

# Ein Raum-Zeit Dünngitterverfahren zur Diskretisierung parabolischer Differentialgleichungen

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch–Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von

Daniel Oeltz

aus

Bornheim-Sechtem

Bonn 2006

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

1. Referent: Prof. Dr. Michael Griebel

2. Referent: Prof. Dr. Rolf Krause

Tag der Promotion: 27. Juni 2006

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn  
[http://hss.ulb.uni-bonn.de/diss\\_online](http://hss.ulb.uni-bonn.de/diss_online) elektronisch publiziert.

## Zusammenfassung

In der vorliegenden Arbeit werden effiziente adaptive Diskretisierungsverfahren zur numerischen Lösung parabolischer Probleme vorgestellt. Hierbei gelingt es erstmalig, aufbauend auf speziellen diskreten Funktionenräumen, den sogenannten Raum-Zeit Dünngitterräumen, parabolische Probleme mit der gleichen Komplexität im Speicher- und Rechenaufwand wie stationäre elliptische Probleme zu lösen. Obwohl wesentlich weniger Freiheitsgrade als bei klassischen parabolischen Diskretisierungsverfahren benötigt werden, erreichen wir mit den vorgestellten Verfahren die (bis auf einen logarithmischen Faktor) gleichen Konvergenzraten wie bei herkömmlichen Diskretisierungen. Hierzu werden lediglich etwas stärkere Glattheitsvoraussetzungen an die Lösung des parabolischen Problems benötigt. Es wird jedoch in dieser Arbeit gezeigt, dass diese Glattheitsvoraussetzungen bei geeigneten Annahmen an das Gebiet, die rechte Seite und die Anfangs- und Randbedingungen für die Lösung parabolischer Probleme erfüllt sind.

Ferner stellen wir für den Fall, dass die zu approximierende Funktion nicht genügend glatt ist, eine adaptive Erweiterung des Verfahrens in Raum und Zeit vor. Die resultierenden adaptiven Diskretisierungen weisen in den numerischen Experimenten für Probleme mit nicht glatten Lösungen nahezu die gleiche Effizienz wie die nicht adaptiven Diskretisierungsverfahren für Probleme mit genügend glatten Lösungen auf. Besonders bemerkenswert ist hierbei, dass das vorgestellte adaptive Verfahren automatisch zu lokalen Zeitschritten (*local time stepping*) führt, deren Umsetzung bei herkömmlichen Diskretisierungen algorithmisch aufwändig ist.

Zur effizienten Lösung der bei der Diskretisierung anfallenden linearen Gleichungssysteme werden in dieser Arbeit Multilevellöser in Raum-Zeit entwickelt. Wir untersuchen die Konvergenzeigenschaften der Löser an numerischen Beispielen, die zeigen, dass die Konvergenzraten von der Feinheit der Diskretisierung unabhängig sind.

Zum Abschluss verwenden wir die Raum-Zeit Dünngitterdiskretisierungen zur numerischen Lösung der zu instationären verteilten Kontrollprobleme gehörenden Sattelpunktsprobleme. Während bisherige Arbeiten zur Diskretisierung dieser Sattelpunktsprobleme auf Grund der hohen Zahl an Freiheitsgraden klassischer Diskretisierungsverfahren hierbei lediglich zwei Ortsdimensionen behandeln, sind wir mit den Raum-Zeit Dünngitterdiskretisierungen in der Lage, erstmals auch Probleme in drei Ortsdimensionen zu behandeln. Hierzu erweitern wir die Multilevellöser und die Adaptivität auf die Lösung von Systemen parabolischer Differentialgleichungen. Unterschiedliche numerische Beispiele demonstrieren dabei die Effizienz der adaptiven Raum-Zeit Dünngitterdiskretisierung zur Lösung der Sattelpunktsprobleme in bis zu drei Ortsdimensionen.

## Danksagung

An dieser Stelle möchte ich es nicht versäumen, den Personen Dank auszusprechen, die auf die eine odere andere Art dazu beigetragen haben, dass es möglich war, diese Arbeit zu beginnen, durchzuhalten und zu einem erfolgreichen Abschluss zu bringen. Zunächst gebührt an dieser Stelle mein Dank Prof. Michael Griebel, der mich in dieses Thema eingeführt und mir mit zahlreichen Ideen und Ratschlägen weiter geholfen hat. Ebenso möchte ich mich bei Prof. Rolf Krause für die Übernahme des Zweitgutachtens und für die Zeit, die er stets für Diskussionen mit mir übrig hatte, bedanken. Dipl. Math. Jürgen Braun und Nino Meurer gilt mein Dank für das Korrekturlesen. Einen ganz besonderen Dank gilt Herrn Dr. Marc Alexander Schweitzer dafür, dass er immer ein offenes Ohr und viele Denkanstöße für mich hatte und mir bei jeglichen Computerproblemen zur Seite stand. Als weitere wichtige Person, ohne die ich nur halb so viel Freude bei der Arbeit gehabt hätte, sei meinem Kollegen Herrn Dipl. Phys. Lukas Jager Dank gesagt, der viele meiner Launen mit getragen hat und mit dem ich jede Idee ausführlich besprechen konnte. Schließlich möchte ich allen Kollegen und Mitarbeitern des Instituts für Numerische Simulation der Universität Bonn für die angenehme und inspirierende Arbeitsatmosphäre danken.

Bonn, im März 2006

Daniel Oeltz

# Kapitel 1

## Einleitung

Die numerische Simulation hat in den letzten Jahren eine immer bedeutendere Rolle in der Forschung und Entwicklung eingenommen. Kostenintensive oder in der Realität gefährliche Experimente werden heute zunehmend durch Simulationen auf dem Computer ersetzt. Zusätzlich können Prozesse simuliert werden, deren Beobachtung nur sehr schwer oder sogar unmöglich wäre, da die zu beobachtenden Phänomene auf zu kleiner oder zu großer Größen- oder Zeitskala stattfinden. Als Anwendungsgebiete lassen sich hier beispielsweise die Simulation von Crash-Tests, der elektronische Windkanal, die Optionspreisbewertung oder die Wettervorhersage anführen. Für die Simulation solcher Prozesse wird zunächst ein mathematisches Modell benötigt, das alle Größen, an deren Beobachtung Interesse besteht, möglichst genau reproduziert. Bei diesen mathematischen Modellen handelt es sich um Näherungen an die Wirklichkeit, wobei nicht jedes Modell unumstritten ist.

Häufig werden solche Modelle durch partielle Differentialgleichungen beschrieben. Aufgrund der den realen Problemen zugrunde liegenden Komplexität sind diese partiellen Differentialgleichungen in aller Regel jedoch nicht analytisch lösbar. Aus diesem Grund versucht man mit Hilfe des Computers lediglich eine Näherungslösung zu berechnen. Dabei hat es nicht nur die außerordentliche Leistungsentwicklung der Computer ermöglicht, solche komplexen Prozesse auf heutigen Rechnern zu lösen. Es ist auch vor allem der Entwicklung effizienter Algorithmen für die entsprechenden Problemstellungen zu verdanken, dass gegenwärtig immer mehr Probleme schnell durch die Anwendung numerischer Verfahren auf dem Computer mit der gewünschten Genauigkeit gelöst werden können.

Für den Gesamt-Rechenaufwand zur Berechnung einer Näherungslösung sind einerseits die Anzahl der benötigten Freiheitsgrade und andererseits die pro Freiheitsgrad benötigten Rechenoperationen von ausschlaggebender Bedeutung. Daher ist die Minimierung der Anzahl benötigter Freiheitsgrade und des Rechenaufwands pro Freiheitsgrad, d.h. die Entwicklung effizienter Algorithmen, das zweite große

Forschungsgebiet des wissenschaftlichen Rechnens und der numerischen Simulation.

In dem Kontext dieses zweiten Zweiges des wissenschaftlichen Rechnens und der numerischen Simulation ist es der Anspruch der vorliegenden Arbeit einen Beitrag zur Bewältigung der angesprochenen Schwierigkeiten zu leisten. Hierbei betrachten wir die numerische Lösung parabolischer Gleichungen. Dieser Typ partieller Differentialgleichungen resultiert aus der Modellierung verschiedenster zeitabhängiger Phänomene in den Natur-, Ingenieurs- und Wirtschaftswissenschaften und beschreibt die Entwicklung bestimmter Größen im Ort, z.B. der Temperatur, unter vorgegebenen äußeren Einflüssen in Abhängigkeit von der Zeit.

Zur Lösung parabolischer Gleichungen existieren bereits Ansätze. Dabei wird im Wesentlichen zwischen drei Verfahrenstypen unterschieden:

- In der ersten Methode, der so genannten Linienmethode (method of lines, MOL), wird zunächst der Ort mittels Finiter Differenzen oder Finiter Elemente diskretisiert was das Ursprungsproblem in ein System gewöhnlicher Differentialgleichungen transformiert, vgl. [70, 106]. Dieses System kann nun mit herkömmlichen Integrationsschemata gelöst werden. Die Genauigkeit im Ort kann dabei durch klassische a posteriori Fehlerschätzer für stationäre Probleme kontrolliert werden. Allerdings ist es technisch schwierig, Veränderungen der Ortsdiskretisierung in der Zeit zu behandeln.
- Verglichen mit der Linienmethode wird bei der Rothe Methode genau entgegengesetzt vorgegangen und zunächst die Zeit diskretisiert [70]. Dies führt zu einer Sequenz von elliptischen Problemen, die durch Standardtechniken adaptiv gelöst werden können. Hierbei ist nun eine adaptive Behandlung der Zeit technisch aufwändig.
- Bei der Discontinuous-Galerkin Diskretisierung [49, 50, 51, 52, 53] werden schließlich Raum und Zeit zugleich diskretisiert wodurch die Theorie im Hinblick auf die Fehlerschätzung und Adaptivität vereinfacht wird. In der Praxis zerfällt diese Methode allerdings, wie bei den beiden vorherigen Ansätzen, zu einem Zeitschrittverfahren, d.h. dass eine Sequenz von Ortsproblemen gelöst werden muß. Damit ist die gleichzeitige Adaptivität bzgl. des Ortes und der Zeit in der praktischen Umsetzung auf dem Computer technisch schwierig und aufwändig.

Die Diskretisierung mit Hilfe dieser drei Verfahren führt letztendlich immer zu einer Sequenz von diskreten Problemen, deren Lösungen die Lösung des Ursprungsproblems an den verschiedenen diskreten Zeitpunkten approximieren. Wird etwa ein geeignetes Diskretisierungsschema auf einem uniformen Gitter mit einem Fehler der Ordnung  $p$  im Ort mit  $N^d$  Freiheitsgraden und einem Fehler der Ordnung  $q$  in der Zeit angewendet, so benötigen wir  $O(N^{d+p/q})$  Freiheitsgrade in Raum und Zeit um insgesamt eine Fehlerordnung von  $p$  zu erhalten. So resultiert beispielsweise für das

implizite Eulerverfahren mit  $q = 1$  und bei entsprechender Ortsdiskretisierung mittels Finiter Differenzen oder linearer Finiter Elemente eine Fehlerordnung  $p = 2$  und ein Gesamtaufwand  $O(N^{d+2})$ , während wir für das Crank-Nicolson Verfahren, für das  $q = 2$  gilt, einen Gesamtaufwand von  $O(N^{d+1})$  feststellen können. Die Lösung besitzt also eine Genauigkeit von  $O(N^{-2})$  wenn wir  $O(N^{d+1})$  bzw.  $O(N^{d+2})$  Freiheitsgrade verwenden. Selbst wenn die bei den Verfahren anfallenden Ortsprobleme mit  $O(N^d)$  Unbekannten noch relativ schnell gelöst werden können, führen die  $O(N)$  oder  $O(N^2)$  Zeitschritte dazu, dass auch auf modernsten Computern sehr lange gerechnet werden muß, bis eine Lösung mit zufrieden stellender Genauigkeit approximiert ist.

Aus diesem Grund ist die Untersuchung und Entwicklung von numerischen Verfahren zur Lösung von parabolischen Gleichungen Gegenstand aktueller Forschung. Hierbei wird von einer für die Lösung von parabolischen Problemen typischen Eigenschaft Gebrauch gemacht: Die im Ort hochfrequenten Lösungsanteile entwickeln sich auf einer größeren Zeitskala als die niederfrequenten Anteile. Auf dieser Beobachtung basierend sind bereits Ansätze zur schnellen Lösung parabolischer Probleme entwickelt worden.

In [27, 28] wird mit der so genannten *frozen  $\tau$*  Technik vorgeschlagen, die Lösung zu verschiedenen Zeitpunkten auf unterschiedlich feinen Gittern zu berechnen. Bei diesem Ansatz wird die Steuerung, welches Gitter zu welchem Zeitschritt zur Berechnung verwendet werden soll, während der Berechnung eines Zeitschrittes über die Größe von Mehrgitterkorrekturen heuristisch bestimmt. Dieses Vorgehen macht eine a priori Fehleranalyse sehr schwer und führt zusätzlich dazu, dass unter Umständen an manchen Zeitschritten auf zu feinen Gittern gerechnet wird, was wiederum die Komplexität verschlechtert.

Ein der *frozen  $\tau$*  Technik ähnliches Verfahren wird in [36, 37] für parabolische Probleme mit periodischen Randbedingungen auf dem Einheitswürfel diskutiert. Hierbei wird der Ort allerdings nicht mit Hilfe einer Gitterhierarchie zerlegt, sondern mit einer Fourierbasis diskretisiert, wobei dann die unterschiedlichen Ortsfrequenzen mit verschiedenen Zeitschritten integriert werden. Durch die Verwendung von Fourierbasen ist dieser Ansatz allerdings auf Probleme mit periodischen Randbedingungen auf dem Einheitswürfel beschränkt, wobei der elliptische Teil des parabolischen Operators durch den Laplace-Operator beschrieben wird.

Die Ergebnisse dieser Arbeiten begründen die Hoffnung, dass wesentlich effizientere Verfahren zur Lösung parabolischer Probleme existieren, als die oben beschriebenen klassischen Zeitschrittverfahren. Insbesondere drängt sich die Frage auf, ob und wie es möglich ist, die durch die Zeit entstandene, um den Faktor  $O(N)$  bzw.  $O(N^2)$  verglichen mit rein stationären Problemen erhöhte Komplexität zu reduzieren.

Betrachten wir die Zeit als zusätzliche Dimension, so sehen wir, dass die Dimension des parabolischen Problems im Vergleich zu einem rein stationären Problem

um eins größer ist. Der Effekt, dass die Zahl der Freiheitsgrade exponentiell mit wachsender Dimension steigt, während die Approximationsrate konstant bleibt, ist als "*Fluch der Dimension*" wohlbekannt [15]. Der Faktor  $O(N)$  bzw.  $O(N^2)$  um den sich die Komplexität der Zeitschrittverfahren verglichen mit Verfahren zur Lösung zeitunabhängiger Probleme verschlechtert, kann demnach einfach als ein Beispiel für den "*Fluch der Dimension*" aufgefasst werden. Dieser "*Fluch der Dimension*" kann allerdings stark verringert werden, wenn man das Problem auf die Approximation von Funktionen aus bestimmten Funktionenklassen einschränkt. Hierbei haben die so genannten Dünngitter in den letzten Jahrzehnten eine immer stärkere Bedeutung erlangt.

Dünngitter wurden bereits erfolgreich bei einer Reihe von Problemen angewendet, wie etwa in der Quantenmechanik [57, 73, 125], zur Lösung von stochastischen Differentialgleichungen [101], bei der Quadratur von hochdimensionalen Funktionen in der Physik und Finanzmathematik [10, 18, 58, 90] und zur Lösung von elliptischen partiellen Gleichungen [5, 6, 29]. Unter zusätzlichen Glattheitsannahmen approximieren Dünngitter Funktionen mit (fast) gleicher Rate wie volle Gitter, obwohl in  $d$ -Dimensionen Dünngitter mit  $O(N \log(N)^{d-1})$  Freiheitsgraden wesentlich weniger Freiheitsgrade benötigen als volle Gitter, die  $O(N^d)$  Unbekannte haben. Tatsächlich wurden in [94, 99] parabolische Probleme mit Hilfe dünner Gitter (in Raum und Zeit) und der Kombitechnik gelöst, während in [5] eine Galerkin-artige Dünngitterdiskretisierung parabolischer Probleme mit Hilfe des unidirektionalen Prinzips vorgestellt wurde. In [91] werden schließlich innerhalb einer  $p$ -Version zur Zeitdiskretisierung dünne Gitter zur Lösung der anfallenden stationären Probleme verwendet.

Obwohl die numerischen Experimente dieser Arbeiten viel versprechend sind, bleiben dennoch die für Dünngitter typischen Probleme erhalten. So ist die Behandlung von beliebigen variablen Koeffizienten im Differentialoperator sehr schwierig und die Regularitätsannahmen an die Lösung des Problems im Fall von Dünngittern sind ausgesprochen restriktiv. Eine weitere Problematik stellt die Behandlung von beliebigen Geometrien dar [1, 46, 47, 98]. Diese Probleme treten im Fall großer Dimensionen, d.h.  $d > 3$ , in den Hintergrund, da in diesem Fall nahezu alle auftretenden Gebiete leicht auf den Einheitswürfel transformiert werden können. Aus diesem Grund scheint die Dünngitterdiskretisierung für hochdimensionale Probleme ( $d > 3$ ) durchaus Erfolg versprechend [56]. Für den Fall  $d \leq 3$  konnten sich Dünngitter aus den genannten Gründen in der Praxis allerdings nicht durchsetzen.

Wir stellen in dieser Arbeit als Verallgemeinerung der klassischen Dünngitter die *Raum-Zeit Dünngitter* vor. Ist das dem parabolischen Problem zugrunde liegende Ortsgebiet  $\Omega$  zeitunabhängig, so besitzt der zugehörige Raum-Zeit Zylinder  $\Omega \times (0, T)$  offensichtlich eine triviale Tensorproduktstruktur, die wir bei den Raum-Zeit Dünngittern ausnutzen werden. Hierbei konstruieren wir die Dünngitterfunktionen als Tensorprodukt von einer eindimensionalen Multilevelbasis in der Zeit und ei-

---

ner  $d$ -dimensionalen Multilevelbasis im Ort. Es wird sich erweisen, dass die oben erwähnten Probleme klassischer Dünngitter mit diesem Prinzip umgangen werden können. Zusätzlich werden wir zeigen, dass damit Raum-Zeit Diskretisierungen möglich sind, die anstelle der  $O(N^{d+1})$  oder  $O(N^{d+2})$  Freiheitsgrade einer klassischen parabolischen Diskretisierung mit beispielsweise dem Crank-Nicolson oder dem impliziten Euler Verfahren lediglich  $O(N^d)$  Freiheitsgrade benötigen, wobei unter geeigneten Glattheitsvoraussetzungen die Konvergenzordnung (nahezu) unverändert erhalten bleibt. Des Weiteren wird in dieser Arbeit ein Multilevelverfahren in Raum-Zeit entwickelt, mit dem die bei der Raum-Zeit Dünngitterdiskretisierung entstehenden linearen Gleichungssystem mit einem zur Zahl der Unbekannten proportionalen Aufwand gelöst werden können. Die mit diesem Multilevelverfahren erzielten Konvergenzraten sind dabei unabhängig von der Feinheit der Diskretisierung. Somit führt die Kombination aus dem Diskretisierungsverfahren mit dem effizienten Löser zu einem Gesamtverfahren, das bei genügender Glattheit eine numerische Lösung des parabolischen Problems mit dem gleichem Aufwand wie die Lösung eines stationären Problems ermöglicht. Für den Fall, dass die Lösung des parabolischen Problems die benötigten Glattheitsvoraussetzungen nicht erfüllt, stellen wir eine adaptive Erweiterung der vorgestellten Diskretisierungsmethoden vor. Es wird anhand einiger numerischer Beispiele gezeigt werden, dass der Einsatz der vorgestellten Adaptivität in Raum-Zeit die Lösung von Problemen mit singulären Lösungen mit einem vergleichbaren Aufwand wie die Lösung von Problemen mit glatten Lösungen erlaubt. Schließlich wird auch die Erweiterung der Raum-Zeit Dünngitterdiskretisierungen auf instationäre verteilte Kontrollprobleme untersucht werden. Die numerische Lösung der zu dem Kontrollproblem gehörigen Sattelpunktsprobleme ist aufgrund der hohen Komplexität herkömmlicher Verfahren bisher lediglich für die Untersuchung von zweidimensionalen Problemen eingesetzt worden. Mit Hilfe der Raum-Zeit Dünngitterdiskretisierungen sind wir hingegen erstmalig in der Lage, auch solche Sattelpunktsprobleme in drei Ortsdimensionen adaptiv zu lösen und geben einige numerische Ergebnisse in bis zu drei Ortsdimensionen an.

Wir gehen bei unserer Untersuchung der Raum-Zeit Dünngitterdiskretisierungen in der vorliegenden Arbeit wie folgt vor. In Kapitel 2 führen wir zunächst die Raum-Zeit Dünngitter ein. Wir untersuchen dabei deren Eigenschaften bezüglich der Dimension und der Approximationsrate in bestimmten Sobolev-Räumen. Dabei betrachten wir zunächst den Fall, dass die in der Konstruktion verwendeten Multilevelsplittings gewisse Normäquivalenzen erfüllen (wie sie etwa für Wavelets gelten). Da die Konstruktion solcher Splittings auf allgemeinen Geometrien allerdings ausgesprochen schwierig ist, untersuchen wir auch den Fall der Verwendung der hierarchischen Basis innerhalb der Dünngitterkonstruktion, der auch in den weiteren Teilen dieser Arbeit betrachtet wird.

