, North-Holland, Amsterdam, 1991
- [Bab60a] BABENKO K. I.: *Approximation of periodic Functions of many Variables by trigonometric Polynomials*, Dokl. Akad. Nauk SSSR, Vol. 132, pp. 247–250, 1960
- [Bab60b] BABENKO K. I.: *Approximation by trigonometrical Polynomials in a certain Class of periodic Functions of several Variables*, Soviet Math. Dokl., Vol. 1, pp. 672–675, 1960
- [Bal94] BALDER R.: *Adaptive Verfahren für elliptische und parabolische Differentialgleichungen auf dünnen Gittern*, Dissertation, Technische Universität München, 1994
- [BaZe92] BALDER R., ZENGER CHR.: *The d-dimensional Helmholtz equation on sparse Grids*, Institut für Informatik der Technischen Universität München, SFB-Report 342/21/92 A, 1992
- [Bra77] BRANDT A.: *Multi-Level Adaptive Techniques (MLAT) for Partial Differential Equations: Ideas and Software*, Mathematical Software III (Ed. Rice .J. R.), Academic Press, New York, pp. 273–314, 1977
- [Bun92] BUNGARTZ, H.-J.: *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*, Dissertation, Technische Universität München, 1992
- [Bun98] BUNGARTZ H.-J.: *Finite Elements of Higher Order on Sparse Grids*, Habilitation, Technische Universität München, 1998
- [BuDo97] BUNGARTZ H.-J., DORNSEIFER T.: *Sparse grids: Recent developments for elliptic partial differential equations*, Institut für Informatik der Technischen Universität München, SFB-Report 342/02/97 A, 1997

- [BDZ96] BUNGARTZ H.-J., DORNSEIFER T. AND ZENGER CHR.: *Tensor product approximation spaces for the efficient numerical solution of partial differential equations*, erscheint in Proc. Int. Workshop on Scientific Computations, Konya, 1996, Nova Science Publishers, Inc.
- [BuGr97] BUNGARTZ H.-J., GRIEBEL M.: *A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives*, eingereicht bei J. Complexity, 1997 (auch als: SFB-Report 256, Nummer 254, Universität Bonn, 1997)
- [CaLi98] CALVO P., LISBONA F.: *Finite Difference Simulation of Capillary Barriers in Infiltration Processes*, Proceedings of the Second International Conference "Finite Difference Methods: Theory and Applications" (CFDM98), Vol. 1, pp. 53–62, Minsk, Weißrußland, eds. A. A. Samarskii, 1998
- [CHSS95] CASTILLO J. E., HYMAN J. M., SHASHKOV M. J. AND STEINBERG S.: *High-Order Mimetic Finite Difference Methods on Nonuniform Grids*, in ICOSAHOM-95, Proc. of the Third International Conference on Spectral and High Order Methods, Houston, Texas, 5-9 June 1995 (Special Issue of Houston Journal of Mathematics, eds. A. V. Ilin and L. R. Scott, 1995, pp. 347-361)
- [Cho67] CHORIN A. J.: *A Numerical Method for Solving Incompressible Viscous Flow Problems*, J. of Comp. Phys., Vol. 2, pp 12–26, 1967
- [Cho68] CHORIN A.J.: *Numerical Solution of the Navier Stokes Equations*, J. Math. Comput., Vol. 22, pp 745–762, 1968
- [CVD97] BODE A., DAUELSBERG M., DURST F., GRIEBEL M., JÜRGENSEN H., KADINSKI L., LUKSCH P., MAY M., MAKAROV Y., SCHIEKOFER T., STRAUCH G.: *PAR-CVD: Entwicklung leistungsfähiger paralleler Berechnungsverfahren zur Untersuchung und Optimierung von CVD-Prozessen*, in Proceedings der BMWF-Tagung HPSC97, München, 1997
- [CVD99] BODE A., DAUELSBERG M., DURST F., GRIEBEL M., JÜRGENSEN H., KADINSKI L., LUKSCH P., MAY M., MAKAROV Y., SCHIEKOFER T., STRAUCH G.: *PAR-CVD: Entwicklung leistungsfähiger paralleler Berechnungsverfahren zur Untersuchung und Optimierung von CVD-Prozessen*, Abschlußbericht, Erlangen/Bonn/München/Aachen, 1995–1999 (in Druck)
- [Dau92] DAUBECHIES I.: *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, 61, SIAM, Philadelphia, 1992
- [Dau95] DAUELSBERG M.: *Entwicklung mathematischer Modelle und numerische Simulation der MOCVD von Verbindungshalbleitern*, Diplomarbeit

- beit, Lehrstuhl für Strömungsmechanik, Universität Erlangen–Nürnberg, 1995
- [DeSc89] DELVOS F.-J., SCHEMPP W.: *Boolean methods in interpolation and approximation*, Pitman Research Notes in Mathematics Series 230, Longman, Essex 1989
- [Dor97] DORNSEIFER T.: *Diskretisierung allgemeiner elliptischer Differentialgleichungen in krummlinigen Koordinatensystemen auf dünnen Gittern*, Dissertation, Technischen Universität München, 1997
- [ELW] ESTEP D. J., LARSON M. G., WILLIAMS R. D.: *Estimating the Error of Numerical Solutions of Systems of nonlinear Reaction–Diffusion Equations*, California Institute of Technology, CACR Center for Advanced Computing Research, Report CACR–145, 1997
- [Fab09] FABER G.: *Über stetige Funktionen*, Mathematische Annalen, Vol. 66, pp. 81–94, 1909
- [FePe96] FERZINGER J.H., PERIC M.: *Computational Methods for Fluid Dynamics*, Springer, Berlin/Heidelberg 1996
- [Fle91] FLETCHER C. A. J.: *Computational Techniques for Fluid Dynamics – Volume II*, Springer Series in Computational Physics, Berlin/Heidelberg 1991
- [For88] FORNBERG B.: *Generation of Finite Difference Formulas on Arbitrarily Spaced Grids*, Math. Comput., Vol. 51(184), pp. 699–706, 1988
- [For97] FORNBERG B.: *Calculation of weights in finite difference formulas*, erscheint in SIAM Review Classroom Notes
- [FPT97] FORTIN M., PEYRET R., TEMAM R.: , J. M'ec., Vol. 10, pp. 357–390, 1971 (auch in Lecture Notes in Physics, Vol. 8, pp. 337–342, Springer Verlag, New York, 1972)
- [GGS82] GHIA U., GHIA K. N., SHIN C. T.: *High-Re Solutions for Incompressible Flow Using the Navier–Stokes Equations and a Multigrid Method*, J. of Comp. Phys., Vol. 48, pp. 387–411, 1982
- [GSS] GRAY SCOTT SIMULATOR:  
<http://www-swiss.ai.mit.edu/~switz/amorphous/jsim/sim/GrayScott.html>
- [Gre87] GRESHO P. M., SANI R. L.: *On Pressure Boundary Conditions for the Incompressible Navier–Stokes–Equations*, International Journal for Numerical Methods in Fluids, auch: Lawrence Livermore National Laboratory, Preprint UCRL-96471



- [Gri90] GRIEBEL M.: *Zur Lösung von Finite-Differenzen- und Finite-Element-Gleichungen mittels der Hierarchischen-Transformations-Mehrgitter-Methode*, Institut für Informatik der Technischen Universität München, SFB-Report 342/4/90 A, 1990
- [Gri91a] GRIEBEL M.: *A parallelizable and vectorizable multi-level algorithm on sparse grids in Parallel Algorithms for partial differential equations*, Notes on Numerical Fluid Mechanics, Vol. 31, pp. 94-100, W. Hackbusch, ed., Vieweg, Braunschweig, 1991
- [Gri91b] GRIEBEL M.: *Parallel Multigrid Methods on Sparse Grids*, in Multigrid Methods III, Proceedings of the 3rd European Conference on Multigrid Methods, Bonn, Oktober 1990, International Series of Numerical Mathematics (Vol. 98), W. Hackbusch and U. Trottenberg, ed., Birkhäuser, Basel, 1991 (auch: Institut für Informatik der Technischen Universität München, SFB-Report 342/30/90 A)
- [Gri97] GRIEBEL M.: *Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences*, Computing, also Notes on Numerical Fluid Mechanics, erscheint in Proceedings Large Scale Scientific Computations, 7. June - 11. June, 1997, Varna, Bulgaria, Vieweg-Verlag, Braunschweig, 1998
- [GrDoNe98] GRIEBEL M., DORNSEIFER T., NEUNHOEFFER T.: *Numerical Simulation in Fluid Dynamics – a Practical Introduction*, SIAM, Philadelphia, 1998
- [GrKn98] GRIEBEL M., KNAPEK S.: *Optimized approximation spaces for operator equations*, Constructive Approximation, submitted, 1998 (auch: SF-Report 256, Nr. 568, Universität Bonn, 1998)
- [GKS98] GRIEBEL M., KNAPEK S., SCHIEKOFER T.: *Galerkin discretization of operator equations using wavelets of tensor product type II: Compression*, Universität Bonn, in Vorbereitung
- [GOS98] GRIEBEL M., OSWALD P., SCHIEKOFER T.: *Sparse grids for boundary integral equations*, angenommen bei Numerische Mathematik (auch: SFB-Report 256, Nr. 554, Universität Bonn, 1998)
- [GrSc98] GRIEBEL M., SCHIEKOFER T.: *An adaptive sparse grid Navier Stokes solver in 3D based on the finite difference approach*, erscheint in Proceedings of ENUMATH 97, Wiley, 1998
- [GrZu98] GRIEBEL M., ZUMBUSCH G.: *Adaptive Sparse Grids for Hyperbolic Conservation Laws*, Proceedings of Seventh International Conference on Hyperbolic Problems, Theory, Numerics, Applications, Birkhäuser, 1998