In Kapitel 3 werden wir dann klassische Regularitätsresultate für lineare parabolische Differentialgleichungen betrachten. Dabei werden wir zeigen, dass die

Lösungen von parabolischen Problemen für die Approximation in den Dünngitterräumen wichtige Glattheitsannahmen erfüllen, sofern gewisse Bedingungen an das zugrunde liegende Gebiet, den elliptischen Teil des parabolischen Operators und die Anfangs- und Randbedingungen erfüllt sind.

In Kapitel 4 diskutieren wir auf den Raum-Zeit Dünngittern basierende Diskretisierungsverfahren. Wir untersuchen eine Raum-Zeit Dünngittervariante des Crank-Nicolson und des Discontinuous-Galerkin Verfahrens. Für das Discontinuous-Galerkin Verfahren mit stückweise linearen oder konstanten Funktionen in der Zeit werden wir zusätzlich a priori Fehlerabschätzungen angeben, die den Fehler nach oben gegen Interpolationsfehler des Raum-Zeit Dünngitters abschätzen. Schließlich führen wir noch das Unidirektionale Prinzip ein, mit dem das Matrix-Vektor Produkt des mit den Raum-Zeit Dünngittern diskretisierten Systems effizient, d.h. mit einer zu der Zahl der Unbekannten proportionalen Anzahl an Rechenschritten, berechnet werden kann, obwohl die zugrunde liegende Matrix nicht dünn besetzt ist. Dazu verallgemeinern wir das bereits in [5] vorgestellte Prinzip auf beliebige Tensorprodukt-Bilinearformen und stellen auch neue, flexiblere Varianten der in dem Algorithmus benötigten Teilalgorithmen Bottom Up und Top Down vor. Schließlich werden wir anhand einiger numerischer Experimente die Konvergenz der vorgestellten Diskretisierungsverfahren untersuchen.

Da nicht nur die Anzahl der zum Erreichen einer vorgegebenen Genauigkeit benötigten Freiheitsgrade, sondern auch die pro Freiheitsgrad benötigte Anzahl an Rechenschritten zur Lösung des Problems für ein effizientes Verfahren von Bedeutung sind, werden wir in Kapitel 5 Verfahren zur Lösung der anfallenden linearen Gleichungssysteme betrachten. In Anlehnung an [61] werden wir einen neuen Multilevel-Löser für die entstehenden Gleichungssysteme als (Block-) Gauss-Seidel Verfahren über ein mittels des Erzeugendensystems assemblierten singulären Systems entwickeln. Ferner untersuchen wir die Konvergenzeigenschaften des Verfahrens an verschiedenen numerischen Beispielen. Die Experimente zeigen, dass das Verfahren mit einer von der Feinheit der Diskretisierung unabhängigen Rate konvergiert.

Um Probleme, deren Lösungen nicht die geforderten Glattheitseigenschaften erfüllen, effizient behandeln zu können, diskutieren wir in Kapitel 6 Adaptivität im Kontext der Raum-Zeit Dünngitterdiskretisierungen. Hierzu übertragen wir einen auf den hierarchischen Koeffizienten der diskreten Lösung basierenden Fehlerindikator für klassische Dünngitter [5, 32, 62] auf die vorgestellten Raum-Zeit Dünngitter und untersuchen die Effizienz dieses Fehlerindikators anhand einiger numerischer Beispiele. In den letzten Jahren hat die Fehlerschätzung bezüglich linearer Funktionale in dem Kontext von adaptiven Finiten Element Diskretisierungen eine immer stärkere Bedeutung erlangt [8, 13]. Damit wird vor allem dem Ziel Rechnung getragen, nicht unbedingt die Lösung selbst möglichst exakt zu bestimmen, sondern eventuell nur gewisse gemittelte Größen der Lösung zu berechnen. Hierzu zeigen wir

eine Möglichkeit der Konstruktion eines Fehlerindikators für lineare Funktionale im Kontext der Raum-Zeit Dünngitterdiskretisierung auf. Zudem diskutieren wir die Effizienz dieses Fehlerindikators an einigen numerischen Ergebnissen.

In Kapitel 7 beschäftigen wir uns mit der Frage, wie die zunächst für den skalaren Fall vorgestellten Diskretisierungsverfahren auf die Diskretisierung von Systemen partieller Differentialgleichungen übertragen werden können. Als besonders interessante Anwendung werden wir dabei unsere Untersuchung auf instationäre verteilte Kontrollprobleme fokussieren, zu deren Lösung gekoppelte Systeme von Vorwärts- und Rückwärtsgleichungen gelöst werden müssen. Wir werden zeigen, dass sich die in den vorherigen Kapiteln untersuchten Algorithmen auf diese Problemklasse übertragen lassen und die gleichen Konvergenzraten des Löser und der Diskretisierung wie in den skalaren Beispielen erreicht werden. Man beachte, dass Raum-Zeit Diskretisierungen dieser Systeme auf vollen Gittern wegen des enormen Speicherplatzbedarfes der Vollgitterräume bisher nur für maximal zwei Ortsdimensionen durchgeführt wurden [22, 23, 24], während mit Raum-Zeit Dünngittern auf Grund der wesentlich geringeren Komplexität, nun erstmals auch Ergebnisse in drei Ortsdimensionen diskutiert werden können.

In Kapitel 8 fassen wir schließlich die Ergebnisse dieser Arbeit kurz zusammen und geben einen Ausblick auf offene Fragestellungen im Kontext der Raum-Zeit Dünngitterdiskretisierungen.



## Kapitel 2

# Raum-Zeit Dünngitter

In diesem Kapitel beschreiben wir die Konstruktion der Raum-Zeit Dünngitterräume und untersuchen einige Eigenschaften dieser Räume etwas genauer. Hierzu führen wir zunächst die benötigte Notation ein. Dann stellen wir die Raum-Zeit Dünngitterkonstruktion vor und untersuchen die Dimension der resultierenden Räume. Im Kontext von [79] diskutieren wir außerdem die Approximationseigenschaft dieser Räume unter der Voraussetzung, dass gewisse Normäquivalenzen für die in der Konstruktion verwendeten Multilevelbasen im Ort und in der Zeit gelten. Da die Konstruktion von Basen, die diese Normäquivalenzen erfüllen, in der Praxis auf beliebigen Geometrien sehr schwierig und aufwändig ist (oder sogar unmöglich), werden wir auch ein Resultat vorstellen, bei dem lediglich obere Abschätzungen für die entsprechenden Multilevelbasen gelten müssen. Wir stellen zudem die eindimensionalen Multilevelbasen vor, die wir in dieser Arbeit innerhalb der Konstruktion der Raum-Zeit Dünngitter für die Zeitkomponente verwenden. Schließlich diskutieren wir die Approximationseigenschaften der Raum-Zeit Dünngitterräume bei Verwendung dieser eindimensionalen Multilevelbasen in der Zeit und der linearen bzw.  $d$ -linearen hierarchischen Basis im Ort, die aus einer Folge von ineinander geschachtelten Finite Element Räumen gewonnen wird.

### 2.1 Notation

In diesem Abschnitt führen wir einige Funktionenräume und die dazugehörige Notation ein, die wir für die theoretische Betrachtung der Raum-Zeit Dünngitterräume benötigen. Diese Räume sind eine Verallgemeinerung der in der Theorie über die klassischen Dünngitterräume vorkommenden Räume dominierender gemischter Glattheit, vergleiche [35, 104, 105].

Mit fett gedruckten Buchstaben bezeichnen wir im Folgenden Multiindizes, d.h.  $\mathbf{k}$  bezeichnet einen Vektor mit Komponenten  $k_i$ ,  $1 \leq i \leq n$ , wobei  $n$  aus dem Kontext

hervorgeht. Wir verwenden die Normen

$$\|\mathbf{k}\|_\infty := \max_{1 \leq i \leq n} |k_i|, \quad \|\mathbf{k}\|_1 := \sum_{i=1}^d |k_i|,$$

und die Abkürzung

$$\partial^{\mathbf{k}} f := \partial_{x_1}^{k_1} \dots \partial_{x_d}^{k_d} f.$$

Sei  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 1$ , ein beschränktes Gebiet. Wie üblich definieren wir für festes  $m \in \mathbb{N}$  und  $1 \leq p < \infty$  die Norm

$$\|u\|_{m,p} := \left( \sum_{\|\mathbf{k}\|_1 \leq m} \int_{\Omega} |\partial^{\mathbf{k}} u|^p dx \right)^{1/p} \quad (2.1)$$

und die Seminorm

$$|u|_{m,p} := \left( \sum_{\|\mathbf{k}\|_1 = m} \int_{\Omega} |\partial^{\mathbf{k}} u|^p dx \right)^{1/p}, \quad (2.2)$$

wobei die so genannte schwache Ableitung  $\partial^{\mathbf{k}} u$  von  $u$ , sofern sie existiert, definiert ist durch

$$\int_{\Omega} \partial^{\mathbf{k}} u \varphi dx = (-1)^{\|\mathbf{k}\|_1} \int_{\Omega} u \partial^{\mathbf{k}} \varphi dx \quad \forall \varphi \in C_0^\infty(\Omega). \quad (2.3)$$

Für  $p = \infty$  erhalten wir analog

$$\|u\|_{m,\infty} := \sum_{\|\mathbf{k}\|_1 \leq m} \|\partial^{\mathbf{k}} u\|_\infty. \quad (2.4)$$

Für  $m \in \mathbb{N}$ ,  $1 \leq p \leq \infty$  bezeichnen wir mit  $H^{m,p}(\Omega)$  wie allgemein üblich den Sobolev-Raum

$$H^{m,p}(\Omega) := \{u \in L^p(\Omega) \mid \text{es existiert } \partial^{\mathbf{k}} u \text{ und } \partial^{\mathbf{k}} u \in L^p(\Omega) \forall \|\mathbf{k}\|_1 \leq m\}.$$

Für den am häufigsten vorkommenden Fall  $p = 2$  kürzen wir die Schreibweisen durch Weglassen von  $p$  ab, d.h. wir definieren  $\|\cdot\|_m := \|\cdot\|_{m,2}$  und  $H^m(\Omega) := H^{m,2}(\Omega)$ .

Sei  $T > 0$  fest und  $\Omega_T := \Omega \times (0, T)$ . Die Regularitätstheorie für die Lösung parabolischer Probleme zeigt, dass die Lösung bzgl. der Ortsvariablen und der Zeit durchaus unterschiedlich glatt sein kann. Wir betrachten daher im Folgenden zusätzlich den Raum

$$H^{2m,m}(\Omega_T) := \{u \in L^2(\Omega_T) \mid \partial^{\mathbf{k}} u \in L^2(\Omega_T) \forall \|\mathbf{k}\|_1 + 2k_{d+1} \leq 2m\}, \quad (2.5)$$

dessen Funktionen im Ort eine doppelt so hohe Glattheit wie in der Zeit besitzen. Wir definieren nun die Räume dominierender gemischter Glattheit mittels Tensorproduktbildung.

**Definition 2.1.** Für  $m, k \in \mathbb{N}$  und  $1 \leq p, q \leq \infty$  definieren wir

$$H_{mix}^{(m,p),(k,q)}(\Omega_T) := H^{m,p}(\Omega) \otimes H^{k,q}((0, T)) \quad (2.6)$$

mit der Norm

$$\|u\|_{H_{mix}^{(m,p),(k,q)}(\Omega_T)} := \sum_{\substack{\|\mathbf{i}\|_1 \leq m \\ j \leq k}} \|\partial_t^j \partial_x^{\mathbf{i}} u\|_{L^p(\Omega) \otimes L^q((0, T))}. \quad (2.7)$$

Falls  $p = q = 2$  gilt, schreiben wir auch kürzer  $H_{mix}^{m,k}(\Omega_T)$  und  $\|\cdot\|_{H_{mix}^{m,k}(\Omega_T)}$  und für  $m = k$  kürzen wir weiter durch  $H_{mix}^m(\Omega_T)$  und  $\|\cdot\|_{H_{mix}^m(\Omega_T)}$  ab.

Für eine etwas ausführlichere Diskussion des Tensorproduktes verweisen wir auf [97, 116]. Wir merken an dieser Stelle nur an, dass wir für zwei Funktionen  $u^\Omega \in H^{m,p}(\Omega)$  und  $u^T \in H^{k,q}((0, T))$  das Tensorprodukt der beiden Funktionen mit dem Produkt dieser beiden Funktionen identifizieren,

$$u^\Omega \otimes u^T = u^\Omega \cdot u^T.$$

Zusätzlich werden wir später ausnutzen, dass im Fall  $p = q = 2$  für ein Orthonormalsystem  $\{e_i^\Omega\}$  von  $H^m(\Omega)$  und  $\{e_j^T\}$  von  $H^k((0, T))$  die Produkte  $e_i^\Omega \cdot e_j^T$  wieder ein Orthonormalsystem von  $H_{mix}^{m,k}(\Omega_T)$  bilden und insbesondere  $H_{mix}^{m,k}(\Omega_T)$  wieder ein Hilbertraum ist.

Man beachte, dass wir bisher lediglich die Produktstruktur des Raum-Zeit Zylinders ausgenutzt haben. Damit haben wir uns lediglich auf in der Zeit konstante Gebiete  $\Omega$  eingeschränkt und benötigen keinerlei Produktstruktur des Ortsgebietes  $\Omega$ . Aus diesem Grund können wir allgemeinere Probleme als im Fall klassischer Dünngitter [79] behandeln, deren Konstruktion nur auf Produktgebieten durchführbar ist. Besitzt hingegen  $\Omega$  Produktstruktur, d.h.  $\Omega = I_1 \times \dots \times I_d$  mit  $I_i \subset \mathbb{R}$ , gilt

$$H^m(I_1) \otimes \dots \otimes H^m(I_d) \otimes H^k((0, T)) \subset H_{mix}^{m,k}(\Omega_T)$$

und die oben definierten Räume sind echt größer als die klassischen Räume dominierender gemischter Ableitungen, die in der Analyse klassischer Dünngitter eine zentrale Rolle spielen.

## 2.2 Konstruktion und Eigenschaften der Raum-Zeit Dünngitter

In diesem Abschnitt führen wir die Raum-Zeit Dünngitter ein, die als Verallgemeinerung der klassischen Dünngitterräume aufgefasst werden können. Klassische Dünngitterräume werden von Tensorprodukten eindimensionaler Multiskalenbasen unterschiedlicher Level aufgespannt, d.h. die Basisfunktionen des Dünngitterraumes sind im allgemeinen stark anisotrop. Hiermit ist es möglich, dass diese Räume für

Funktionen, die genügend glatt sind, mit  $N$  Freiheitsgraden eine Approximationsgenauigkeit von  $O(N^{-\alpha})$  erreichen, wobei  $\alpha > 0$  unabhängig von der Dimension ist. Dies ist der große Vorteil gegenüber anderen klassischen Ansatzräumen wie zum Beispiel Finiten Elementen, deren Approximationsordnung sich exponentiell mit der Dimension verschlechtert.

Dementsprechend sind in den letzten Jahren die Dünngitterräume insbesondere für Probleme in hohen Dimensionen untersucht und angewendet worden, wie zum Beispiel in der Quantenmechanik [73, 57, 125], bei stochastischen Differentialgleichungen [101], bei hochdimensionaler Quadratur in der Physik und Finanzmathematik [10, 18, 58, 90] und zur Lösung von hauptsächlich elliptischen Differentialgleichungen [5, 6, 29]. Insbesondere sei an dieser Stelle auf den Übersichtsartikel [35] und die Referenzen dort verwiesen.

Dabei beruht das Konstruktionsprinzip der dünnen Gitter wesentlich auf der Tensorproduktbildung eindimensionaler Funktionen. Diese Tensorproduktbildung führt allerdings zu Schwierigkeiten.

So ist der Algorithmus, das so genannte Unidirektionale Prinzip, zur effizienten Berechnung des Matrix-Vektor Produktes bei der Galerkin-Diskretisierung mit Dünngittern für elliptische Probleme zunächst nur im Fall von Koeffizientenfunktionen anwendbar, die Tensorproduktstruktur besitzen. In [46, 93] werden entsprechende Modifikationen vorgeschlagen, so dass der Algorithmus auch auf allgemeinere Fälle angewendet werden kann. Allerdings wird dazu die zugrunde liegende Bilinearform derart modifiziert, dass unklar ist, mit welcher Rate die diskrete Lösung des derartig diskretisierten Problems gegen die exakte Lösung konvergiert.

Eine weitere Schwierigkeit besteht in der Behandlung allgemeiner Gebiete, da die Tensorproduktkonstruktion zunächst nur auf Gebieten mit Produktstruktur durchgeführt werden kann. Für Gebiete, die sich als Vereinigung von  $d$ -dimensionalen Quadern darstellen lassen, können herkömmliche Gebietszerlegungsmethoden in Kombination mit der Dünngittertechnik, wie in [98] beschrieben, angewendet werden. Für krummlinig berandete Gebiete wurde in [33, 46, 47, 98] vorgeschlagen, die entsprechenden Gebiete mit Hilfe geeigneter Abbildungen auf den Einheitswürfel zu transformieren. Schließlich wird in [1] eine Art Finite Element Ansatz betrachtet, bei dem das zugrunde liegende Gebiet durch ein Gitter aufgelöst wird, wobei in jedem Quader des Gitters lokal ein Dünngitterraum aufgesetzt wird. Bei diesen Lösungsansätzen treten allerdings weitere Schwierigkeiten auf. Während in den letzten Jahren die Entwicklung von Gittergeneratoren vorangetrieben wurde und dementsprechend groß die Auswahl von verfügbaren Gittergeneratoren ist, ist kein Programmpaket verfügbar, das die benötigten Gebietszerlegungen mit den entsprechenden Transformationen auf den Einheitswürfel berechnet. Des Weiteren führt eine entsprechende Gebietstransformation in der Regel zu variablen Koeffizientenfunktionen, deren Behandlung im Dünngitterfall, wie oben diskutiert, nicht ohne weiteres mit der gleichen Effizienz möglich ist, wie im Fall von Tensorprodukt-Ko-

effizientenfunktionen.

Wie wir sehen, beruhen die beschriebenen Probleme hauptsächlich auf der Tensorproduktkonstruktion der Dünngitterräume. Die Idee für zeitabhängige Probleme besteht nun darin, die dem Dünngitterprinzip immanente Idee nur auf die natürliche Produktstruktur zwischen Raum und Zeit anzuwenden.

Für  $\Omega \subset \mathbb{R}^d$  und  $T > 0$  seien dazu im folgenden endlich dimensionale Räume  $V_j^\Omega$  und  $V_j^T$ ,  $j \in \mathbb{N}$ , mit

$$L^2(\Omega) = \sum_{j \geq 0} V_j^\Omega, \quad L^2((0, T)) = \sum_{j \geq 0} V_j^T,$$

und  $V_j^\Omega \subset V_{j+1}^\Omega$ ,  $V_j^T \subset V_{j+1}^T$  gegeben. Des Weiteren seien  $W_j^\Omega$  und  $W_j^T$  dazugehörige Überschußräume, d.h. es gilt

$$\begin{aligned} V_j^\Omega &= V_{j-1}^\Omega \oplus W_j^\Omega, \quad j \geq 1, \\ V_j^T &= V_{j-1}^T \oplus W_j^T, \quad j \geq 1, \end{aligned}$$

und  $W_0^\Omega := V_0^\Omega$ ,  $W_0^T := V_0^T$ . Ferner seien für jedes  $j \in \mathbb{N}$  im Folgenden Basen  $\{\psi_{j,i}^\Omega\}$  bzw.  $\{\psi_{j,i}^T\}$  von  $W_j^\Omega$  bzw.  $W_j^T$  gegeben. Insbesondere läßt sich damit jede Funktion  $u^\Omega$ , bzw.  $u^T$  eindeutig mit Hilfe geeigneter Koeffizienten  $u_{j,i}^\Omega \in \mathbb{R}$  bzw.  $u_{j,i}^T \in \mathbb{R}$  darstellen als

$$u^\Omega = \sum_{j,i} u_{j,i}^\Omega \psi_{j,i}^\Omega, \quad u^T = \sum_{j,i} u_{j,i}^T \psi_{j,i}^T.$$

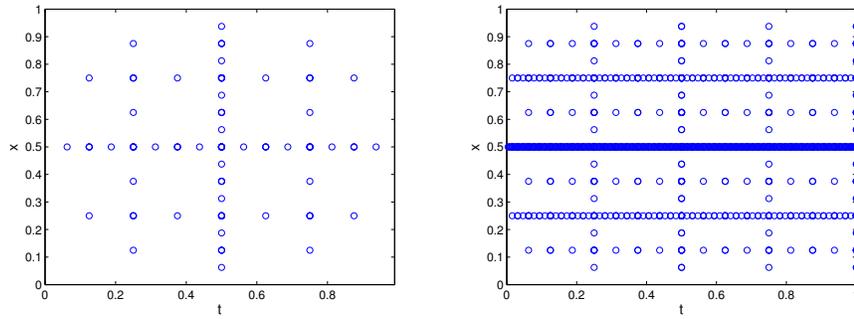
Später in diesem Kapitel werden wir beispielsweise im Ort die aus einer Sequenz von Finiten Element Räumen konstruierten hierarchischen Überschußräume und die dazugehörige hierarchische Basis [123] verwenden. Allerdings können auch gänzlich andere Zerlegungen Verwendung finden, wie etwa Zerlegungen mit Hilfe von Fourier- oder Polynombasen. Um die Allgemeinheit des Konstruktionsprinzips zu zeigen, verzichten wir an dieser Stelle deshalb zunächst darauf, die Räume  $W_j^\Omega$  und  $W_j^T$  näher zu spezifizieren. In den späteren Untersuchungen werden wir die für die Theorie benötigten Bedingungen an die Räume formulieren und erst zum Ende dieses Kapitels werden wir unsere Betrachtungen auf konkret gegebene Zerlegungen einschränken.

Da  $L^2(\Omega_T) = L^2(\Omega) \otimes L^2((0, T))$  gilt, erhalten wir

$$L^2(\Omega_T) = \sum_{\mathbf{j} \in \mathbb{N}^2} V_{\mathbf{j}},$$

mit  $V_{\mathbf{j}} = V_{j_1}^\Omega \otimes V_{j_2}^T$  und die Überschußräume

$$W_{\mathbf{j}} = W_{(j_1, j_2)} := W_{j_1}^\Omega \otimes W_{j_2}^T \tag{2.8}$$



**Abbildung 2.1.** Die zu  $\tilde{V}_4^0$  (links) und  $V_4^0$  (rechts) für  $d = 1$  gehörigen Dünngitter, wobei zur Konstruktion der Multilevelhierarchien die lineare hierarchische Basis in Raum und Zeit verwendet wurde (die Punkte markieren die Stellen, an denen die entsprechenden Basisfunktionen den Wert 1 annehmen).

sowie die Basen  $\{\psi_{\mathbf{j},\mathbf{i}} := \psi_{j_1,i_1}^\Omega \cdot \psi_{j_2,i_2}^T\}$ . Zur Vereinfachung der Notation schreiben wir, sofern dies nicht zu Mißverständnissen führt, die Normen ohne Angabe des Integrationsgebiets, d.h. wir schreiben beispielsweise  $\|\cdot\|_{L^2}$  anstelle von  $\|\cdot\|_{L^2(\Omega)}$ .

**Definition 2.2.** Für  $l \in \mathbb{N}$  definieren wir die Vollgitterräume

$$V_l^\infty := \bigoplus_{\substack{j_1 \leq l, \\ j_2 \leq 2l}} W_{\mathbf{j}} \quad (2.9)$$

und

$$\tilde{V}_l^\infty := \bigoplus_{\|\mathbf{j}\|_\infty \leq l} W_{\mathbf{j}} \quad (2.10)$$

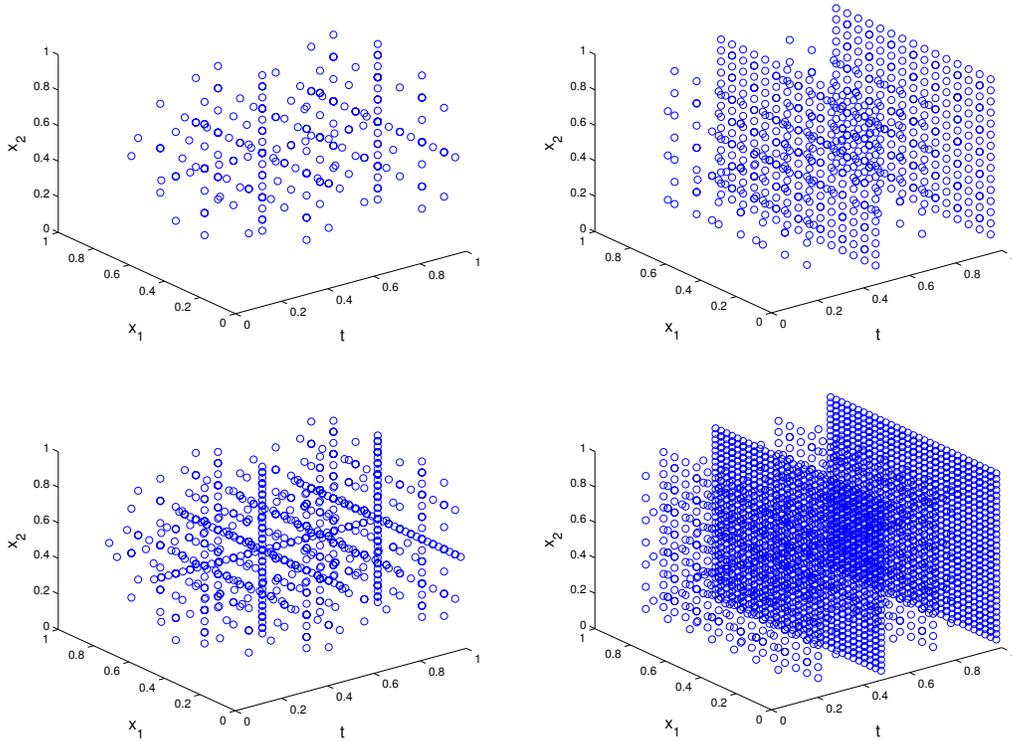
sowie die Raum-Zeit Dünngitterräume

$$V_l^0 := \bigoplus_{2j_1 + j_2 \leq 2l} W_{\mathbf{j}} \quad (2.11)$$

und

$$\tilde{V}_l^0 := \bigoplus_{\|\mathbf{j}\|_1 \leq l} W_{\mathbf{j}}. \quad (2.12)$$

Unsere Definition des Raum-Zeit Dünngitterraumes  $\tilde{V}_l^0$  ist ähnlich zu den klassischen Dünngitterräumen, vgl. [34, 35, 63]. Allerdings sind die Träger der Basisfunktionen isotrop bzgl. des Raumes und eine Anisotropie ist lediglich zwischen Raum und Zeit vorhanden. Man beachte, dass der Vollgitterraum  $V_l^\infty$  und der Dünngitterraum  $V_l^0$  in der Zeit doppelt so fein wie im Ort aufgelöst sind und deshalb mehr Unbekannte enthalten als die entsprechenden Räume  $\tilde{V}_l^\infty$  und  $\tilde{V}_l^0$ .



**Abbildung 2.2.** *Klassische Dünngitter in Raum-Zeit ( $d = 2$ ) (links) und Raum-Zeit Dünngitter (rechts) auf Level  $l = 4$  (oben) und Level  $l = 5$  (unten).*

Diese Eigenschaft ist besonders deutlich in Abbildung 2.1 zu sehen. Hier haben wir zur Konstruktion der Dünngitterräume  $V_4^0$  und  $\tilde{V}_4^0$  für  $\Omega = (0, 1)$ ,  $T = 1.0$  als Überschubräume  $W_j^\Omega$  und  $W_j^T$  die hierarchischen Überschubräume, die durch die stückweise lineare hierarchische Basis  $\{\psi_{j,i}^\Omega\}$  bzw.  $\{\psi_{j,i}^T\}$  [55, 123, 124] in Ort und Zeit aufgespannt werden, verwendet. Die Punkte in der Abbildung markieren die Orte, an denen die Basisfunktionen  $\{\psi_{\mathbf{j},\mathbf{i}} = \psi_{j_1,i_1}^\Omega \cdot \psi_{j_2,i_2}^T\}$  der Dünngitter  $V_4^0$  und  $\tilde{V}_4^0$  jeweils den Wert 1 annehmen. Besondere Beachtung verdient dabei die Tatsache, dass in dem Fall  $d = 1$  die Raum-Zeit Dünngitter mit den aus der linearen hierarchischen Basis konstruierten klassischen Dünngittern übereinstimmen. Allerdings ist dies offensichtlich für  $d > 1$  nicht mehr der Fall. Während bei den klassischen Dünngittern in jeder Zeitscheibe wiederum ein klassisches Dünngitter im Ort verwendet wird, so haben wir bei den Raum-Zeit Dünngittern in jeder Zeitscheibe ein volles Gitter. In Abbildung 2.2 sind dazu für den Fall  $d = 2$  klassische und Raum-Zeit Dünngitter dargestellt, wobei wir hier die bilineare hierarchische Basis auf einem uniformen Gitter im Ort und die eindimensionale lineare hierarchische Basis in der Zeit zur Konstruktion des Raum-Zeit Dünngitters verwendet haben. In der Abbil-

ung sind die Punkte an den Stellen, an denen die Basisfunktionen des jeweiligen Dünngitterraumes den Wert 1 annehmen.

Anhand der Definition 2.2 ist erkennbar, dass die zum Level  $l$  gehörigen Dünngitterräume eine echte Teilmenge der Vollgitterräume sind. Daher besitzen die Raum-Zeit Dünngitterräume offensichtlich weniger Freiheitsgrade als die Vollgitterräume. Eine quantitative Aussage über die Dimensionen dieser Räume liefert das folgende Lemma.

**Lemma 2.3.** *Unter der Annahme, dass für die Überschußräume  $\dim(W_j^\Omega) = O(2^{d \cdot j})$  und  $\dim(W_j^T) = O(2^j)$  gilt, erhalten wir die Abschätzungen*

$$\dim(V_l^\infty) = O(2^{(d+2)l}), \quad (2.13)$$

$$\dim(V_l^0) = \begin{cases} O(2^{2l}) & \text{für } d = 1, \\ O(2^{2l}) & \text{für } d = 2, \\ O(2^{dl}) & \text{für } d > 2, \end{cases} \quad (2.14)$$

$$\dim(\tilde{V}_l^\infty) = O(2^{(d+1)l}), \quad (2.15)$$

$$\dim(\tilde{V}_l^0) = \begin{cases} O(2^l) & \text{für } d = 1, \\ O(2^{dl}) & \text{für } d > 1. \end{cases} \quad (2.16)$$

**Beweis.** Die Abschätzungen (2.13) und (2.15) ergeben sich unmittelbar aus der Definition dieser Räume. Aus der Definition (2.11) für den Dünngitterraum  $V_l^0$  folgt

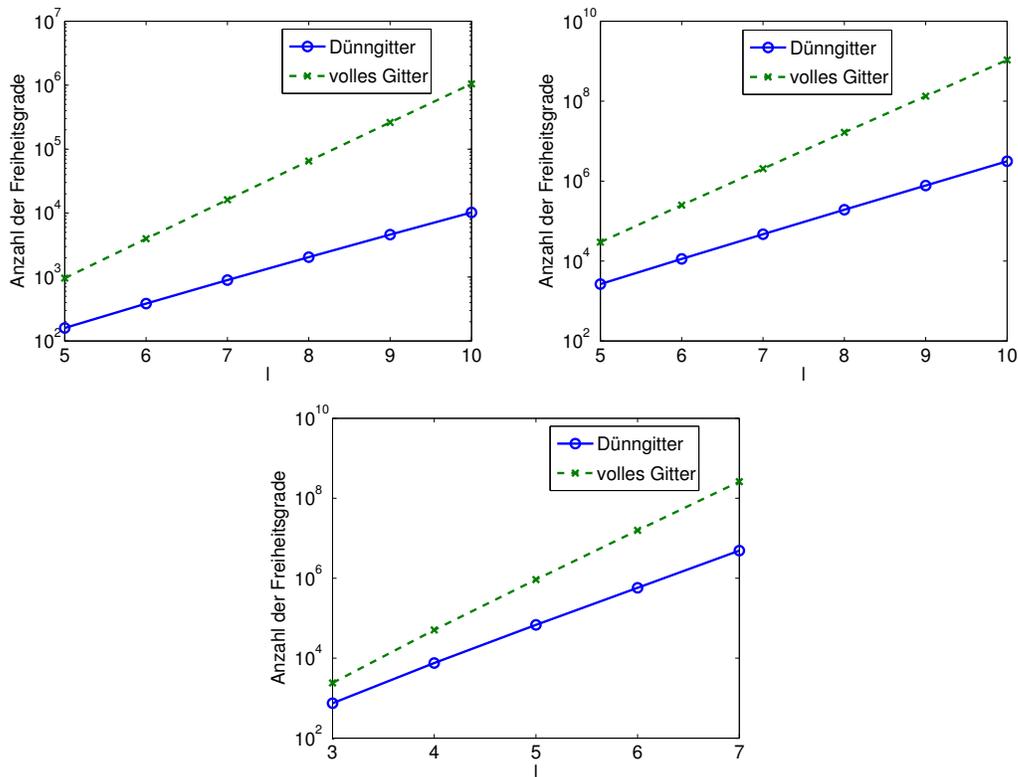
$$\begin{aligned} \dim(V_l^0) &= \sum_{j_1 \leq l} \left( \dim(W_{j_1}^\Omega) \sum_{j_2 \leq 2l - 2j_1} \dim(W_{j_2}^T) \right) \\ &\leq c \sum_{j_1 \leq l} \dim(W_{j_1}^\Omega) 2^{2l - 2j_1} \\ &\leq c 2^{2l} \sum_{j_1 \leq l} 2^{(d-2)j_1}. \end{aligned}$$

Da  $\sum_{j_1 \leq l} 2^{(d-2)j_1} \leq 2$  für  $d = 1$  und  $\sum_{j_1 \leq l} 2^0 = l + 1$  im Falle  $d = 2$  gilt, ist (2.14) für  $d \leq 2$  gezeigt. Für  $d > 2$  folgt

$$2^{2l} \sum_{j_1 \leq l} 2^{(d-2)j_1} = 2^{2l} 2^{(d-2)l} \sum_{j_1 \leq l} 2^{(d-2)(j_1-l)} \leq c 2^{dl}$$

und somit auch die Behauptung. Für  $\tilde{V}_l^0$  gilt

$$\begin{aligned} \dim(\tilde{V}_l^0) &= \sum_{j_1=0}^l \dim(W_{j_1}^T) \dim(V_{l-j_1}^\Omega) \leq c \sum_{j_1=0}^l 2^{j_1} 2^{d(l-j_1)} \\ &= c 2^{d \cdot l} \sum_{j_1=0}^l 2^{(1-d)j_1}, \end{aligned}$$



**Abbildung 2.3.** Anzahl der Freiheitsgrade eines vollen Gitters und des mit Hilfe der hierarchischen Basis in der Zeit und im Ort konstruierten Raum-Zeit Dünngitters  $\tilde{V}_l^0$  für  $d = 1$ , (oben links),  $d = 2$  (oben rechts) und  $d = 3$  (unten).

und wir erhalten (2.16).  $\square$

Das obige Lemma zeigt, dass die Dimension der Dünngitterräume (bis auf einen logarithmischen Faktor) von gleicher Ordnung wie die Dimension der entsprechenden Approximationsräume  $V_l^\Omega$  im Ort ist. Damit besitzen die Raum-Zeit Dünngitter wesentlich weniger Freiheitsgrade als die vollen Gitter. Besonders anschaulich ist dies in Abbildung 2.3 zu beobachten, wo die Dimensionen des mit Hilfe der eindimensionalen hierarchischen Basis in der Zeit und der  $d$ -linearen hierarchischen Basis im Ort konstruierten Dünngitters  $\tilde{V}_l^0$  und des vollen Gitters  $\tilde{V}_l^\infty$  für  $d = 1, 2, 3$  in Abhängigkeit des Levels  $l$  dargestellt sind.

Vergleichen wir die Dimension der Raum-Zeit Dünngitter allerdings mit der Dimension der klassischen Dünngitterräume, die im Raum-Zeit Fall (hier muss auch die Zeit als zusätzliche Dimension beachtet werden) von der Ordnung  $O(2^{ld})$  ist, so sehen wir, dass die soeben definierten Räume wesentlich mehr Freiheitsgrade besitzen. Dies scheint auf den ersten Blick ein Rückschritt zu sein. Es wird sich in dieser

Arbeit allerdings noch herausstellen, dass die Raum-Zeit Dünngitter dafür andere Vorteile gegenüber den klassischen Dünngittern besitzen. So reduziert sich etwa das Problem der Behandlung einer allgemeinen Geometrie  $\Omega$  auf das Problem der Konstruktion einer Multilevelbasis auf  $\Omega$ , das z.B. auch schon aus anderen Verfahren wie etwa den Mehrgitterverfahren bekannt und teilweise auch gelöst ist. Es wird sich in dem Abschnitt über die Diskretisierung parabolischer Probleme zusätzlich zeigen, dass Raum-Zeit Dünngitter problemlos auch für Operatoren mit jedem beliebigen, in der Zeit konstanten, elliptischen Operator zweiter Ordnung effizient anwendbar sind, während wir, wie eingangs diskutiert, bei klassischen Dünngittern dabei auf Schwierigkeiten stoßen.

Wir beschäftigen uns nun mit der Approximationseigenschaft der soeben eingeführten Räume. Dabei wird sich herausstellen, dass wir hier unter schwächeren Glattheitsannahmen als bei klassischen Dünngittern üblich, nahezu die gleichen Approximationsraten von vollen Gitter erhalten. Hier wird ein weiterer Vorteil gegenüber den klassischen dünnen Gittern erkennbar, die wesentlich stärkere Glattheitsvoraussetzungen benötigen. Wir werden später sogar anhand von klassischen Regularitätsresultaten feststellen, dass die Lösungen parabolischer Probleme die benötigten Glattheitsvoraussetzungen erfüllen.

Selbstverständlich kommen wir bei unserer Analyse bzgl. der Approximationseigenschaft nicht ohne geeignete Annahmen an die Räume  $V_j^\Omega$  und  $V_j^T$  aus. Wie bei klassischen Dünngittern unterscheiden wir im Wesentlichen zwischen zwei Fällen.

Zunächst betrachten wir die Approximationseigenschaft unter der Annahme, dass die Sequenz der Räume  $V_j^\Omega$  und  $V_j^T$  gewisse Normäquivalenzen erfüllt. Für ein einfaches rechteckiges Gebiet  $\Omega$  existieren eine Reihe von Multiskalenbasen, wie etwa orthogonale Wavelets, Spline-Wavelets, Prewavelets, biorthogonale Wavelets, lifting Wavelets und ähnliche Konstruktionen, die aus Translation und Dilatation einer Mutterfunktion gewonnen werden, siehe dazu auch [38, 40, 41, 44, 103] und die Referenzen dort.

Da für beliebige Gebiete die Konstruktion solcher Räume ausgesprochen schwierig ist [42, 43, 87, 102], werden wir anschließend untersuchen, wie sich die Konvergenzrate ändert, wenn nur noch einseitige Normabschätzungen vorhanden sind, wie dies etwa bei der Verwendung der hierarchischen Basis der Fall ist. Wir folgen in diesem Abschnitt im wesentlichen der ausführlichen Darstellung in [79], in der sich analoge Resultate für klassische Dünngitter finden.

Wie üblich nennen wir zwei Normen  $\|\cdot\|$ ,  $\|\|\cdot\|\|$  äquivalent,  $\|\cdot\| \simeq \|\|\cdot\|\|$ , wenn es Konstanten  $c_1, c_2 > 0$  gibt, so dass

$$c_1\|v\| \leq \|\|v\|\| \leq c_2\|v\|$$

für alle  $v$  gilt. Wir nehmen nun an, dass für fest gegebene  $s, t > 0$  und für jedes  $v^\Omega \in H^s(\Omega)$ ,  $v^\Omega = \sum_j w_j^\Omega$ ,  $w_j^\Omega \in W_j^\Omega$ , und jedes  $v^T \in H^t((0, T))$ ,  $v^T = \sum_j w_j^T$ ,

$w_j^T \in W_j^T$  die Normäquivalenzen

$$\|v^\Omega\|_{H^s}^2 \simeq \sum_j 2^{2sj} \|w_j^\Omega\|_{L^2}^2, \quad (2.17)$$

$$\|v^T\|_{H^t}^2 \simeq \sum_j 2^{2tj} \|w_j^T\|_{L^2}^2, \quad (2.18)$$

gelten. Wir erhalten aus diesen Normäquivalenzen dann direkt entsprechende Aussagen für das Tensorprodukt.