- [GrRo92] GROSSMANN CH., ROOS H.-G.: *Numerik partieller Differentialgleichungen*, Teubner Studienbücher, Stuttgart 1992
- [Hac86] HACKBUSCH W.: *Theorie und Numerik elliptischer Differentialgleichungen*, Teubner, Stuttgart 1986
- [Hal92] HALLATSCHKE K.: *Fouriertransformation auf dünnen Gittern mit hierarchischen Basen*, Numer. Math., Vol. 63(1), pp. 83–97, 1992
- [HäTa86] HÄUSER J., TAYLOR C.: *Numerical Grid Generation in Computational Dynamics*, Pineridge Press, Swansea, 1986
- [HoJo91] HORN R. A., JOHNSON C. R.: *Topics in matrix analysis*, Cambridge University Press, Cambridge/New York, 1991
- [JMT97] JOHN V., MAUBACH J. M., TOBISKA L.: *Nonconforming streamline-diffusion-finite-element-methods for convection-diffusion problems*, Numer. Math., 78, pp. 165–188, 1997
- [Kač85] KAČUR J.: *Method of Rothe in evolution equations*, Teubner, Leipzig, 1985
- [Kad96] KADINSKI L.: *Mathematische Modellierung und numerische Simulation von CVD-Prozessen in der Halbleitertechnik*, Dissertation, Universität Erlangen–Nürnberg, 1996
- [KnSt93] KNUPP P., STEINBERG S.: *Fundamentals of Grid Generation*, CRC Press Inc., London, 1993
- [Knu73] KNUTH D. W.: *The Art of Computer Programming, Volume III: Sorting and Searching*, Addison-Wesley, 1973
- [Kor61] KOROBOW N. M.: *Properties and Calculation of optimal Coefficients*, Trudy Mat. Inst. Steklov, Vol. 60, pp. 195–199, 1961
- [LMOS93] LEE K. J., MCCORMICK W. D., OUYANG Q., SWINNEY H. L.: *Pattern Formation by Interacting Chemical Fronts*, Science, Vol. 261, pp. 192–194, 1993
- [LMR94] LOUIS A. K., MAASS P., RIEDER A.: *Wavelets*, Teubner Studienbücher, Stuttgart 1994
- [MaSt84] MANOHAR R., STEPHENSON J. W.: *High order difference schemes for linear partial differential equations*, SIAM J. Sci. Stat. Comput., Vol. 5(1), pp. 69–78, 1984
- [Mat97] MATLAB 5.2: The MathWorks, Inc.