**Lemma 2.4.** *Es gelten für feste  $s, t > 0$  die Normäquivalenzen (2.17) und (2.18). Dann erhalten wir die Normäquivalenz*

$$\|u\|_{H_{mix}^{s,t}(\Omega_T)}^2 \simeq \sum_{\mathbf{j}} 2^{2\|(sj_1, tj_2)\|_1} \|w_{\mathbf{j}}\|_{L^2(\Omega_T)}^2, \quad (2.19)$$

für  $u \in H_{mix}^{s,t}(\Omega_T)$ ,  $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$ ,  $w_{\mathbf{j}} \in W_{\mathbf{j}}$ . Gelten die Normäquivalenzen zusätzlich für  $s = 0$ , d.h. die Multiskalenzerlegung ist  $L^2$ -stabil, dann gilt für  $u \in H^s(\Omega_T)$

$$\|u\|_{H^s(\Omega_T)}^2 \simeq \sum_{\mathbf{j}} 2^{2s\|(j_1, j_2)\|_\infty} \|w_{\mathbf{j}}\|_{L^2(\Omega_T)}^2. \quad (2.20)$$

Zum Beweis des Lemmas benutzen wir die Tensorproduktarstellung (2.6) von  $H_{mix}^{s,t}(\Omega_T)$  und die Darstellung

$$H^s(\Omega_T) = (H^s(\Omega) \otimes L^2((0, T))) \cap (L^2(\Omega) \otimes H^s((0, T)))$$

sowie die folgenden zwei Propositionen über additive Unterraumzerlegungen aus [65], die wir der Vollständigkeit halber hier kurz wiederholen. Hierzu bezeichnen wir einen Hilbertraum  $H$  mit dem dazugehörigen Skalarprodukt  $a(\cdot, \cdot)$  mit  $\{H; a\}$ . Seien zwei Hilberträume  $H_1, H_2$  und dazu zwei Sequenzen von abgeschlossenen Unterräumen  $V_{1i} \subset H_1, V_{2i} \subset H_2, i \in \mathbb{N}$ , gegeben so dass

$$H_1 = \sum_i V_{1i} \text{ und } H_2 = \sum_i V_{2i}$$

gilt. Für eine Folge  $b_{li}(\cdot, \cdot)$  von Bilinearformen auf  $V_{li}, l = 1, 2$ , nennen wir  $\{H_l; a_l\} = \sum_i \{V_{li}; b_{li}\}$  eine stabile, additive Unterraumzerlegung, wenn

$$a_l(u, u) \simeq \inf_{\substack{u_i \in V_{li}, \\ u = \sum_i u_i}} \sum_i b_{li}(u_i, u_i) \quad (2.21)$$

gilt. Für das Tensorprodukt und den Schnitt von stabilen Zerlegungen gelten dann die folgenden zwei Propositionen.

**Proposition 2.5.** *Wenn die Zerlegungen  $\{H_l; a_l\} = \sum_i \{V_{li}; b_{li}\}$ ,  $l = 1, 2$ , stabil sind, dann ist die Tensorprodukt-Zerlegung*

$$\{H_1 \otimes H_2; a_1 \otimes a_2\} = \sum_{i_1} \sum_{i_2} \{V_{1i_1} \otimes V_{2i_2}; b_{1i_1} \otimes b_{2i_2}\}$$

*stabil.*

**Proposition 2.6.** *Seien  $\{\alpha_{1i}\}_{i_1} \subset \mathbb{R}_+$  und  $\{\alpha_{2i}\}_i \subset \mathbb{R}_+$ , zwei Folgen reeller positiver Zahlen. Weiter seien die Bilinearform  $b(\cdot, \cdot)$  auf den Hilberträumen  $H_1$  und  $H_2$  und stabile Unterraumzerlegungen*

$$\{H_l; a_l\} = \sum_i \{V_i; \alpha_{li} b\}, \quad l = 1, 2$$

*gegeben, wobei die Summen direkt sind. Dann gilt für alle  $\alpha_1 > 0$  und  $\alpha_2 > 0$ , dass die Zerlegung*

$$\{H_1 \cap H_2; \alpha_1 a_1 + \alpha_2 a_2\} = \sum_i \{V_i; (\alpha_1 \alpha_{1i} + \alpha_2 \alpha_{2i}) b\}$$

*stabil ist.*

Für einen Beweis dieser beiden Aussagen verweisen wir auf [79] und beginnen nun mit dem Beweis des Lemmas 2.4.

**Beweis.** Mit Proposition 2.5 und der Stabilität der Multilevel-Zerlegungen (2.17) und (2.18) von  $H^s(\Omega)$  sowie  $H^t((0, T))$  erhalten wir direkt die Stabilität der Zerlegung

$$\{H_{mix}^{s,t}(\Omega_T); \|\cdot\|_{H_{mix}^{s,t}}\} = \sum_{\mathbf{j}} \{W_{\mathbf{j}}; 2^{2\|(s \cdot j_1, t \cdot j_2)\|_1} \|\cdot\|_{L^2}\}.$$

Damit ist (2.19) gezeigt. Aus Proposition 2.5 folgt die Stabilität der Zerlegungen

$$\begin{aligned} \{H^s(\Omega) \otimes L^2((0, T)); \|\cdot\|_{H^s(\Omega) \otimes L^2((0, T))}\} &= \sum_{\mathbf{j}} \{W_{\mathbf{j}}; 2^{2(s \cdot j_1)} \|\cdot\|_{L^2}\}, \\ \{L^2(\Omega) \otimes H^s((0, T)); \|\cdot\|_{L^2(\Omega) \otimes H^s((0, T))}\} &= \sum_{\mathbf{j}} \{W_{\mathbf{j}}; 2^{2(s \cdot j_2)} \|\cdot\|_{L^2}\}. \end{aligned}$$

Nun verwenden wir die Darstellung

$$H^s(\Omega_T) = (H^s(\Omega) \otimes L^2((0, T))) \cap (L^2(\Omega) \otimes H^s((0, T)))$$

aus [79] und Proposition 2.6 liefert zusammen mit der Abschätzung

$$2^{2sj_1} + 2^{2sj_2} \leq 2 \max\{2^{2sj_1}, 2^{2sj_2}\} \leq 2 \cdot 2^{2s\|\mathbf{j}\|_\infty}$$

die Ungleichung (2.20).  $\square$

Mit Hilfe dieser Normäquivalenzen erhalten wir nun die folgenden Resultate über die Approximationseigenschaften der Voll- bzw. Dünngitterräume.

**Satz 2.7.** *Seien  $t \geq s > 0$  fest gegeben. Für die Multiskalenzerlegung bzgl. der Zeit sei die Normäquivalenz (2.18) sowohl für  $s$  als auch für  $t$  erfüllt. Die Multiskalenzerlegung bzgl. des Ortes erfülle die Normäquivalenz (2.17) mit  $2s$  und  $2t$ . Zusätzlich gelte  $W_j^\Omega \subset H^{2s}(\Omega)$  sowie  $W_j^T \subset H^s((0, T))$ . Dann gibt es eine Konstante  $c > 0$  so dass für alle  $u \in H^{2t,t}(\Omega_T)$  und alle  $l \in \mathbb{N}$*

$$\inf_{v \in V_l^\infty} \|u - v\|_{H^{2s,s}(\Omega_T)}^2 \leq c 2^{4(s-t)l} \|u\|_{H^{2t,t}(\Omega_T)}^2 \quad (2.22)$$

gilt. Ist  $u \in H_{mix}^{2t,t}(\Omega_T)$ , so erhalten wir

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H^{2s,s}(\Omega_T)}^2 \leq c 2^{4(s-t)l} \|u\|_{H_{mix}^{2t,t}(\Omega_T)}^2 \quad (2.23)$$

mit einer Konstanten  $c > 0$  unabhängig von  $u$  und  $l$ . Gelten die Normäquivalenzen (2.17) bzgl. des Ortes mit  $s, t$  und ist  $W_j^\Omega \subset H^s(\Omega)$ , dann gilt für  $u \in H^t(\Omega_T)$

$$\inf_{v \in \tilde{V}_l^\infty} \|u - v\|_{H^s(\Omega_T)}^2 \leq c 2^{2(s-t)l} \|u\|_{H^t(\Omega_T)}^2 \quad (2.24)$$

und für  $u \in H_{mix}^t(\Omega_T)$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H^s(\Omega_T)}^2 \leq c 2^{2(s-t)l} \|u\|_{H_{mix}^t(\Omega_T)}^2. \quad (2.25)$$

**Beweis.** Wir beschränken uns auf den Fall (2.23) da die anderen Fälle analog folgen. Für  $u \in H_{mix}^{2t,t}(\Omega_T)$ ,  $u = \sum_j w_j$ ,  $w_j \in W_j$ , erhalten wir mit Lemma 2.4 und den Voraussetzungen die Normäquivalenzen

$$\begin{aligned} \|u\|_{H_{mix}^{2s,s}(\Omega_T)}^2 &\simeq \sum_j 2^{2s\|(2j_1, j_2)\|_1} \|w_j\|_{L^2(\Omega_T)}^2, \\ \|u\|_{H_{mix}^{2t,t}(\Omega_T)}^2 &\simeq \sum_j 2^{2t\|(2j_1, j_2)\|_1} \|w_j\|_{L^2(\Omega_T)}^2. \end{aligned}$$

Damit folgt

$$\begin{aligned}
\inf_{v \in V_l^0} \|u - v\|_{H^{2s,s}(\Omega_T)}^2 &\leq \|u - \sum_{2j_1+j_2 \leq 2l} w_{\mathbf{j}}\|_{H^{2s,s}(\Omega_T)}^2 \\
&\leq c \sum_{2j_1+j_2 > 2l} 2^{2s\|(2j_1,j_2)\|_\infty} \|w_{\mathbf{j}}\|_{L^2(\Omega_T)}^2 \\
&= c \sum_{2j_1+j_2 > 2l} 2^{2s\|(2j_1,j_2)\|_\infty - 2t\|(2j_1,j_2)\|_1} 2^{2t\|(2j_1,j_2)\|_1} \|w_{\mathbf{j}}\|_{L^2(\Omega_T)}^2 \\
&\leq c \max_{2j_1+j_2 > 2l} 2^{2s\|(2j_1,j_2)\|_\infty - 2t\|(2j_1,j_2)\|_1} \|u\|_{H_{mix}^{2t,t}(\Omega_T)}^2 \\
&\leq c 2^{4(s-t)l} \|u\|_{H_{mix}^{2t,t}(\Omega_T)}^2.
\end{aligned}$$

□

Das obige Resultat zeigt, dass wir unter etwas stärkeren Voraussetzungen an die Glattheit der zu approximierenden Funktion  $u$  für die Raum-Zeit Dünngitter  $V_l^0$  bzw.  $\tilde{V}_l^0$  die gleiche Approximationsordnung wie für die Räume  $V_l^\infty$  bzw.  $\tilde{V}_l^\infty$  erhalten.

Wie bereits zu Anfang erwähnt, ist die Konstruktion von Multiskalenzerlegungen, die die entsprechenden Normäquivalenzen erfüllen, auf allgemeinen Geometrien schwierig oder gar unmöglich. Das folgende Lemma zeigt jedoch, dass wir nur eine obere Normabschätzung benötigen, um ein ähnlich gutes Approximationsverhalten für die Raum-Zeit Dünngitter zu erhalten.

**Lemma 2.8.** *Es seien  $\|\cdot\|_\star$  und  $\|\cdot\|_\Delta$  zwei vorgegebene Normen. Für feste  $t > s \geq 0$ , gebe es eine Konstante  $\bar{c} > 0$ , so dass für jedes  $u$  mit  $\|u\|_\star < \infty$ ,  $\|u\|_\Delta < \infty$ ,  $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$ ,  $w_{\mathbf{j}} \in W_{\mathbf{j}}$ ,*

$$\|w_{\mathbf{j}}\|_\star \leq \bar{c} 2^{s\|\mathbf{j}\|_\infty - t\|\mathbf{j}\|_1} \|u\|_\Delta \text{ für alle } \mathbf{j} \in \mathbb{N}^2 \quad (2.26)$$

*gilt. Dann gibt es eine Konstante  $c > 0$ , so dass*

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_\star \leq \begin{cases} c 2^{(s-t)l} \|u\|_\Delta & \text{für } s = 0, t > 0, \\ c 2^{(s-t)l} \|u\|_\Delta & \text{für } s > 0, t > 0. \end{cases} \quad (2.27)$$

**Beweis.** Für den Fall klassischer Dünngitter mit  $\|\cdot\|_\star = \|\cdot\|_{H^s}$  sowie  $\|\cdot\|_\Delta = \|\cdot\|_{H_{mix}^t}$  findet sich ein Beweis für die obige Behauptung in [79], der sich auf unseren Fall übertragen lässt. Aus Gründen der Vollständigkeit und um zu zeigen, dass die Parameter  $s, t$  nicht mit Sobolev-Normen zusammenhängen müssen, führen wir den Beweis dennoch an dieser Stelle.

Es gilt mit  $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$ ,  $w_{\mathbf{j}} \in W_{\mathbf{j}}$ ,

$$\begin{aligned} \inf_{v \in \tilde{V}_l^0} \|u - v\|_{\star} &\leq \sum_{\|\mathbf{j}\|_1 > l} \|w_{\mathbf{j}}\|_{\star} \leq c \sum_{\|\mathbf{j}\|_1 > l} 2^{s\|\mathbf{j}\|_{\infty} - t\|\mathbf{j}\|_1} \|u\|_{\Delta} \\ &\leq c \|u\|_{\Delta} \sum_{\|\mathbf{j}\|_1 > l} 2^{s\|\mathbf{j}\|_{\infty} - t\|\mathbf{j}\|_1}. \end{aligned}$$

Für  $s > 0$  erhalten wir

$$\sum_{\|\mathbf{j}\|_1 > l} 2^{s\|\mathbf{j}\|_{\infty} - t\|\mathbf{j}\|_1} = \sum_{k=l+1}^{\infty} 2^{-tk} \sum_{\|\mathbf{j}\|_1 = k} 2^{s\|\mathbf{j}\|_{\infty}}$$

und schätzen weiter ab

$$\begin{aligned} \sum_{\|\mathbf{j}\|_1 = k} 2^{s\|\mathbf{j}\|_{\infty}} &\leq 2 \sum_{\|\mathbf{j}\|_1 = k} 2^{sj_1} = 2 \sum_{j_1=0}^k 2^{sj_1} \\ &\leq 2 \cdot 2^{sk} \sum_{j_1=0}^k 2^{s(j_1-k)} \leq c 2^{sk}. \end{aligned}$$

Damit gilt

$$\sum_{\|\mathbf{j}\|_1 > l} 2^{s\|\mathbf{j}\|_{\infty} - t\|\mathbf{j}\|_1} \leq c 2^{(s-t)l},$$

woraus die Behauptung für  $s > 0$  folgt. Für den Fall  $s = 0$  gilt

$$\begin{aligned} \sum_{\|\mathbf{j}\|_1 > l} 2^{-t\|\mathbf{j}\|_1} &\leq \sum_{k=l+1}^{\infty} 2^{-tk} \sum_{\|\mathbf{j}\|_1 = k} 1 = \sum_{k=l+1}^{\infty} 2^{-tk} (k+1) \\ &= 2^{-tl} \sum_{k=1}^{\infty} 2^{-tk} (k+l+1) \\ &\leq c 2^{-tl} l \sum_{k=1}^{\infty} 2^{-tk} k. \end{aligned}$$

Die unendliche Summe ist konvergent und somit folgt die Behauptung.  $\square$

Ein ähnliches Resultat erhalten wir für den Dünngitterraum  $V_l^0$ .

**Lemma 2.9.** *Es seien  $\|\cdot\|_{\star}$  und  $\|\cdot\|_{\Delta}$  zwei vorgegebene Normen. Für feste  $t > s \geq 0$ , gebe es eine Konstante  $\bar{c} > 0$ , so dass für jedes  $u$  mit  $\|u\|_{\star} < \infty$ ,  $\|u\|_{\Delta} < \infty$ ,  $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$*

$$\|w_{\mathbf{j}}\|_{\star} \leq \bar{c} 2^{s\|(2j_1, 2j_2)\|_{\infty} - t\|(2j_1, 2j_2)\|_1} \|u\|_{\Delta} \text{ für alle } \mathbf{j} \in \mathbb{N}^2 \quad (2.28)$$

gilt. Dann gibt es eine Konstante  $c > 0$  unabhängig von  $l$  und  $u$  mit

$$\inf_{v \in V_l^0} \|u - v\|_* \leq \begin{cases} c 2^{(s-t)2l} \cdot l \|u\|_\Delta & \text{für } s = 0, t > 0, \\ c 2^{(s-t)2l} \|u\|_\Delta & \text{für } s > 0, t > 0. \end{cases} \quad (2.29)$$

**Beweis.** Es gilt für  $u = \sum_j w_j$

$$\inf_{v \in V_l^0} \|u - v\|_* \leq c \cdot \|u\|_\Delta \sum_{2j_1 + j_2 > 2l} 2^{s\|(2j_1, j_2)\|_\infty - t\|(2j_1, j_2)\|_1}.$$

Für  $s > 0$  gilt

$$\begin{aligned} \sum_{2j_1 + j_2 > 2l} 2^{s\|(2j_1, j_2)\|_\infty - t\|(2j_1, j_2)\|_1} &\leq \sum_{k \geq 2l+1} 2^{-tk} \sum_{2j_1 + j_2 = k} 2^{s\|(2j_1, j_2)\|_\infty} \\ &\leq 2 \sum_{k \geq 2l+1} 2^{-tk} \sum_{2j_1 + j_2 = k} 2^{sj_2} \leq c \sum_{k \geq 2l+1} 2^{-tk} \sum_{j_2=0}^k 2^{sj_2} \\ &\leq c \sum_{k \geq 2l+1} 2^{(s-t)k} \end{aligned}$$

womit die Behauptung für  $s > 0$  folgt. Im Falle  $s = 0$  gilt

$$\begin{aligned} \sum_{2j_1 + j_2 > 2l} 2^{-t\|(2j_1, j_2)\|_1} &\leq \sum_{k=2l+1}^{\infty} 2^{-tk} \sum_{2j_1 + j_2 = k} 1 \\ &\leq c \sum_{k=2l+1}^{\infty} 2^{-tk} k \leq c 2^{-2tl} \sum_{k=0}^{\infty} 2^{-tk} k. \end{aligned}$$

□

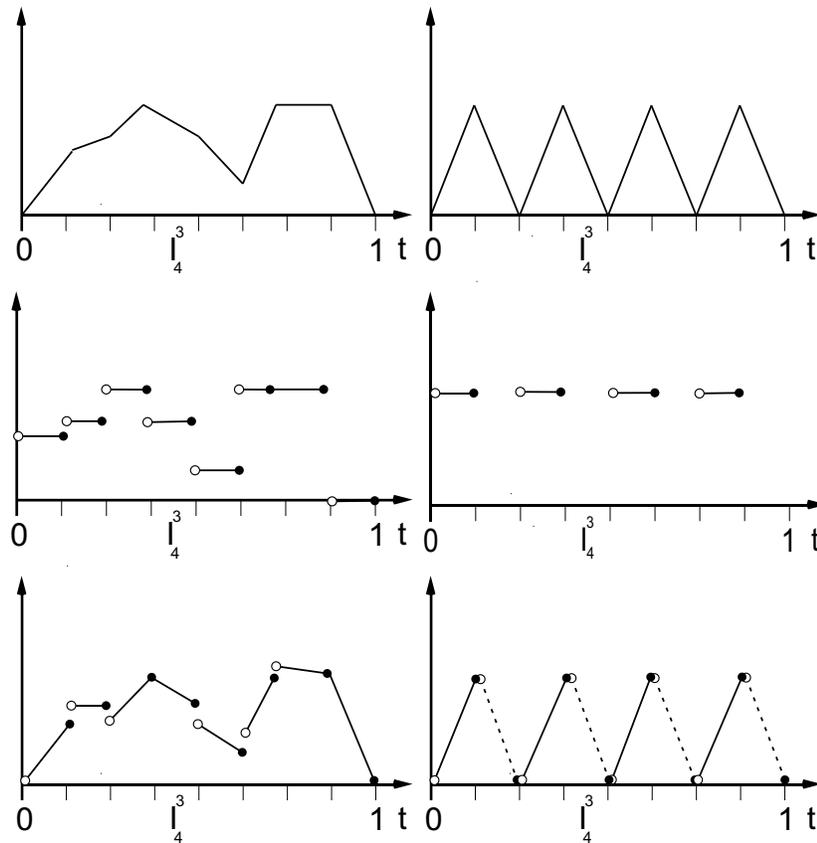
Wir wählen für festes  $t > s \geq 0$  in Lemma 2.8  $\|\cdot\|_* = \|\cdot\|_{H^s}$  und  $\|\cdot\|_\Delta = \|\cdot\|_{H_{mix}^t(\Omega_T)}$  und nehmen an, dass für jedes  $u \in H^t(\Omega_T)$  die Abschätzung (2.26) gilt. Damit erhalten wir im Falle  $s = 0$  die Abschätzung

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{L^2(\Omega_T)} \leq c 2^{-t} l \|u\|_{H_{mix}^t(\Omega_T)}$$

die sich verglichen mit den Ergebnissen des Satzes 2.7 lediglich um den logarithmischen Faktor  $l$  unterscheidet, während wir für  $s > 0$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H^s(\Omega_T)} \leq c 2^{s-t} \|u\|_{H_{mix}^t(\Omega_T)}$$

und somit sogar die gleichen Raten wie im Fall von Lemma 2.8 erhalten. Als Konsequenz aus Lemma 2.8 und Lemma 2.9 resultiert die Feststellung, dass es ausreicht, Zerlegungen zu verwenden, die obere Normabschätzungen der Art (2.26) und (2.28) erfüllen, um die (bis auf einen logarithmischen Faktor) gleiche Approximationsordnung wie im Fall von Zerlegungen, die die Normäquivalenzen (2.17) und (2.18) erfüllen, zu erhalten.



**Abbildung 2.4.** Beispielfunktion aus den Räumen  $V_{3,1}^T$  (oben links),  $V_{3,0}^T$  (mitte links) und  $V_{3,1}^{T,disc}$  (unten links) sowie die entsprechenden Basisfunktionen der hierarchischen Überschlußräume  $W_{3,1}^T$  (oben rechts),  $W_{3,0}^T$  (mitte rechts) und  $W_{3,1}^{T,disc}$  (unten rechts).

## 2.3 Eindimensionale Multiskalenzerlegungen

Nachdem wir in dem vorherigen Abschnitt die Approximationseigenschaften der Raum-Zeit Dünngitterräume diskutiert haben, stellen wir in diesem Abschnitt kurz drei mögliche Multilevel-Zerlegungen von  $L^2((0, T))$  und dazugehörige eindimensionale Multilevelbasen vor. Diese Zerlegungen werden wir später zur Konstruktion der Raum-Zeit Dünngitterräume für die verschiedenen Diskretisierungsverfahren benötigen. Zur Vereinfachung der Darstellung beschränken wir uns dabei auf Funktionen mit Nullrandwerten. Die Überlegungen lassen sich jedoch direkt auf Funktionen mit beliebigen Randwerten übertragen.

Wir nehmen zur Vereinfachung der Darstellung  $T = 1$  an. Für festes  $l \in \mathbb{N}$

unterteilen wir das Intervall  $(0, 1]$  in  $2^l$  Teilintervalle  $I_k^l$ ,  $1 \leq k \leq 2^l$ ,

$$I_k^l = ((k-1)2^{-l}, k2^{-l}].$$

Mit  $V_{l,1}^T$  bezeichnen wir den Raum der auf jedem Teilintervall  $I_k^l$  linearen, stetigen Funktionen mit der üblichen nodalen Basis  $\varphi_{l,1,k}^T$ ,  $1 \leq k < 2^l$ ,

$$\varphi_{l,1,k}^T(t) := \begin{cases} 1 - |k - 2^l \cdot t| & \text{für alle } t \in I_k^l \cup I_{k+1}^l, \\ 0 & \text{sonst.} \end{cases} \quad (2.30)$$

Als Überschußräume  $W_{l,1}^T$  wählen wir hier die durch die klassische lineare hierarchische Basis  $\psi_{l,1,k}^T := \varphi_{l,1,2k-1}^T$ ,  $1 \leq k < 2^{l-1}$ , auf Level  $l$  aufgespannten Räume,

$$V_{l,1}^T = V_{l-1,1}^T \oplus W_{l,1}^T = V_{l-1,1}^T \oplus \text{span}\{\psi_{l,1,k}^T \mid 1 \leq k < 2^{l-1}\}. \quad (2.31)$$

Wir betrachten nun die Zerlegung in die Räume  $V_{l,0}^T$  der auf jedem Teilintervall  $I_k^l$  stückweise konstanten Funktionen. Dieser Raum wurde bereits bei der Konstruktion klassischer Dünngitter in [5] verwendet. Eine Basis dieses Raumes bilden die Funktionen  $\varphi_{l,0,k}^T$ ,  $1 \leq k < 2^l$ , die auf dem Teilintervall  $I_k^l$  konstant eins sind und überall sonst verschwinden,

$$\varphi_{l,0,k}^T(t) := \begin{cases} 1 & \text{falls } t \in I_k^l, \\ 0 & \text{sonst.} \end{cases} \quad (2.32)$$

Hierbei verwenden wir im folgenden die Überschußräume  $W_{l,0}^T$ , die durch die stückweise konstante hierarchische Basis  $\psi_{l,0,k}^T := \varphi_{l,0,2k-1}^T$ ,  $1 \leq k < 2^{l-1}$ , aufgespannt werden, d.h.

$$V_{l,0}^T := V_{l-1,0}^T \oplus W_{l,0}^T = V_{l-1,0}^T \oplus \text{span}\{\psi_{l,0,k}^T \mid 1 \leq k < 2^{l-1}\}. \quad (2.33)$$

Schließlich betrachten wir noch die Räume  $V_{l,1}^{T,disc}$  der auf jedem Teilintervall  $I_k^l$  linearen Funktionen. Die Funktionen  $\varphi_{l,k,disc}^T$  und  $\tilde{\varphi}_{l,k,disc}^T$ ,

$$\begin{aligned} \varphi_{l,disc,k}^T(t) &:= \begin{cases} 1 - |k - 2^l t| & \text{falls } t \in I_k^l, \\ 0 & \text{sonst,} \end{cases} \\ \tilde{\varphi}_{l,disc,k}^T(t) &:= \begin{cases} 1 - |(k-1) - 2^l t| & \text{falls } t \in I_k^l, \\ 0 & \text{sonst,} \end{cases} \end{aligned}$$

bilden eine Basis von  $V_{l,1}^{T,disc}$ . Bemerkenswert ist, dass  $V_{l,1}^T \subset V_{l,1}^{T,disc}$  gilt und insbesondere die Dimension dieses Raumes nahezu doppelt so groß wie die Dimension des Raumes  $V_{l,1}^T$  ist, was später bei der Kosten-Nutzen Diskussion der unterschiedlichen Diskretisierungsverfahren von Bedeutung sein wird. Ein möglicher Überschußraum  $W_{l,1}^{T,disc}$  wird durch die hierarchische Basis

$$\psi_{l,disc,k}^T := \varphi_{l,disc,2k+1}^T, \quad \tilde{\psi}_{l,disc,k}^T := \tilde{\varphi}_{l,disc,2k}^T,$$

mit  $1 \leq k < 2^{l-1}$ , aufgespannt,

$$\begin{aligned} V_{l,1}^{T,disc} &= V_{l-1,1}^{T,disc} \oplus W_{l,1}^{T,disc} \\ &= V_{l-1,1}^{T,disc} \oplus \text{span}(\{\psi_{l,disc,k}^T \mid 1 \leq k \leq 2^{l-1}\} \cup \{\tilde{\psi}_{l,disc,k}^T \mid 1 \leq k \leq 2^{l-1}\}). \end{aligned} \quad (2.34)$$

Abbildung 2.4 zeigt für  $l = 3$  für jeden der soeben definierten Räume Beispielfunktionen und die jeweiligen Basisfunktionen der hierarchischen Überschußräume.

Im folgenden Abschnitt untersuchen wir die Raum-Zeit Dünngitter im Hinblick auf Lemma 2.8, die aus der Konstruktion mit den soeben vorgestellten eindimensionalen Multilevel-Zerlegungen in der Zeit und der aus einer Sequenz von Finite Element Räumen gewonnenen linearen bzw.  $d$ -linearen hierarchischen Basis konstruiert werden.

## 2.4 Verwendung der hierarchischen Basis

Sei  $\mathcal{T}_j$ ,  $j \in \mathbb{N}$ , eine geschachtelte Folge von Triangulierungen (in Dreiecke oder Vierecke bzw. Tetraeder oder Hexaeder) des Gebietes  $\Omega \subset \mathbb{R}^d$ ,  $1 \leq d \leq 3$ ,  $\mathcal{T}_j \subset \mathcal{T}_{j+1}$ , d.h. die Menge  $\mathcal{N}_j$  der Knoten von  $\mathcal{T}_j$  ist eine echte Teilmenge der Knoten  $\mathcal{N}_{j+1}$  von  $\mathcal{T}_{j+1}$ ,

$$\mathcal{N}_j \subset \mathcal{N}_{j+1}.$$

Wir bezeichnen den Durchmesser von  $T \in \mathcal{T}_j$  mit  $h(T)$  und definieren

$$\rho(T) := \sup\{\text{diam}(B) \mid B \subset T, B \text{ ist eine Kugel}\}.$$

Es sei  $\mathcal{T}_j$  eine Folge von *regulären* Triangulierungen, d.h. es gibt von  $j$  unabhängige Konstanten  $\sigma > 0$  und  $c > 0$ , so dass

$$\frac{h(T)}{\rho(T)} \leq \sigma, \quad h_j := \max_{T \in \mathcal{T}_j} h(T) \leq c2^{-j}, \quad \text{für alle } j \text{ und } T \in \mathcal{T}_j$$

gilt. Des Weiteren nehmen wir an, dass die Anzahl der Knoten jeder Triangulierung  $\mathcal{T}_j$  durch

$$|\mathcal{N}_j| \leq c2^{dj} \quad (2.35)$$

mit  $c > 0$  unabhängig von  $j$  abgeschätzt werden kann. Sei  $V_j^\Omega$  der zur Triangulierung  $\mathcal{T}_j$  gehörige Finite Element Raum der stückweise linearen bzw.  $d$ -linearen Funktionen. Ist eine Abbildung  $P_j : C^0(\Omega) \rightarrow V_j^\Omega$  gegeben, so hat für jedes  $u \in C^0(\Omega)$  die Funktion  $u_l := P_l u \in V_l^\Omega$  die Darstellung

$$u_l = P_0 u_l + \sum_{j=0}^{l-1} (P_{j+1} u_l - P_j u_l), \quad (2.36)$$

mit den Überschußräumen

$$W_{j+1}^\Omega := \text{range}(P_{j+1} - P_j) \quad (2.37)$$

und  $W_0^\Omega := V_0^\Omega$ . Wählen wir zum Beispiel  $P_j u$  als den Interpolanten von  $u \in C^0(\Omega)$  in  $V_j^\Omega$ , d.h.

$$P_j u(x) = u(x) \text{ für alle } x \in \mathcal{N}_j, \quad (2.38)$$

so erhalten wir eine Zerlegung in Überschubräume  $W_j^\Omega$ , die von der hierarchischen Basis [123] aufgespannt werden. Wählen wir hingegen  $P_j u$  als die  $L^2$ -Projektion auf  $V_j$ , so spannen  $L^2$ -orthogonale Waveletbasen die resultierenden Überschubräume auf.

Im folgenden wählen wir für  $P_j$  den Interpolationsoperator (2.38). Sei  $\varphi_{j,i}^\Omega$ ,  $1 \leq i \leq \dim(V_j^\Omega)$  die nodale Basis des Finite Element Raumes  $V_j^\Omega$ . Eine Teilmenge dieser nodalen Basis bildet wiederum eine Basis der hierarchischen Überschubräume und wir bezeichnen die Elemente dieser hierarchischen Basis mit  $\psi_{j,i}^\Omega$ ,  $1 \leq i \leq \dim(W_j^\Omega)$ . Zur Vereinfachung nehmen wir (nach Ummummerierung) an, dass der Index  $i$  eines Basiselements  $\psi_{j,i}^\Omega$  den nodalen Punkt  $n_i \in \mathcal{N}_j$  bezeichnet, an dem die Basisfunktion eins ist, d.h. für beliebiges  $n_k \in \mathcal{N}_j$  gilt

$$\psi_{j,i}^\Omega(n_k) = \begin{cases} 1 & \text{falls } k = i, \\ 0 & \text{sonst} \end{cases}.$$

Wir erhalten aus der Definition (2.37) des Raumes  $W_j^\Omega$  für die Koeffizienten  $w_{j,i}^\Omega$  der hierarchischen Überschüsse  $w_j^\Omega = \sum_{i=1}^{\dim(W_j)} w_{j,i}^\Omega \psi_{j,i}^\Omega$  die Darstellung

$$w_{j,i}^\Omega = (P_j u - P_{j-1} u)(n_i). \quad (2.39)$$

Analoges gilt für die eindimensionalen Multilevelzerlegungen des vorigen Abschnitts.

Wir untersuchen nun mit Hilfe von Lemma 2.8 bzw. Lemma 2.9 die Approximati- onseigenschaften der Raum-Zeit Dünngitter, die mit Hilfe der Multilevelzerlegungen des vorigen Abschnitts in der Zeit und der ( $d$ -) linearen hierarchischen Basis  $\{\psi_{j,i}^\Omega\}$  im Ort konstruiert werden. Dazu nehmen wir an, dass die Folge von Finite Element Räumen *affin äquivalent* ist, d.h. jedes Element lässt sich durch eine affine Abbil- dung auf ein Referenzelement zurückführen. Zusätzlich benötigen wir die folgende klassische Interpolationsfehlerabschätzung für Finite Elemente, siehe z.B. [25, 39].

**Lemma 2.10.** *Für eine reguläre Folge von Triangulierungen und linearen bzw.  $d$ -linearen, affin äquivalenten Finite Elementen gibt es eine Konstante  $c > 0$ , so dass für alle  $u \in H^2(\Omega)$*

$$\|u - P_j u\|_{L^2(\Omega)} \leq c 2^{-2j} \|u\|_{H^2(\Omega)}, \quad (2.40)$$

$$\|u - P_j u\|_{H^1(\Omega)} \leq c 2^{-j} \|u\|_{H^2(\Omega)}, \quad (2.41)$$

*unabhängig von  $j$  gilt. Für  $u \in H^{2,\infty}(\Omega)$  gilt außerdem*

$$\|u - P_j u\|_{L^\infty(\Omega)} \leq c \cdot 2^{-2j} \|u\|_{H^{2,\infty}(\Omega)}. \quad (2.42)$$

Wir erhalten damit für die hierarchischen Koeffizienten  $w_{j,i}^\Omega$  mit Hilfe der Relation (2.39) und der Interpolationsfehlerabschätzung (2.42) die Abschätzung des nächsten Lemmas.

**Lemma 2.11.** *Für eine reguläre Folge von Triangulierungen und linearen bzw. d-linearen affin äquivalenten Finiten Elementen gibt es eine Konstante  $c > 0$ , so dass für alle  $u \in H^{2,\infty}(\Omega)$ ,  $u = \sum_j w_j^\Omega = \sum_j \sum_i w_{j,i}^\Omega \psi_{j,i}^\Omega$ ,  $w_j^\Omega \in W_j^\Omega$ ,*

$$|w_{j,i}^\Omega| \leq \|P_j u - P_{j-1} u\|_{L^\infty(\Omega)} \leq c \cdot 2^{-2j} \|u\|_{H^{2,\infty}(\Omega)} \quad (2.43)$$

gilt. Weiter ist

$$\|w_j^\Omega\|_{L^2(\Omega)} \leq c 2^{-2j} \|u\|_{H^2(\Omega)}, \quad (2.44)$$

$$\|w_j^\Omega\|_{L^\infty(\Omega)} \leq c 2^{-2j} \|u\|_{H^{2,\infty}(\Omega)}. \quad (2.45)$$

Dabei ist zu berücksichtigen, dass die eindimensionale Zerlegung (2.31) in der Zeit genau der eindimensionalen hierarchischen Basis entspricht, so dass wir das gleiche Ergebnis für diese Zerlegung erhalten. Ebenso erhalten wir für die Zerlegung (2.34) das gleiche Resultat, da  $u \in H^2((0, T))$  und damit stetig ist.

Da wir annehmen, dass unsere Finiten Element Räume  $V_j^\Omega$  affin äquivalent sind und die zugrundeliegende Triangulierung regulär ist, gelten für die hierarchische Basis  $\psi_{j_1, i_1}^\Omega$  die Normabschätzungen

$$\|\psi_{j_1, i_1}^\Omega\|_{H^1(\Omega)} \leq c 2^{-(d-2) \cdot (j_1/2)}, \quad (2.46)$$

$$\|\psi_{j_1, i_1}^\Omega\|_{L^2(\Omega)} \leq c 2^{-dj_1/2}, \quad (2.47)$$

siehe z.B. [39]. Wir erhalten nun die folgende Aussage über die Approximationseigenschaft des Dünngitterraumes  $\tilde{V}_l^0$  bei Verwendung der ( $d$ -) linearen hierarchischen Basis im Ort und der linearen hierarchischen Basis in der Zeit.

**Satz 2.12.** *Sei  $V_j^\Omega$  eine Folge von ( $d$ -) linearen, affin äquivalenten Finite Element Räumen, deren zugrunde liegende Triangulierungen regulär sind. Dann gibt es eine Konstante  $c > 0$ , so dass für den aus (2.12) mit Hilfe der ( $d$ -) linearen hierarchischen Basis im Ort und der linearen hierarchischen Basis in der Zeit konstruierten Raum-Zeit Dünngitterraum  $\tilde{V}_l^0$  und für alle  $u \in H_{mix}^2(\Omega_T)$  bzw.  $u \in H_{mix}^{(2,\infty),(2,\infty)}(\Omega_T)$*

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{L^2(\Omega_T)} \leq c 2^{-2l} \|u\|_{H_{mix}^2(\Omega_T)}, \quad (2.48)$$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H_{mix}^{1,0}(\Omega_T)} \leq c 2^{-l} \|u\|_{H_{mix}^{(2,\infty),(2,\infty)}(\Omega_T)}, \quad (2.49)$$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)} \leq c 2^{-2l} \|u\|_{H_{mix}^{(2,2),(2,\infty)}(\Omega_T)} \quad (2.50)$$

*gilt.*

**Beweis.** Da wir die lineare hierarchische Basis in der Zeit verwenden, können wir Lemma 2.11 auch bzgl. der Zeit anwenden. Damit erhalten wir für  $u^T \in H^2((0, T))$  bzw.  $u^T \in H^{2,\infty}((0, T))$ ,  $u^T = \sum_j w_j^T$ ,

$$\|w_j^T\|_{L^2((0,T))} \leq c 2^{-2j} \|u^T\|_{H^2((0,T))}, \quad (2.51)$$

$$\|w_j^T\|_{L^\infty((0,T))} \leq c 2^{-2j} \|u^T\|_{H^{2,\infty}((0,T))}. \quad (2.52)$$

Mit einem Tensorproduktargument für  $u \in H_{mix}^2(\Omega_T)$ , bzw.  $u \in H_{mix}^{(2,2),(2,\infty)}(\Omega_T)$ ,  $u = \sum_j w_j$ , folgt

$$\|w_j\|_{L^2(\Omega_T)} \leq c 2^{-2\|\mathbf{j}\|_1} \|u\|_{H_{mix}^2(\Omega_T)}, \quad (2.53)$$

$$\|w_j\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)} \leq c 2^{-2\|\mathbf{j}\|_1} \|u\|_{H_{mix}^{(2,2),(2,\infty)}(\Omega_T)}. \quad (2.54)$$

Somit liefert die Anwendung des Lemmas 2.8 für  $s = 0$  und  $t = 2$  die Ungleichungen (2.48) und (2.50).

Sei nun  $u \in H_{mix}^{(2,\infty),(2,\infty)}(\Omega_T)$ ,  $u = \sum_j \sum_i w_{j,i} \psi_{j,i}$ , dann gilt wiederum mit einem Tensorproduktargument

$$|w_{j,i}| \leq c 2^{-2\|\mathbf{j}\|_1} \|u\|_{H_{mix}^{(2,\infty),(2,\infty)}(\Omega_T)}.$$

Des Weiteren gilt mit (2.46) und (2.47) für  $\psi_{j,i} = \psi_{j_1,i_1}^\Omega \cdot \psi_{j_2,i_2}^T$

$$\|\psi_{j,i}\|_{H_{mix}^{1,0}(\Omega_T)} \leq c 2^{-((d-2) \cdot j_1 + j_2)/2}.$$

Da die Folge der Triangulierungen regulär ist, kann die Anzahl der Zellen, die einen Knoten gemeinsam haben, durch eine Konstante unabhängig von dem Level  $j_1$  nach oben abgeschätzt werden [75]. Insbesondere erhalten wir damit eine obere Abschätzung für die Anzahl an Basisfunktionen  $\psi_{j_1,i_1}^\Omega$  auf einem fest vorgegebenem Level  $j_1$ , deren Träger nicht disjunkt zu einer weiteren beliebig vorgegebenen Basisfunktion des gleichen Levels sind. Ebenso folgt daraus eine solche Abschätzung für den Raum-Zeit Fall und die Basisfunktionen  $\psi_{j,i}$ .

Insgesamt erhalten wir für  $w_j = \sum_i w_{j,i} \psi_{j,i}$

$$\begin{aligned} \|w_j\|_{H_{mix}^{1,0}(\Omega_T)}^2 &\leq c \sum_i |w_{j,i}|^2 \|\psi_{j,i}\|_{H_{mix}^{1,0}(\Omega_T)}^2 \\ &\leq c \sum_i 2^{-4\|\mathbf{j}\|_1} 2^{2j_1} 2^{-dj_1 - j_2} \|u\|_{H_{mix}^{(2,\infty),(2,\infty)}(\Omega_T)}^2. \end{aligned}$$

Mit Hilfe von (2.35) schließen wir, dass die Summe  $O(2^{-dj_1 - j_2})$  Summanden besitzt. Somit gilt

$$\begin{aligned} \|w_j\|_{H_{mix}^{1,0}(\Omega_T)}^2 &\leq c 2^{-4\|\mathbf{j}\|_1} 2^{2j_1} 2^{-dj_1 - j_2} \|u\|_{H_{mix}^{(2,\infty),(2,\infty)}(\Omega_T)}^2 \sum_i 1 \\ &\leq c 2^{2\|\mathbf{j}\|_\infty - 4\|\mathbf{j}\|_1} \|u\|_{H_{mix}^{(2,\infty),(2,\infty)}(\Omega_T)}^2 \end{aligned}$$

und Lemma 2.8 mit  $s = 1$  und  $t = 2$  zeigt (2.49).  $\square$

Verwenden wir anstelle der stückweise linearen, stetigen Zerlegung in der Zeit nun die stückweise lineare (aber unstetige) Zerlegung (2.34), so erhalten wir analog die folgenden Resultate.