- [MaSc97] MAY M., SCHIEKOFER T.: *An Abstract Data Type for Parallel Simulations based on Sparse Grids*, in Proceedings of the 3<sup>rd</sup> European PVM Conference, Munich, Germany, October 7-9th 1996, Bode A., Dongarra J., Ludwig T. und Sunderam V., Editoren, Springer Verlag, Berlin/Heidelberg, Lecture Notes in Computer Science, Vol. 1156, pp. 59–67, 1997
- [MaCh98] MAZHUKIN V. I., CHUIKO M. M.: *Solution of two-dimensional Stefan Problem by the Method of dynamic Adaption*, Proceedings of the Second International Conference “Finite Difference Methods: Theory and Applications” (CFDM98), Vol. 2, pp. 135–145, Minsk, Weißrußland, eds. A. A. Samarskii, 1998
- [Meh88] MEHLHORN K.: *Datenstrukturen und Effiziente Algorithmen: Band 1, Sortieren und Suchen*, Teubner Verlag, 1988
- [Mei92] MEIMANOV A. M.: *The Stefan Problem*, de Gruyter Expositions in Mathematics, de Gruyter, Bonn, 1992
- [Mey80] MEYBERG K.: *Algebra Teil 1*, Hanser Verlag, 1980, 2. Auflage
- [OMP] SILICON GRAPHICS DEVELOPER’S TOOLBOX : *OpenMP API (application programming interface)*, <https://toolbox.sgi.com/toolbox/>
- [Pea93] PEARSON J. E. : *Complex Patterns in a Simple System Science*, Vol. 261, pp. 189–192, 1993
- [Per89] PEREVERZEV S. V.: *An estimate of the complexity of the approximate solution of Fredholm equations of the second kind with differentiable kernels*, Ukr. Math. J., 41(2), pp. 1225–1227, 1989
- [PeTa90] PEYRET T., TAYLOR T. D.: *Computational Methods for Fluid Flow*, Springer Series in Computational Physics (corrected third printing), Stuttgart 1990
- [Pfl96] PFLAUM C.: *Diskretisierung elliptischer Differentialgleichungen mit dünnen Gittern*, Dissertation, Technische Universität München, 1996
- [Pro97] PROHL A.: *Projection and Quasi-Compressibility Methods for Solving the Incompressible Navier–Stokes Equations*, Teubner, Stuttgart 1997
- [Rek82] REKTORYS K.: *The method of discretization in time and partial differential equations*, Dordrecht, Boston, 1982
- [RiMo67] RICHTMYER R. D., MORTON K. W.: *Difference Methods for Initial-Value Problems*, Wiley–Interscience, 1967

- [ScZu98] SCHIEKOFER T., ZUMBUSCH G.: *Software concepts of a sparse grid finite difference code*, in Proceedings of the 14th GAMM-Seminar Kiel on Concepts of Numerical Software, 1998, (in Druck)
- [Sch70] SCHWARZ H. A.: *Gesammelte Mathematische Abhandlungen*, Nummer 2, pp. 133-143, Springer, Berlin, 1890  
Erstpublikation: Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich, Vol. 15, pp. 272-286, 1870
- [Sid95] SIDILKOVER D., ASCHER U. M.: *A Multigrid Solver for the Steady State Navier–Stokes–Equations using the Pressure–Poisson Formulation*, *Matematica Aplicada e Computational*, Vol. 14, pp. 249-274, 1995
- [SFB256] SFB 256, Universität Bonn: *Grape Manual*,  
<http://www.iam.uni-bonn.de/sfg256/grape>
- [SG] SILICON GRAPHICS: *C Language Reference Manual*, Chapter 11, Multiprocessing C/C++ Compiler Directives
- [SBG96] SMITH B., BJORSTAD P., GROPP W.: *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, 1996
- [Smi85] SMITH G. D.: *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford Applied Mathematics and Computing Science Series, 3rd edition, Clarendon Press, Oxford 1985
- [Smo60] SMOLYAK S. A.: *Interpolation and Quadrature Formulas for the Classes  $W_s^a$  and  $E_s^a$* , *Soviet Math. Dokl.*, Vol. 1, pp. 384–387, 1960
- [Smo63] SMOLYAK S. A.: *Quadrature and interpolation formulas for tensor products of certain classes of functions*, *Dokl. Akad. Nauk SSSR*, Vol. 4, pp. 240–243, 1963
- [Spr97] SPRENGEL F.: *Periodic interpolation and wavelets on sparse grids*, *Numerical Algorithms*, Vol. 17, pp. 147–169, 1998
- [StKu77] STEGER J. L., KUTLER P.: , *AIAA J.* 15, pp. 581-590, 1977
- [Str89] STRIKWERDA J. C.: *Finite Difference Schemes and Partial Differential Equations*, The Wadsworth & Brooks/Cole Mathematics Series, Pacific Grove, California, 1989
- [Tem69] TEMAM R.: *Sur l'approximation de la solution des 'equations de Navier–Stokes par la m'ethod des pas fractionnaires*, *Archiv. Ration. Mech. Anal.*, Vol. 32, pp. 135-153, 1969