**Satz 2.13.** *Sei  $V_j^\Omega$  eine Folge von (d-) linearen, affin äquivalenten Finite Element Räumen, deren zugrunde liegenden Triangulierungen regulär sind. Dann gibt es eine Konstante  $c > 0$ , so dass für den aus (2.12) mit Hilfe der (d-) linearen hierarchischen Basis im Ort und der aus der Zerlegung (2.34) in der Zeit konstruierten Raum-Zeit Dünngitterraum  $\tilde{V}_l^0$*

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{L^2(\Omega_T)} \leq c 2^{-2l} l \|u\|_{H_{mix}^{2,2}(\Omega_T)}, \quad (2.55)$$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H_{mix}^{1,0}(\Omega_T)} \leq c 2^{-l} \|u\|_{H_{mix}^{(2,\infty),(1,\infty)}(\Omega_T)}, \quad (2.56)$$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)} \leq c 2^{-2l} l \|u\|_{H_{mix}^{(2,2),(2,\infty)}(\Omega_T)} \quad (2.57)$$

*gilt.*

Da alle eindimensionalen Funktionen  $u^T \in H^{1,\infty}((0, T))$  bzw.  $u^T \in H^2((0, T))$  stetig sind, übertragen sich die Argumente in dem Beweis von Satz 2.12 auf den Fall, dass wir die Zerlegung (2.34) in der Raum-Zeit Dünngitterkonstruktion verwenden.

Betrachten wir nun den Fall, dass wir die stückweise konstante hierarchische Zerlegung (2.33) in der Zeit verwenden und schreiben abkürzend  $\psi_{j,i}^T := \psi_{j,0,i}^T$ . Mit Hilfe der klassischen Interpolationsfehlerabschätzung für die stückweise konstante Interpolation erhalten wir die Existenz einer Konstanten  $c > 0$ , so dass für jede Funktion  $u^T \in H^{1,\infty}((0, T))$ ,  $u^T = \sum_j w_j^T = \sum_j \sum_i w_{j,i}^T \psi_{j,i}^T$ ,

$$|w_{j,i}^T| \leq c \cdot 2^{-j} \|u^T\|_{H^{1,\infty}((0,T))} \quad (2.58)$$

gilt, wobei  $c > 0$  von  $u$ ,  $j$  und  $i$  unabhängig ist. Durch leichtes nachrechnen folgt, dass für die stückweise konstanten Basisfunktionen in der Zeit

$$\|\psi_{j_2,i_2}^T\|_{L^2((0,T))} = 2^{-j_2/2} \quad (2.59)$$

gültig ist. Damit erhalten wir das folgende Ergebnis über die Approximationseigenschaften des entsprechenden Raum-Zeit Dünngitters.

**Satz 2.14.** *Sei  $V_j^\Omega$  eine Folge von (d-) linearen, affin äquivalenten Finite Element Räumen, deren zugrunde liegenden Triangulierungen regulär sind. Dann gibt es eine Konstante  $c > 0$ , so dass für den aus (2.11) mit Hilfe der (d-) linearen hierarchischen*

Basis im Ort und der stückweise konstanten Zerlegung (2.33) konstruierten Raum-Zeit Dünngitterraum  $V_l^0$

$$\inf_{v \in V_l^0} \|u - v\|_{L^2(\Omega_T)} \leq c 2^{-2l} \|u\|_{H_{mix}^{2,1}(\Omega_T)}, \quad (2.60)$$

$$\inf_{v \in V_l^0} \|u - v\|_{H_{mix}^{1,0}(\Omega_T)} \leq c 2^{-l} \|u\|_{H_{mix}^{(2,2),(1,\infty)}(\Omega_T)}, \quad (2.61)$$

$$\inf_{v \in V_l^0} \|u - v\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)} \leq c 2^{-2l} \|u\|_{H_{mix}^{(2,2),(1,\infty)}(\Omega_T)} \quad (2.62)$$

gilt. Verwenden wir anstelle von  $V_l^0$  den Raum  $\tilde{V}_l^0$  mit stückweise konstanten Funktionen in der Zeit, so erhalten wir entsprechend

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{L^2(\Omega_T)} \leq c 2^{-l} \|u\|_{H_{mix}^{2,1}(\Omega_T)}, \quad (2.63)$$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H_{mix}^{1,0}(\Omega_T)} \leq c 2^{-l} \|u\|_{H_{mix}^{2,1}(\Omega_T)}, \quad (2.64)$$

$$\inf_{v \in \tilde{V}_l^0} \|u - v\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)} \leq c 2^{-l} \|u\|_{H_{mix}^{(2,2),(1,\infty)}(\Omega_T)}. \quad (2.65)$$

**Beweis.** Wir beschränken uns auf den Beweis der Abschätzungen bezüglich des Raumes  $V_l^0$  mit Hilfe von Lemma 2.9. Die Abschätzungen für  $\tilde{V}_l^0$  folgen analog aus Lemma 2.8. Für  $u^T \in H^{1,\infty}((0, T))$  gilt für die hierarchische Zerlegung  $u^T = \sum_j w_j^T$

$$\|w_j^T\|_{L^2((0,T))} \leq c 2^{-j} \|u^T\|_{H^{1,\infty}((0,T))}, \quad (2.66)$$

wobei  $c > 0$  von  $j$  und  $u$  unabhängig ist. Dieses Resultat wird wie (2.58) mit Hilfe von Interpolationsfehlerabschätzungen und der Dreiecksungleichung gezeigt. Über ein Tensorproduktargument, siehe [66], erhalten wir damit für  $u \in H_{mix}^{(2,2),(1,\infty)}(\Omega_T)$ ,  $u = \sum_j w_j$ ,

$$\|w_j\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)} \leq c 2^{-(2j_1+j_2)} \|u\|_{H_{mix}^{(2,2),(1,\infty)}(\Omega_T)}. \quad (2.67)$$

Lemma 2.9 liefert nun mit  $s = 0$ ,  $t = 1$ ,  $\|\cdot\|_\star = \|\cdot\|_{H^{(0,2),(0,\infty)}}$  und  $\|\cdot\|_\Delta = \|\cdot\|_{H_{mix}^{(2,2),(1,\infty)}(\Omega_T)}$  die Abschätzung (2.62).

Analog zu (2.67) ist für jedes  $u \in H_{mix}^{2,1}(\Omega_T)$ ,  $u = \sum_j w_j$  die Abschätzung

$$\|w_j\|_{L^2(\Omega_T)} \leq c 2^{-(2j_1+j_2)} \|u\|_{H_{mix}^{2,1}(\Omega_T)} \quad (2.68)$$

mit  $c > 0$  unabhängig von  $u$  und  $j$  gültig. Daraus folgt wiederum mit Lemma 2.9 für  $s = 0$ ,  $t = 1$ ,  $\|\cdot\|_\star = \|\cdot\|_{L^2(\Omega_T)}$  und  $\|\cdot\|_\Delta = \|\cdot\|_{H_{mix}^{2,1}(\Omega_T)}$  die Ungleichung (2.60).

Sei nun  $u \in H_{mix}^{(2,\infty),(1,\infty)}(\Omega_T)$ ,  $u = \sum_j \sum_i w_{j,i} \psi_{j,i}$ . Dann ergeben (2.43) und (2.58) über ein Tensorproduktargument

$$|w_{j,i}| \leq c \cdot 2^{-2j_1-j_2} \|u\|_{H_{mix}^{(2,\infty),(1,\infty)}(\Omega_T)}.$$

Des Weiteren gilt mit (2.46) und (2.59) für  $\psi_{\mathbf{j},\mathbf{i}} = \psi_{j_1,i_1}^\Omega \cdot \psi_{j_2,i_2}^T$

$$\|\psi_{\mathbf{j},\mathbf{i}}\|_{H_{mix}^{1,0}(\Omega_T)}^2 \leq c \cdot 2^{-(d-2) \cdot j_1 - j_2}.$$

Daraus folgt für  $w_{\mathbf{j}} = \sum_{\mathbf{i}} w_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}}$

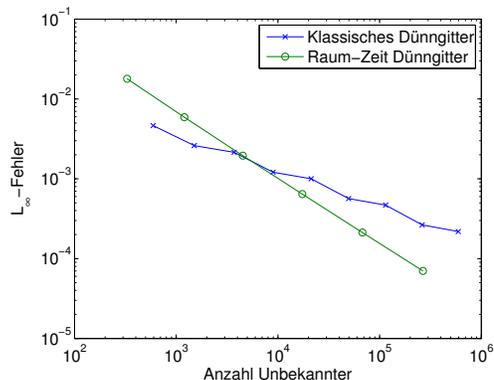
$$\begin{aligned} \|w_{\mathbf{j}}\|_{H_{mix}^{1,0}(\Omega_T)}^2 &\leq \sum_{\mathbf{i}} |w_{\mathbf{j},\mathbf{i}}|^2 \|\psi_{\mathbf{j},\mathbf{i}}\|_{H_{mix}^{1,0}(\Omega_T)}^2 \\ &\leq c \sum_{\mathbf{i}} 2^{-4j_1 - 2j_2} 2^{2j_1} 2^{-dj_1 - j_2} \|u\|_{H_{mix}^{(2,\infty),(1,\infty)}(\Omega_T)}^2 \\ &\leq c 2^{-4j_1 - 2j_2} 2^{\|(2j_1, j_2)\|_\infty} 2^{-dj_1 - j_2} \sum_{\mathbf{i}} \|u\|_{H_{mix}^{(2,\infty),(1,\infty)}(\Omega_T)}^2 \\ &\leq c 2^{\|(2j_1, j_2)\|_\infty - 2\|(2j_1, j_2)\|_1} \|u\|_{H_{mix}^{(2,\infty),(1,\infty)}(\Omega_T)}^2 \end{aligned}$$

und Lemma 2.9 mit  $s = 1/2$  und  $t = 1$  zeigt die Behauptung.  $\square$

Zum Abschluss dieses Kapitels ist es angebracht, den angesprochenen Vergleich zwischen den klassischen Dünngittern und den Raum-Zeit Dünngittern intensiver zu behandeln. Während wir die Verwendung der Raum-Zeit Dünngitter bisher vor allem mit der einfacheren Handhabbarkeit komplizierter Geometrien und der Behandlung von Differentialoperatoren mit variablen Koeffizienten motiviert haben, folgt aus den theoretischen Betrachtungen, dass wir unter schwächeren Glattheitsannahmen als bei klassischen dünnen Gittern die gleichen Approximationsraten erhalten. Gleichzeitig besitzen die Raum-Zeit Dünngitter zu vorgegebenem Level  $l$  wesentlich mehr Freiheitsgrade als die entsprechenden klassischen Dünngitter. Dies wirft die Frage auf, ob wir für manche Funktionen, die die von den klassischen Dünngittern benötigten Glattheitsannahmen nicht erfüllen, dennoch ein besseres Kosten-Nutzen Verhältnis für die Raum-Zeit Dünngitter erwarten können. Dabei verstehen wir unter dem Kosten-Nutzen Verhältnis die Relation von der Anzahl an verwendeten Freiheitsgraden  $N$  zu dem in einer vorgegebenen Norm  $\|\cdot\|$  gemessenen Fehler  $\|e_l\|$ ,  $e_l := u - u_l$ , zwischen der exakten Funktion  $u$  und der approximierenden Funktion  $u_l$  des entsprechenden Raumes. In engem Zusammenhang mit dem Kosten-Nutzen Verhältnis steht die so genannte  $\varepsilon$ -Komplexität [35, 107, 122] einer numerischen Methode. Hierbei bezeichnet die  $\varepsilon$ -Komplexität den Aufwand zur Berechnung einer Näherungslösung mit einer vorgegebenen Genauigkeit  $\varepsilon$ . Hierbei ist insbesondere der Exponent  $\alpha$  von Interesse, der für eine gegebene Anzahl an Unbekannten  $N$  die Genauigkeit durch  $\varepsilon(N) = O(N^{-\alpha})$  bestimmt. In diesem Sinne besitzt ein Verfahren ein besseres Kosten-Nutzen Verhältnis als ein anderes Verfahren, wenn der Wert  $\alpha$  für das erstere Verfahren echt größer als der Wert  $\alpha$  des zweiten Verfahrens ist.

Wir betrachten nun  $\Omega_T = (0, 1)^2 \times (0, 1)$  und messen den Interpolationsfehler der Funktion

$$u(x, t) = \exp(-t) |x_1 - x_2|^{8/5}$$



**Abbildung 2.5.**  $L^\infty$ - Interpolationsfehler für  $u(x, t) = \exp(-t) \cdot |x_1 - x_2|^{8/5}$  auf  $\Omega_T = (0, 1)^2 \times (0, 1)$  des klassischen, aus dem Tensorprodukt der eindimensionalen hierarchischen Basis konstruierten Dünngitters und des Raum-Zeit Dünngitters, bei dem im Ort die bilineare hierarchische Basis auf einer uniformen Verfeinerung und in der Zeit die lineare hierarchische Basis verwendet wurde.

auf dem klassischen Dünngitter und dem Raum-Zeit Dünngitter. Dabei gilt  $u \in H_{mix}^2(\Omega_T)$ , womit die Funktion die für die obige Theorie benötigten Glattheitsvoraussetzungen erfüllt. Gleichzeitig erfüllt  $u$  nicht die Glattheitsvoraussetzungen aus der Theorie der klassischen Dünngitter, d.h.  $u \notin H^2((0, 1)) \otimes H^2((0, 1)) \otimes H^2((0, 1))$ . Abbildung 2.5 zeigt nun den  $L^\infty$ -Interpolationsfehler in Abhängigkeit von der Anzahl der benötigten Freiheitsgrade für das klassische Dünngitter, das aus der linearen hierarchischen Basis konstruiert wurde, und dem aus der bilinearen hierarchischen Basis im Ort und der linearen hierarchischen Basis in der Zeit konstruierten Raum-Zeit Dünngitter  $\tilde{V}_l^0$ . Insgesamt ist die Steigung der zu dem Raum-Zeit Dünngitter gehörenden Kurve steiler und somit der Wert  $\alpha$  größer als für das klassische Dünngitter. Eine näherungsweise Bestimmung über die Steigungen der Regressionsgeraden durch die beiden Kurven führt zu den Werten  $\alpha = 0.83$  für das Raum-Zeit Dünngitter und  $\alpha = 0.44$  für das klassische Dünngitter. Somit liefert das Raum-Zeit Dünngitter, obwohl wir zunächst mit einem größeren Fehler pro Unbekannter starten, bereits für  $V_l^7$  einen wesentlich kleineren Fehler pro Unbekannter als das klassische Dünngitter. Dies zeigt, dass wir in dem hier betrachteten Fall tatsächlich für das Raum-Zeit Dünngitter ein wesentlich besseres Kosten-Nutzen Verhältnis als für das klassische Dünngitter erhalten.

Wie in diesem Kapitel deutlich geworden ist, ist es möglich, die Idee klassischer Dünngitter auf die natürliche Tensorproduktstruktur zwischen Raum und Zeit zu übertragen. Benötigt werden hierzu lediglich Multilevelzerlegungen von  $L^2((0, T))$  und  $L^2(\Omega)$ . Mit Hilfe dieser Konstruktion können wesentliche Nachteile klassischer Dünngitter, wie etwa die Behandlung komplizierter Geometrien und variabler Ko-

---

effizienten, behoben oder zumindest stark verbessert werden. Verwendet man im Ort und in der Zeit die aus einer Sequenz von ineinander geschachtelten Finite Element Räumen konstruierten hierarchischen Überschüßräume in der Raum-Zeit Dünngitterkonstruktion, so sind die Dimensionen der Raum-Zeit Dünngitterräume und der Finite Element Räume im Ort von (bis auf einen logarithmischen Faktor im Fall einer Ortsdimension) gleicher Ordnung. Gleichzeitig liefern die Raum-Zeit Dünngitterräume unter etwas stärkeren Glattheitsvoraussetzungen die gleichen Approximationsraten wie volle Gitter. Dies resultiert in der weiterführenden Fragestellung, ob diese Glattheitsannahmen an die Lösung parabolischer Probleme realistisch sind, was wir in dem nachfolgenden Kapitel untersuchen werden.



## Kapitel 3

# Parabolische Probleme und Regularität

Wie in dem vorhergehenden Kapitel deutlich wurde, liefern die Raum-Zeit Dünngitter unter etwas stärkeren Glattheitsvoraussetzungen die gleichen Interpolationsfehlerraten wie volle Gitter. Da wir an der Lösung parabolischer Gleichungen interessiert sind, stellt sich daher die Frage, unter welchen Voraussetzungen die Lösungen parabolischer Gleichungen die benötigte Regularität aufweisen. Dieser Fragestellung gehen wir in diesem Kapitel nach. Hierzu wiederholen wir zunächst klassische Regularitätsresultate für lineare parabolische Probleme, wobei wir der Darstellung in [121] folgen. Im Anschluss zeigen wir, dass die dort vorkommenden Räume zu den  $H_{mix}^{m,n}(\Omega_T)$  Räumen, die für die Interpolationseigenschaften der Raum-Zeit Dünngitterräume maßgeblich sind, identisch sind.

### 3.1 Notation

Sei  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , ein beschränktes Gebiet und  $\Omega_T := \Omega \times (0, T)$ . Wir bezeichnen mit  $C^{k,1}(\Omega)$  die Menge der  $k$ -mal stetig differenzierbaren Funktionen, deren  $k$ -te Ableitungen lipschitzstetig sind. Für  $\Omega, \tilde{\Omega} \subset \mathbb{R}^d$  nennen wir eine bijektive Abbildung

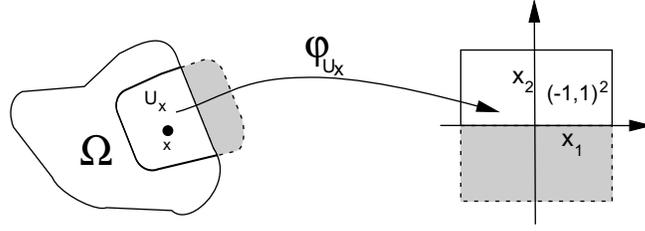
$$\varphi = (\varphi_1, \dots, \varphi_d) : \Omega \subset \mathbb{R}^d \rightarrow \tilde{\Omega} \subset \mathbb{R}^d$$

einen  $(k, 1)$ -Diffeomorphismus, wenn  $\varphi_i \in C^{k,1}(\Omega)$  für alle  $1 \leq i \leq d$  und für die Umkehrabbildung  $\varphi^{-1} = (\tilde{\varphi}_1, \dots, \tilde{\varphi}_d)$   $\tilde{\varphi}_i \in C^{k,1}(\tilde{\Omega})$  für alle  $1 \leq i \leq d$  gelten.

Das Gebiet  $\Omega$  heißt  $C^{k,1}$  glatt, falls es für jedes  $x \in \Omega$  eine Umgebung  $U_x$  gibt, so dass ein  $(k, 1)$ -Diffeomorphismus  $\varphi_{U_x} : U_x \rightarrow (-1, 1)^d$  mit

$$\begin{aligned}\varphi_{U_x}(U_x \cap \partial\Omega) &= \{x \in (-1, 1)^d \mid x_d = 0\}, \\ \varphi_{U_x}(U_x \cap \Omega) &= \{x \in (-1, 1)^d \mid 0 < x_d < 1\}, \\ \varphi_{U_x}(U_x \setminus \bar{\Omega}) &= \{x \in (-1, 1)^d \mid -1 < x_d < 0\}\end{aligned}$$

existiert, vgl. Abbildung 3.1. Für einen Hilbertraum  $V$  mit Norm  $\|\cdot\|_V$  bezeichne



**Abbildung 3.1.** Diffeomorphismus  $\varphi_{U_x}$ , der eine Umgebung  $U_x$  des Punktes  $x$  auf das Quadrat  $(-1, 1)^2$  abbildet.

$L^2((0, T); V)$  die Menge aller schwach meßbaren Funktionen, für die

$$\int_0^T \|u(t)\|_V^2 < \infty$$

gilt. Wir definieren weiter

$$W((0, T); H^1(\Omega)) := \{f \mid f \in L^2((0, T); H^1(\Omega)) \text{ und } \partial_t f \in L^2((0, T); H^{-1})\} \quad (3.1)$$

und versehen den Raum mit dem Skalarprodukt

$$(f, g) := \int_0^T ((f(t), g(t))_{H^1(\Omega)} + (\partial_t f(t), \partial_t g(t))_{H^{-1}(\Omega)}) dt.$$

Hierbei ist das Skalarprodukt des Dualraumes  $H^{-1}(\Omega)$  von  $H^1(\Omega)$  durch

$$(v, w)_{H^{-1}(\Omega)} := (\theta(v), \theta(w))_{H^1(\Omega)}$$

definiert, wobei die Abbildung  $\theta : H^{-1}(\Omega) \rightarrow H^1(\Omega)$  jedem  $v \in H^{-1}(\Omega)$  das Element  $\theta(v) \in H^1(\Omega)$  zuordnet, so dass

$$(\theta(v), u)_{H^1(\Omega)} = v(u) \text{ für alle } u \in H^1(\Omega)$$

gilt.

Wir führen nun die für die Regularitätstheorie parabolischer Probleme wichtigen  $W^k((0, T); H^m(\Omega))$ -Räume ein.

**Definition 3.1.** Für  $k, m \in \mathbb{N}$  sei der Raum  $W^k((0, T); H^m(\Omega))$  die Menge aller Funktionen  $u : (0, T) \rightarrow H^m(\Omega)$  für die

$$\frac{d^j u}{dt^j} \in L^2((0, T); H^m(\Omega)) \text{ für alle } 0 \leq j \leq k \quad (3.2)$$

gilt. Dabei sind die Ableitungen im Distributionensinne zu verstehen. Die Norm auf  $W^k((0, T); H^m(\Omega))$  ist definiert durch

$$\|u\|_{W^k((0, T); H^m(\Omega))}^2 := \sum_{j=0}^k \int_0^T \left\| \frac{d^j u}{dt^j} \right\|_{H^m(\Omega)}^2 dt. \quad (3.3)$$

Für  $k \geq 0$  ist  $W^k((0, T); \Omega)$  ein Hilbertraum.

### 3.2 Regularität der Lösung parabolischer Differentialgleichungen

Im folgenden sei  $L$  ein linearer elliptischer Differentialoperator zweiter Ordnung, d.h.

$$Lu = -\nabla \cdot (A\nabla u) + b \cdot \nabla u + cu \quad (3.4)$$

mit  $A = (a_{ij})_{i,j=1}^d : \Omega_T \rightarrow \mathbb{R}^{d \times d}$ ,  $b = (b_i)_{i=1}^d : \Omega_T \rightarrow \mathbb{R}^d$  und  $c : \Omega_T \rightarrow \mathbb{R}$ . Für eine gegebene rechte Seite  $f : \Omega_T \rightarrow \mathbb{R}$ , eine Anfangsbedingung  $u_0 : \Omega \rightarrow \mathbb{R}$  und eine Randbedingung  $g : \partial\Omega \times (0, T]$  suchen wir eine Funktion  $u : \Omega_T \rightarrow \mathbb{R}$ , die die Gleichungen

$$\partial_t u + Lu = f \text{ in } \Omega_T, \quad (3.5)$$

$$u = g \text{ auf } \partial\Omega \times (0, T], \quad (3.6)$$

$$u(\cdot, 0) = u_0(\cdot) \text{ auf } \Omega \quad (3.7)$$

erfüllt. Wenn wir nun die Gleichung (3.5) mit einer Testfunktion  $\varphi \in H_0^1(\Omega)$  multiplizieren und partiell integrieren, so erhalten wir die zugehörige schwache Formulierung des Problems,

$$\begin{aligned} (\partial_t u, \varphi)_{L^2(\Omega)} + a(t; u, \varphi) &= (f, \varphi)_0 \text{ für alle } \varphi \in H_0^1(\Omega), t \in (0, T], \\ u(0, \cdot) &= u_0(\cdot) \text{ f.ü. in } \Omega, \\ u(t, x) &= g(t, x) \text{ für alle } t \in (0, T) \text{ f.ü. in } \partial\Omega, \end{aligned} \quad (3.8)$$

wobei

$$a(t; u, v) := \int_{\Omega} (\nabla u)^T A \nabla v + b v \cdot \nabla u + c u v \, dx. \quad (3.9)$$

Wir suchen somit eine Funktion  $u$ , die jedem  $t \in (0, T]$  eine Funktion  $u(t, \cdot) \in H^1(\Omega)$  zuordnet und die im Distributionensinne in  $t$  differenzierbar ist. Aus dieser Betrachtungsweise heraus nehmen Funktionen mit Werten in dem Hilbertraum  $H^1(\Omega)$ , deren Ableitungen in  $H^{-1}(\Omega)$  liegen, also den Funktionen, die in dem Raum  $W((0, T); H^1(\Omega))$  enthalten sind, eine bedeutende Rolle bei der Behandlung parabolischer Probleme ein.

Um die Existenz einer Lösung von (3.8) zu garantieren, stellen wir die folgenden Bedingungen an die Bilinearform  $a(t; \cdot, \cdot)$

A1) Für  $\varphi, \psi \in H^1(\Omega)$  fest sei  $a(t; \varphi, \psi)$  meßbar auf  $[0, T]$ .

A2) Es existiere eine Konstante  $C > 0$  unabhängig von  $t \in [0, T]$ , so dass

$$|a(t; \varphi, \psi)| \leq C \cdot \|\varphi\|_{H^1(\Omega)} \|\psi\|_{H^1(\Omega)}, \text{ für alle } t \in [0, T], \varphi, \psi \in H^1(\Omega) \quad (3.10)$$

gilt.

A3) Es existiere eine Konstante  $c > 0$  unabhängig von  $t \in [0, T]$ , so dass

$$c \cdot \|\varphi\|_{H^1(\Omega)}^2 \leq a(t; \varphi, \varphi) \text{ für alle } \varphi \in H_0^1(\Omega) \quad (3.11)$$

erfüllt ist.

Unter diesen Voraussetzungen an  $a(t; \cdot, \cdot)$  läßt sich nun die Lösbarkeit des Problems (3.8) zeigen.

**Satz 3.2.** *Die Randfunktion  $g$  lasse sich derart in das Innere des Gebietes  $\Omega$  fortsetzen, dass  $g \in W((0, T); H^1(\Omega))$  ist. Es gelten A1, A2 und A3 für die Bilinearform  $a(\cdot; \cdot, \cdot)$ . Dann besitzt das Problem (3.8) für jedes  $f \in L^2((0, T), H^{-1}(\Omega))$  und  $u_0 \in H^1(\Omega)$  eine eindeutige Lösung  $u \in W(0, T; H^1(\Omega))$ .*

Für den Beweis verweisen wir auf [121] und untersuchen nun die Regularität der Lösung  $u$  des Problems (3.8). Dabei ist, wie auch bei elliptischen Problemen, die Glattheit der Lösung von der Glattheit des Gebietes  $\Omega$  abhängig, so dass Bedingungen an das Gebiet unumgänglich sind. Im Folgenden nehmen wir daher an, dass  $\Omega$  beschränkt und von der Glattheit  $C^{1+k,1}$  ist. Zusätzlich gelte für die Koeffizientenfunktionen des elliptischen Operators  $L(t)$

$$K1) \ a_{ij}(\cdot, t), b_i(\cdot, t), c(\cdot, t) \in C^k(\bar{\Omega}) \text{ für alle } t \in [0, T],$$

$$K2) \ \partial_t^l a_{ij}, \partial_t^l b_i, \partial_t^l c \in C(\bar{\Omega} \times [0, T]) \text{ für } 0 \leq l \leq k.$$

Wir erhalten dann das folgende Resultat bzgl. der Differenzierbarkeit der Lösung  $u$  von (3.8) nach  $t$ .

**Satz 3.3.** *Für festes  $k \in \mathbb{N}$  seien die Voraussetzungen A1, A2, A3, K1 und K2 erfüllt. Weiterhin sei die Randfunktion  $g$  in das Innere von  $\Omega$  fortsetzbar mit*

$$g \in W^k((0, T); H^1(\Omega)) \text{ und } \partial_t^{k+1} g \in L^2((0, T); H^{-1}(\Omega)).$$

*Zusätzlich seien für die Rand- und Anfangsbedingungen die Kompatibilitätsbedingungen*

$$\partial_t^l u(\cdot, 0) - \partial_t^l g(\cdot, 0) \in H_0^1(\Omega), \text{ für alle } 0 \leq l \leq k-1, \quad (3.12)$$

$$\partial_t^k u(\cdot, 0) - \partial_t^k g(\cdot, 0) \in L^2(\Omega), \quad (3.13)$$

*erfüllt. Dann besitzt das Problem (3.8) für jedes  $f \in W^k((0, T); H^{-1}(\Omega))$  genau eine Lösung  $u$  und es gilt*

$$u \in W^k((0, T); H^1(\Omega)) \text{ und } \partial_t^{k+1} u \in L^2((0, T); H^{-1}(\Omega)). \quad (3.14)$$

Die Kompatibilitätsbedingungen (3.12) und (3.13) sind nur die Kurzschreibweise für

$$u(\cdot, 0) = u_0(\cdot), \quad \partial_t u(\cdot, 0) = f(\cdot, 0) - L(\cdot, 0)u_0, \quad \text{u.s.w.},$$

die wir durch Differentiation und Auswerten an der Stelle  $t = 0$  aus (3.5) erhalten.

Um nun Aussagen über die Glattheit der Lösung bzgl. des Ortes zu erhalten, erhöhen wir die Differenzierbarkeitsvoraussetzungen bzgl. der Ortsvariablen  $x$ . Wir nehmen daher für  $q \in \mathbb{N}$  an, dass die Bedingungen des vorherigen Satzes nun für  $k + q$  gelten. Wir erhalten dann aus [121] die folgende Regularitätsaussage.

**Satz 3.4.** *Für  $k, q \in \mathbb{N}$  gelte die Bedingung K1 mit  $k + q$  anstelle von  $k$ . Das Gebiet  $\Omega$  sei beschränkt und  $C^{1+q+k,1}$  glatt. Für die Umkehrabbildung  $L^{-1}(t) : H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$  gelte*

$$\|\partial_t^n L^{-1}(t)f\|_{H^{2+q}(\Omega)} \leq c_q \|f\|_{H^q(\Omega)} \quad (3.15)$$

mit  $c_q$  unabhängig von  $t$  für alle  $f \in H^{2+q}(\Omega) \cap H_0^1(\Omega)$  und  $0 \leq n \leq k - 1$ . Des Weiteren seien die Voraussetzungen des vorigen Satzes erfüllt und es sei

$$f \in W^{k-1}((0, T); H^q(\Omega)) \quad \text{und} \quad \partial_t^k f \in L^2((0, T); H^{-1}(\Omega)). \quad (3.16)$$

Wir definieren rekursiv die Zahlen  $m_j$ ,  $1 \leq j \leq k$ , durch

$$m_1 = 2 + \min(1, q), \quad m_{j+1} = 2 + \min(q, m_j).$$

Für  $\tilde{q} := \max_{1 \leq j \leq k} m_j$  sei die gegebene Randfunktion  $g$  in das Innere von  $\Omega$  mit

$$g \in W^k((0, T); H^{2+\tilde{q}}(\Omega)) \quad \text{und} \quad \partial_t^{k+1} g \in L^2((0, T); H^{-1}(\Omega)) \quad (3.17)$$

fortsetzbar. Dann besitzt das Problem (3.8) genau eine Lösung  $u$  und es gilt

$$u \in W^{k-j}((0, T); H^{2+\min(q, m_j)}) \quad \text{für alle } 1 \leq j \leq k. \quad (3.18)$$

Die Bedingung (3.15) an  $L(t)$  ist trivialerweise erfüllt, wenn die Koeffizientenfunktionen des elliptischen Operators nicht von der Zeit abhängen.

Die beiden Sätze zeigen, dass es für fest gewählte  $k, m \in \mathbb{N}$  stets möglich ist, geeignete Bedingungen an die rechte Seite, die Randbedingung und die Anfangsbedingung, die Glattheit des Gebietes  $\Omega$  und die Koeffizienten des elliptischen Operators  $L(t)$  zu stellen, so dass für die Lösung  $u$  der schwachen Formulierung

$$u \in W^k((0, T); H^m(\Omega)) \quad (3.19)$$

gilt. Wie wir im vorigen Kapitel bereits diskutiert haben, spielen die Räume  $H_{mix}^{m,k}(\Omega_T)$  dominierender gemischter Glattheit eine wichtige Rolle in dem Kontext der Raum-Zeit Dünngitterräume. Präziser formuliert können wir Funktionen aus diesen Räumen

mit Hilfe der Raum-Zeit Dünngitterräume mit der gleichen (bis auf einen logarithmischen Faktor) Approximationsordnung wie mit den Vollgitterräumen approximieren. Deshalb wäre für die Diskretisierung parabolischer Probleme mit Raum-Zeit Dünngittern eine Regularitätstheorie bzgl. dieser Räume von Bedeutung. Wie der folgende Satz allerdings zeigt, entsprechen die Räume  $H_{mix}^{m,k}(\Omega_T)$  den Räumen  $W^k((0, T); H^m(\Omega))$ .

**Satz 3.5.** Für  $k, m \in \mathbb{N}$  gilt

$$W^k((0, T); H^m(\Omega)) = H_{mix}^{m,k}(\Omega_T). \quad (3.20)$$

**Beweis.** Sei  $u \in W^k((0, T); H^m(\Omega))$  und  $\{e_n(\cdot)\}$  eine orthonormale Basis von  $H^m(\Omega)$ . Wir definieren

$$u_n(t) := (u(t), e_n)_{H^m(\Omega)}, \quad (3.21)$$

und für festes  $t \in (0, T)$  gilt  $u(t)(\cdot) = \sum_n u_n(t) e_n(\cdot)$ . Des Weiteren erhalten wir für die Folge  $w_j := \sum_{n=1}^j u_n(t) \cdot e_n$

$$\lim_{j \rightarrow \infty} w_j = u \text{ in } W^k((0, T); H^m(\Omega))$$

und somit ist  $w_j$  insbesondere eine Cauchy-Folge in  $W^k((0, T); H^m(\Omega))$ . Wir haben außerdem

$$\int_0^T |u_n(t)|^2 dt \leq \int_0^T \|u(t)\|_{H^m(\Omega)}^2 dt < \infty,$$

weshalb  $u_n(t) \in L^2((0, T))$  gilt. Für  $1 \leq l \leq k$  ist

$$\partial_t^l u_n(t) = \left( \frac{d^l u}{dt^l}(t), e_n \right)_{H^m(\Omega)} \in L^2(0, T). \quad (3.22)$$

Damit gilt  $u_n(t) \in H^k(0, T)$  und somit  $w_j \in H^m(\Omega) \otimes H^k((0, T))$ . Da nun  $w_j$  eine Cauchy-Folge in  $W^k((0, T); H^m(\Omega))$  ist und  $\|w_j - w_l\|_{W^k((0, T); H^m(\Omega))} = \|w_j - w_l\|_{H_{mix}^{m,k}(\Omega_T)}$ , ist  $w_j$  auch eine Cauchy-Folge in  $H_{mix}^{m,k}(\Omega_T)$ . Aus der Vollständigkeit von  $H_{mix}^{m,k}(\Omega)$  folgt dann  $u = \lim_{t \rightarrow \infty} w_j \in H_{mix}^{m,k}(\Omega)$ .

Sei nun  $u \in H_{mix}^{m,k}(\Omega)$ . Dann gibt es eine Folge  $u_n^T(t) \in H^k((0, T))$ , eine Folge  $u_n^\Omega(x) \in H^m(\Omega)$  und eine Folge von Koeffizienten  $\lambda_n$ , so dass

$$u_n := \sum_{i=1}^n \lambda_i u_i^\Omega(x) \otimes u_i^T(t)$$

gegen  $u$  in  $H_{mix}^{m,k}(\Omega)$  konvergiert. Somit ist  $u_n$  Cauchy Folge in  $H_{mix}^{m,k}(\Omega)$  und insbesondere auch eine Cauchy-Folge in  $W^k((0, T); H^m(\Omega))$ . Da  $W^k((0, T); H^m(\Omega))$  vollständig ist, folgt damit  $u = \lim_n u_n \in W^k((0, T); H^m(\Omega))$ .  $\square$

Satz 3.4 liefert im Hinblick auf Satz 3.5 somit Regularitätsaussagen bzgl. der  $H_{mix}^{m,k}(\Omega_T)$  Räume. Nehmen wir an, dass Gebiet  $\Omega$  sei von der Glattheit  $C^{4,1}$  und es gelte

$$\begin{aligned} f &\in W^2((0, T); L^2(\Omega)), \partial_t^3 f \in L^2((0, T); H^{-1}(\Omega)), \\ g &\in W^3((0, T); H^2(\Omega)), \partial_t^4 g \in L^2((0, T); H^{-1}(\Omega)), \end{aligned}$$

und die Anfangs- und Randbedingungen seien ausreichend kompatibel. Ist des Weiteren der elliptische Operator  $L$  zeitunabhängig, so folgt aus Satz 3.4 für die Lösung  $u$  des Problems  $u \in H_{mix}^2(\Omega_T)$ . Diese Regularität benötigen wir jedoch gerade in den Sätzen 2.12 und 2.13 zur Abschätzung des Approximationsfehlers bzgl. der  $L^2$ -Norm. Schwächen wir die obigen Glattheitsvoraussetzungen an das Gebiet, die rechte Seite und die Randbedingungen weiter ab, so dass  $\Omega$  von der Glattheit  $C^{3,1}$  ist und

$$\begin{aligned} f &\in W^1((0, T); L^2(\Omega)), \partial_t^2 f \in L^2((0, T); H^{-1}(\Omega)), \\ g &\in W^2((0, T); H^2(\Omega)), \partial_t^3 g \in L^2((0, T); H^{-1}(\Omega)), \end{aligned}$$

gilt, so erhalten wir für die Lösung  $u \in H_{mix}^{2,1}(\Omega_T)$ . Diese Glattheitsvoraussetzung haben wir in Satz 2.14 für die Approximationsfehlerabschätzungen der aus der hierarchischen Basis im Ort und der stückweise konstanten hierarchischen Basis in der Zeit konstruierten Raum-Zeit Dünngitterräume  $V_l^0$  und  $\tilde{V}_l^0$  bzgl. der  $L^2$ -Norm gefordert.

Diese Beispiele zeigen, dass die Glattheitsvoraussetzungen in Kapitel 2 keineswegs willkürlich gewählt sind, sondern für die Lösungen parabolischer Probleme durchaus zutreffen.

In diesem Kapitel haben wir einige klassische Resultate der Regularitätstheorie für lineare parabolische Probleme wiederholt. Es stellt sich dabei als ein wesentliches Resultat heraus, dass die Räume, in denen die Lösungen der parabolischen Gleichungen nach den Regularitätsresultaten enthalten sind, gerade den in (2.6) definierten Räumen dominierender gemischter Raum-Zeit Ableitungen entsprechen. Dieses Ergebnis zeigt, dass die für die Theorie der Raum-Zeit Dünngitter benötigten Glattheitsannahmen tatsächlich für die Lösungen von parabolischen Problemen zu erwarten sind. Somit haben wir eine weitere Motivation dafür, auf den Raum-Zeit Dünngittern basierende Diskretisierungsverfahren zu entwickeln, was wir im nun folgenden Kapitel vertiefen werden.



## Kapitel 4

# Diskretisierung

In diesem Kapitel diskutieren wir die Diskretisierung parabolischer Probleme mit Hilfe der Raum-Zeit Dünngitter. Hierzu verwenden wir als Ausgangspunkt die schwache Formulierung (3.8) des parabolischen Problems. Zur Diskretisierung von parabolischen Problemen existieren im wesentlichen drei Methoden, die sich in der Reihenfolge der Diskretisierung von Raum und Zeit unterscheiden.

Bei der (vertikalen) Linienmethode wird in einem ersten Schritt der Ort, beispielsweise mittels Finiter Elemente, diskretisiert. Das dadurch entstandene System gewöhnlicher Differentialgleichungen kann anschließend mit herkömmlichen Zeitintegratoren, wie z.B. dem Runge-Kutta Verfahren, gelöst werden. Hierbei wird der Fehler im Ort mit Hilfe von Fehlerschätzern für stationäre Probleme [2, 17] kontrolliert.

Bei der Rothe-Methode interpretiert man das kontinuierliche Problem hingegen als gewöhnliche Differentialgleichung in einem Hilbertraum, wobei die zeitliche Komponente durch geeignete Zeitintegratoren diskretisiert wird. Die in jedem Zeitschritt durchzuführende Ortsdiskretisierung wird dabei als Störung der Zeitintegration aufgefasst. Somit können Kriterien entwickelt werden, mit denen entschieden werden kann, wie der Fehler der Ortsdiskretisierung beschränkt sein muß, damit sich der Fehler mit der Zeitdiskretisierung nicht zu sehr aufschaukelt und beschränkt bleibt. Zur Kontrolle des Ortsfehlers werden, wie schon zuvor bei der (vertikalen) Linienmethode, klassische Fehlerschätzer für stationäre Probleme herangezogen [83, 19, 21].

Ein dritter Ansatz, bei dem Ort und Zeit gleichzeitig diskretisiert werden, ist die Discontinuous-Galerkin Methode [49, 50]. Hierbei kommen geeignete Raum-Zeit Fehlerschätzer zur Kontrolle des Gesamtfehlers zur Anwendung.

Da sowohl bei der Rothe-Methode als auch bei der (vertikalen) Linienmethode Ort und Zeit getrennt behandelt werden, erscheinen diese Verfahren auf den ersten Blick als wenig geeignet, um auf den Raum-Zeit Dünngitterfall übertragen zu werden. Betrachten wir beispielsweise die Linienmethode, so ist zu verschiedenen Zeitpunkten jeweils ein stationäres Problem zu lösen, das in der Regel die Lösung

an bestimmten vorherigen Zeitpunkten aber keine Lösung an späteren Zeitpunkten benötigt. Zur Berechnung einer Raum-Zeit Dünngitterinterpolierenden einer Funktion an einem festen Punkt im Ort zu einem festen Zeitpunkt benötigen wir jedoch bereits Werte der zu interpolierenden Funktion an dem gleichen Ortspunkt zu einem späteren Zeitpunkt. Dieses Beispiel illustriert, dass Raum-Zeit Dünngitter wenig geeignet für ein schrittweises Vorgehen in der Zeit sind. Dennoch lassen sich herkömmliche Zeitschrittverfahren auf den Raum-Zeit Dünngitterfall übertragen, indem die Zeitdiskretisierung als Petrov-Galerkin Verfahren interpretiert wird. Da hierbei der Ansatzraum meist aus in der Zeit stetigen Funktionen besteht, wird diese Vorgehensweise auch als Continuous Galerkin Verfahren bezeichnet [48, 54]. In [5] wurde bereits das Crank-Nicolson und das implizite Eulerverfahren in diesem Sinne als Petrov-Galerkin Verfahren interpretiert, wobei die Verwendung von klassischen Dünngitterräumen als Ansatz- und Testraum zu einer Dünngitterdiskretisierung führt. Wir diskutieren diese Vorgehensweise im Kontext der Raum-Zeit Dünngitter in Abschnitt 4.1, wobei eine Stabilitätsabschätzung die eindeutige Lösbarkeit des diskreten Problems zeigt.