- [Tem94] TEMLYAKOV V. N.: *Approximation of periodic functions*, Nova Science Publishers, New York, 1994
- [Tho95] THOMAS J. W.: *Numerical Partial Differential Equations - Finite Difference Methods*, Springer Texts in Applied Mathematics 22, Springer, New York, 1995
- [Tho85] THOMPSON J. F.: *Numerical Grid Generation*, North-Holland, Amsterdam, 1982
- [TWM85] THOMPSON J. F., WARSI Z. U. A., MASTIN C. W.: *Numerical Grid Generation, Foundations and Applications*, North-Holland, Amsterdam, 1985
- [Vor95] VAN DER VORST H. A.: *Bi-CGSTAB : A Fast And Smoothly Converging Variant of BI-CG for the solution of nonsymmetric linear Systems*, SIAM J. Sci. Stat. Com., Vol. 13, No. 2, pp. 631-644, 1992
- [WaWo95] WASILKOWSKI G. W., WOŹNIAKOWSKI H.: *Explicit cost bounds of algorithms for multivariate tensor product problems*, J. Complexity, Vol. 11, pp. 1-56, 1995
- [Wil] WILLSON R.: *Xmorphia*, Center for Advanced Computing Research, California Institute of Technology  
<http://www.cacr.caltech.edu/exhibit/ismap/image.html>  
siehe auch:  
<http://gauss.wias-berlin.de:800/support/Mosaic/Bsp-Xmorph.html>
- [XuSh92] XU J.-C., SHANN W.-C.: *Galerkin-wavelet methods for two-point boundary value problems*, Numer. Math., 63, pp. 123-144, 1992
- [Yse86] YSERENTANT H.: *On the multi-level splitting of finite element spaces*, Numer. Math., 49, pp. 379-412, 1986
- [Yse92] YSERENTANT H.: *Hierarchical bases*, in Proc. ICIAM91, Washington, R.E. O'Malley et al eds, SIAM, Philadelphia, 1992
- [Zen90] ZENGER CHR.: *Sparse Grids*, in Parallel Algorithms for Partial Differential Equations; Proceedings of the Sixth GAMM-Seminar, Kiel, 19.01.-21.01.1990, Notes on Numerical Fluid Mechanics Vol. 31, W. Hackbusch, ed., Vieweg, Braunschweig, 1991 (auch: Institut für Informatik der Technischen Universität München, SFB-Report 342/18/90 A, 1990)
- [Zen91] ZENGER CHR., HALLATSCHEK K.: *Fouriertransformation auf dünnen Gittern mit hierarchischen Basen*, Institut für Informatik der Technischen Universität München, SFB-Report 342/9/91 A, 1991

## **Lebenslauf**

Name : Thomas Schiekofer  
geboren am : 08. Februar 1969  
geboren in : München

## **Schulische Ausbildung**

1975 – 1979 Grundschule Höhenkirchen  
1979 – 1988 Gymnasium Ottobrunn  
1988 Abitur

## **Grundwehrdienst**

Oktober 1988 – Dezember 1989

## **Studium an der Technischen Universität München**

WS 1989/1990 Aufnahme des Studiums der Mathematik  
mit Nebenfach Informatik an der  
Technischen Universität München  
Herbst 1991 Vordiplom in Mathematik  
Herbst 1994 Hauptdiplom in Mathematik

## **berufliche Tätigkeit**

01.11.1991 – 28.02.1993 wissenschaftliche Hilfskraft am Institut für  
Angewandte Mathematik der Technischen  
Universität München  
01.04.1993 – 31.07.1993 wissenschaftliche Hilfskraft am Institut für  
Angewandte Mathematik der Technischen  
Universität München  
01.07.1995 – 31.05.1996 Wissenschaftlicher Angestellter am Institut  
für Informatik der Technischen  
Universität München  
seit 01.06.1996 Wissenschaftlicher Angestellter am Institut  
für Angewandte Mathematik der  
Friedrich-Wilhelms-Universität Bonn