Da bei dem Discontinuous-Galerkin Verfahren Raum und Zeit gleichzeitig diskretisiert werden, ist das Discontinuous-Galerkin Verfahren direkt auf den Raum-Zeit Dünngitterfall übertragbar. Dieses Vorgehen beschreiben wir in Abschnitt 4.2 und untersuchen dort auch das Konvergenzverhalten der Discontinuous-Galerkin Diskretisierung im Fall von stückweise konstanten oder stückweise linearen Funktionen in der Zeit, wobei wir eine a-priori Fehlerabschätzung für die Dünngitter Discontinuous-Galerkin Diskretisierung erhalten.

Die durch die Diskretisierung mit den Raum-Zeit Dünngittern und dem Discontinuous Galerkin oder dem Crank-Nicolson Verfahren anfallenden linearen Gleichungssysteme sind in der Regel nicht dünn besetzt. Allerdings kann, wie schon bei klassischen Dünngittern, das Matrix-Vektor Produkt mit einem zur Zahl der Freiheitsgrade proportionalen Aufwand berechnet werden. Dies geschieht mit Hilfe des *Unidirektionalen Prinzips*, das erstmals in [5] für klassische Dünngitter beschrieben wurde, und das wir in Abschnitt 4.3 auf die Raum-Zeit Dünngitter verallgemeinern. Da wir hier mit beliebigen linearen bzw.  $d$ -linearen hierarchischen Basen im Ort arbeiten, müssen wir dafür insbesondere die benötigten Teilalgorithmen *Top Down* und *Bottom Up* anpassen.

Schliesslich untersuchen wir in Abschnitt 4.4 anhand einiger numerischer Ergebnisse das Konvergenzverhalten des Crank-Nicolson und des Discontinuous Galerkin Verfahrens mit stückweise konstanten bzw. stückweise linearen Funktionen in der Zeit.

## 4.1 Das Crank-Nicolson Verfahren

Ersetzen wir in der schwachen Formulierung

$$\begin{aligned} (\partial_t u, \varphi)_{L^2(\Omega)} + a(t; u, \varphi) &= (f, \varphi)_0 \text{ für alle } \varphi \in H_0^1(\Omega), t \in (0, T], \\ u(0, \cdot) &= u_0(\cdot), \text{ f.ü. in } \Omega, \\ u(t, x) &= g(t, x) \text{ für alle } t \in (0, T) \text{ f.ü. in } \partial\Omega \end{aligned}$$

des parabolischen Problems die punktweise in  $t$  geforderte Gleichheit durch Testen mit beliebigen  $H_0^1(\Omega) \otimes L^2((0, T))$ -Funktionen  $\varphi$ , so erhalten wir die äquivalente Formulierung

$$\begin{aligned} \int_0^T \left( \frac{d}{dt} u, \varphi \right)_{L^2(\Omega)} + a(t; u, \varphi) dt &= \int_0^T (f, \varphi)_0 dt \text{ für alle } \varphi \in H_0^1(\Omega) \otimes L^2((0, T)) \\ u(0, \cdot) &= u_0(\cdot), \text{ f. ü. in } \Omega, \\ u(x, t) &= g(x, t) \text{ für alle } t \in (0, T] \text{ f. ü. in } \partial\Omega. \end{aligned} \quad (4.1)$$

Wir wählen die Formulierung (4.1) als Ausgangspunkt für eine Petrov-Galerkin Diskretisierung. Dazu beschäftigen wir uns nun mit der Wahl zweier geeigneter endlich-dimensionaler Teilräume von  $H_0^1(\Omega) \otimes L^2((0, T))$ , die wir als Ansatz- bzw. Testraum verwenden können. Hierzu betrachten wir zunächst die schwache Formulierung einer gewöhnlichen Differentialgleichung, d.h. wir suchen eine Funktion  $u$ , so dass

$$\int_0^T \frac{d}{dt} u \cdot v dt = \int_0^T f \cdot v dt \text{ für alle } v \in L^2((0, T)), \quad (4.2)$$

$$u(0) = u_0, \quad (4.3)$$

gilt. Im folgenden sei das Zeitgebiet  $(0, T)$  durch die  $N + 1$  Zeitpunkte  $0 = t_0 < t_1 < \dots < t_N = T$  unterteilt und wir definieren  $I_k := (t_{k-1}, t_k)$  für  $1 \leq k \leq N$ . Seien  $V^{Ansatz}$  und  $V^{Test}$  stetige Funktionenräume auf dem Einheitsintervall  $(0, 1)$  wobei für ein fest gewähltes  $r > 0$

$$\dim(V^{Ansatz}) = r + 1 \text{ und } \dim(V^{Test}) = r$$

gelten soll. Bezeichnet  $\pi_k : I_k \rightarrow (0, 1)$  die affin lineare Transformation auf das Einheitsintervall, so definieren wir damit die diskreten Räume

$$\begin{aligned} V_N^{Ansatz} &:= \{ \varphi \in C(0, T) \mid (\varphi|_{I_k}) \circ \pi_k \in V^{Ansatz} \text{ für alle } 1 \leq k \leq N \}, \\ V_N^{Test} &:= \{ \varphi \in L^2(0, T) \mid (\varphi|_{I_k}) \circ \pi_k \in V^{Test} \text{ für alle } 1 \leq k \leq N \}. \end{aligned}$$

Verwenden wir  $V_N^{Ansatz}$  als Ansatzraum und  $V_N^{Test}$  als Testraum einer Petrov-Galerkin Diskretisierung von (4.2), wobei wir  $f$  durch den Interpolanten in dem Ansatzraum ersetzen, so erhalten wir damit je nach Wahl der Räume  $V^{Ansatz}$  und  $V^{Test}$  bekannte

Zeitdiskretisierungsverfahren. Aufgrund der Forderung, dass die Funktionen des Ansatzraumes stetig sein sollen, werden diese Methoden auch als Continuous Galerkin Verfahren bezeichnet, vgl. [11, 48, 54]. Im übrigen erklärt sich mit der Stetigkeitsforderung auch die Bedingung, dass die Dimension von  $V^{Ansatz}$  um genau eins größer sein soll als die Dimension des Testraumes  $V^{Test}$ . Damit erhalten wir, dass der Ansatzraum  $V_N^{Ansatz}$  genau einen Freiheitsgrad mehr als der Testraum  $V_N^{Test}$  besitzt, d.h.

$$\dim(V_N^{Ansatz}) = N \cdot r + 1 = \dim(V_N^{Test}) + 1,$$

und mit der Forderung, dass die Anfangsbedingung erfüllt sein soll, bleiben genauso viele Freiheitsgrade, wie in dem Testraum vorhanden sind, übrig.

Wählen wir nun beispielsweise zu fest gewähltem Parameter  $\theta$

$$V^{Test} := \text{span}\{1\} \text{ und } V^{Ansatz} := \text{span}\{1, \phi(t) = t + 3(2\theta - 1)t(1 - t)\},$$

so erhalten wir das  $\theta$ -Schema. Als Spezialfall führt die Wahl  $\theta = 1$  zu dem impliziten Eulerverfahren, vgl. [11]. Wählen wir hingegen für vorher fest gewähltes  $r > 1$  den Ansatzraum  $V^{Ansatz}$  als den Raum aller Polynome vom Grad  $r$ ,

$$V^{Ansatz} := \text{span}\{x^i \mid 1 \leq i \leq r\}$$

und den Testraum  $V^{Test}$  als den Raum aller Polynome vom Grad  $r - 1$ ,

$$V^{Test} := \text{span}\{x^i \mid 1 \leq i \leq r - 1\},$$

so erhalten wir das  $(r, r)$ -Padé Schema der Ordnung  $2r$ . Für  $r = 1$  resultiert insbesondere das Crank-Nicolson Verfahren

$$\frac{u(t_k) - u(t_{k-1})}{t_k - t_{k-1}} = \frac{f(t_k) + f(t_{k-1})}{2} \text{ für alle } 1 \leq k \leq N.$$

Für die Spezialfälle des impliziten Eulerverfahrens und des Crank-Nicolson Verfahrens ist die Interpretation als Petrov-Galerkin Verfahren bereits in [5] unabhängig von den früher zitierten Arbeiten angeführt und zur Dünngitterdiskretisierung parabolischer Probleme umgesetzt worden.

Wir konzentrieren uns in diesem Abschnitt auf die Crank-Nicolson Diskretisierung. Mit der obigen Interpretation des Crank-Nicolson Verfahrens zur Lösung der gewöhnlichen Differentialgleichung erhalten wir unmittelbar die Interpretation der Crank-Nicolson Diskretisierung des parabolischen Problems mit Finiten Elementen im Ort als Petrov-Galerkin Verfahren mit der durch (4.1) gegebenen Bilinearform. Hierbei ist der Ansatzraum als das Tensorprodukt des zu dem Crank-Nicolson Verfahren gehörigen Ansatzraumes in der Zeit und dem Finite Element Raum zu wählen. Analog ist der Testraum als das Tensorprodukt des Testraumes der Crank-Nicolson Diskretisierung in der Zeit und dem Finite Element Raum gegeben.

Es ist nun naheliegend, anstelle der Vollgitterräume die entsprechenden Dünn-  
gitterräume zu verwenden, d.h. als Ansatzraum verwenden wir den mit Hilfe der  
hierarchischen Basis im Ort und der linearen hierarchischen Basis in der Zeit kon-  
struierten Raum  $\tilde{V}_l^0$ , den wir im folgenden mit  $\tilde{V}_l^{0,Ansatz}$  bezeichnen. Der Testraum  
 $\tilde{V}_l^{0,Test}$  ist dann als der aus der hierarchischen Basis im Ort und der stückweise  
konstanten Zerlegung in der Zeit konstruierte Raum  $\tilde{V}_l^0$  zu wählen. Wir erhalten  
damit das diskretisierte Problem: Finde  $u_l \in \tilde{V}_l^{0,Ansatz}$ , so dass  $u_l$  die Anfangs- und  
Randbedingungen erfüllt und

$$B(u_l, \varphi) = \int_0^T (f, \varphi)_0 dt \text{ für alle } \varphi \in \tilde{V}_l^{0,Test}, \quad (4.4)$$

mit

$$B(u_l, \varphi) := \int_0^T \left( \frac{d}{dt} u_l, \varphi \right)_{L^2(\Omega)} + a(t; u_l, \varphi) dt$$

gilt. Wir nehmen zur Vereinfachung der Notation  $T = 1$  an und wählen im Folgenden  
für festes  $l$  die diskreten Zeitpunkte  $t_k := k \cdot 2^{-l}$ . Wir erhalten dann die folgende  
Stabilitätsaussage.

**Lemma 4.1.** *Die Bilinearform  $a(t; \cdot, \cdot)$  sei unabhängig von  $t$ , d.h.  $a(t; \cdot, \cdot) = a(\cdot, \cdot)$   
und es gelte  $a(v(\cdot, T), v(\cdot, T)) \geq 0$  für alle  $v \in \tilde{V}_l^{0,Ansatz}$ . Ist  $u_0 = 0$ , so gilt für die  
Lösung  $u_l$  der Dünngitter-Crank-Nicolson Diskretisierung (4.4)*

$$\|\partial_t u_l\|_{L^2(\Omega_T)} \leq \|f\|_{L^2(\Omega_T)} \quad (4.5)$$

und

$$\|u_l(t_k)\|_{L^2(\Omega)} \leq 2^{-l/2} \cdot k^{1/2} \|f\|_{L^2(\Omega_T)}. \quad (4.6)$$

**Beweis.** Da die Ableitung bzgl. der Zeit für jedes  $u_l \in \tilde{V}_l^{0,Ansatz}$  in dem Testraum  
liegt, d.h.  $\partial_t u_l \in \tilde{V}_l^{0,Test}$ , setzen wir  $\varphi = \partial_t u_l$  in (4.4) ein und erhalten mit partieller  
Integration

$$\int_0^T (\partial_t u_l, \partial_t u_l)_{L^2(\Omega)} dt + \frac{1}{2} a(u_l(\cdot, T), u_l(\cdot, T)) = \int_0^T f \partial_t u_l dt.$$

Da nach Voraussetzung  $a(u_l(\cdot, T), u_l(\cdot, T)) \geq 0$  gilt, folgt daraus die Ungleichung  
(4.5). Weiter gilt

$$\begin{aligned} \|u_l(t_k)\|_{L^2(\Omega)} &= \left\| \int_0^{t_k} \partial_t u_l(\tau) d\tau \right\|_{L^2(\Omega)} \\ &\leq \|\partial_t u_l\|_{L^2(\Omega) \otimes L^1((0, t_k))} \\ &\leq 2^{-l/2} \cdot k^{1/2} \cdot \|\partial_t u_l\|_{L^2(\Omega) \otimes L^2((0, t_k))} \\ &\leq 2^{-l/2} \cdot k^{1/2} \cdot \|f\|_{L^2(\Omega)} \end{aligned}$$

und die Abschätzung 4.6 ist gezeigt.  $\square$

Die Stabilitätsabschätzung zeigt, dass das Problem (4.4) eine eindeutige Lösung besitzt.

**Lemma 4.2.** *Die Bilinearform  $a(t; \cdot, \cdot)$  sei unabhängig von  $t$ ,  $a(t; \cdot, \cdot) = a(\cdot, \cdot)$ , und koerziv. Dann besitzt das diskrete Problem (4.4) eine eindeutige Lösung.*

Insbesondere zeigt das Lemma, dass das anfallende lineare Gleichungssystem, mit dessen Lösung wir uns später beschäftigen werden, eindeutig lösbar ist. Der Beweis dieses Lemmas folgt sofort aus dem vorherigen Resultat. Sind nämlich  $u_l$  und  $\tilde{u}_l$  zwei Lösungen von (4.4), so ist  $u_l - \tilde{u}_l$  eine Lösung von (4.4) mit  $u_l(\cdot, 0) - \tilde{u}_l(\cdot, 0) = 0$  und  $f = 0$ . Aus der Abschätzung (4.5) folgt damit  $u_l = \tilde{u}_l$ .

## 4.2 Das Discontinuous-Galerkin Verfahren

Bei der Discontinuous-Galerkin Methode [49, 50, 51, 52, 53] wird das Zeitgebiet  $(0, T]$  zunächst in Teilintervalle  $I_k = (t_{k-1}, t_k]$ ,  $1 \leq k \leq N$ , mit diskreten Zeitpunkten  $0 = t_0 < t_1 < \dots < t_N = T$  zerlegt. Für eine gegebene Ortsdiskretisierung  $V_h^\Omega$  mit Finiten Elementen und für fest gewähltes  $q \in \mathbb{N}$  betrachtet man dann den Raum

$$V := \left\{ v \mid v(x, t) \Big|_{I_k} = \sum_{j=0}^q t^j v_j(x) \text{ mit } v_j \in V_h^\Omega \right\}$$

von in der Zeit stückweise stetigen Funktionen. Dabei ist  $q$  der maximale Polynomgrad der Ansatzfunktionen in der Zeit. Man beachte, dass die Funktionen dieses Raumes im Allgemeinen unstetig an den Intervallgrenzen sind. Um dies in der folgenden Formulierung berücksichtigen zu können, benötigen wir noch die Definition der Sprungfunktionen

$$[v]_k := v_k^+ - v_k^- \text{ mit } v_k^{+(-)} := \lim_{s \rightarrow 0^{+(-)}} v(t_k + s)$$

für jedes  $v \in V$ . Bei der Discontinuous-Galerkin Diskretisierung des Problems (3.8) sucht man nun ein  $u_l \in V$ , so dass  $u_l$  die Anfangs- und Randbedingungen und

$$B(u_l, v) = \int_0^T (f, v)_{L^2(\Omega)} dt \text{ für alle } v \in V, \quad (4.7)$$

erfüllt, wobei

$$B(v, w) := \sum_{k=1}^N \int_{I_k} (\partial_t v, w)_{L^2(\Omega)} + a(t; v, w) dt + \sum_{k=0}^{N-1} ([v]_k, w_k^+)_{L^2(\Omega)} \quad (4.8)$$

und  $(u_l)_0^- := u_0$ . Für  $q = 0$  erhalten wir damit das implizite Euler Verfahren während  $q = 1$  ein Verfahren ähnlich dem subdiagonalen Padé-Schema der Ordnung  $(2, 1)$  liefert, vgl. [49].

Wie wir sehen, ist die abstrakte Formulierung (4.7) formal unabhängig von der Wahl des Ansatzraumes  $V$ . Wir können hier also ohne weiteres auch unsere Raum-Zeit Dünngitterräume verwenden, wobei wir weiterhin in der Bilinearform  $B(\cdot, \cdot)$  über die Zeitpunkte des in den Raum-Zeit Dünngittern vorkommenden feinsten Zeitlevels summieren. In Analogie zu dem obigen Raum  $V$  verwenden wir hierbei Raum-Zeit Dünngitter, die aus der stückweise konstanten oder der stückweise linearen Multilevel-Zerlegung des Abschnitts 2.3 in der Zeit und der hierarchischen Basis im Ort konstruiert werden. Wir erhalten damit die zu den klassischen Discontinuous-Galerkin Diskretisierungen analogen Raum-Zeit Dünngitterverfahren.

In [49] wird gezeigt, dass für lineare Finite Elemente im Ort mit geeigneten Glattheitsvoraussetzungen an die kontinuierliche Lösung  $u$  das Verfahren in der Zeit global von der Ordnung  $q + 1$  ist und an den Zeitschritten  $t_k$  sogar eine Ordnung von  $2q + 1$  erreicht wird. Wir untersuchen nun analog die Konvergenzrate des Raum-Zeit Dünngitterverfahrens mit stückweise konstanten bzw. stückweise linearen Funktionen in der Zeit. Hierbei werden wir von der Ungleichung

$$a \cdot b \leq \frac{\varepsilon}{2} a^2 + \frac{1}{2\varepsilon} b^2 \quad (4.9)$$

für beliebige  $a, b \in \mathbb{R}$  und  $\varepsilon > 0$  Gebrauch machen. Zusätzlich schreiben wir abkürzend  $a(\cdot, \cdot)$  anstelle von  $a(t; \cdot, \cdot)$ , obwohl die Bilinearform weiterhin von der Zeit abhängig sein kann. Wir erhalten zunächst das folgende Lemma.

**Lemma 4.3.** *Für alle  $v \in H_{mix}^{1,0}(\Omega_T)$  mit  $\partial_t(v|_{I_k}) \in L^2(\Omega \times I_k)$  für alle  $1 \leq k \leq N$  und  $v_0^- = 0$  gilt*

$$B(v, v) = \int_0^T a(v, v) dt + \frac{1}{2} D(v), \quad (4.10)$$

wobei  $D(v) := \|v_N^-\|_{L^2(\Omega)}^2 + \sum_{k=0}^{N-1} \|[v]_k\|_{L^2(\Omega)}^2$ .

**Beweis.** Partielle Integration liefert

$$\int_{I_k} (\partial_t v, v)_{L^2(\Omega)} dt = \frac{1}{2} \left( \|v_k^-\|_{L^2(\Omega)}^2 - \|v_{k-1}^+\|_{L^2(\Omega)}^2 \right).$$

Unter Verwendung von (4.9) mit  $\varepsilon = 1$  und  $v_0^- = 0$  erhalten wir

$$\begin{aligned} B(v, v) &= \sum_{k=0}^{N-1} \left( (v_k^+ - v_k^-, v_k^+)_{L^2(\Omega)} + \frac{1}{2} \left( \|v_k^-\|_{L^2(\Omega)}^2 - \|v_k^+\|_{L^2(\Omega)}^2 \right) \right) \\ &\quad + \int_0^T a(v, v) dt + \frac{1}{2} \|v_N^-\|_{L^2(\Omega)}^2 \\ &= \int_0^T a(v, v) dt + \sum_{k=0}^{N-1} \left( \frac{1}{2} \|v_k^+\|_{L^2(\Omega)}^2 - (v_k^-, v_k^+) + \frac{1}{2} \|v_k^-\|_{L^2(\Omega)}^2 \right) \\ &\quad + \frac{1}{2} \|v_N^-\|_{L^2(\Omega)}^2 \end{aligned}$$

und die Behauptung ist gezeigt.  $\square$

Ist  $a(t; \cdot, \cdot)$  positiv definit, so kann das obige Lemma als positive Definitheit des Discontinuous-Galerkin Operators  $B(\cdot, \cdot)$  interpretiert werden. Insbesondere folgt aus der Identität (4.10), dass das bei der Diskretisierung anfallende lineare Gleichungssystem eine eindeutige Lösung besitzt. Sind nämlich  $u_l$  und  $\tilde{u}_l$  zwei Lösungen des diskreten Problems (4.7), so ist  $u_l - \tilde{u}_l$  eine Lösung des diskreten Problems (4.7) mit  $f = 0$  und  $u_0 = 0$ . Deshalb gilt  $B(u_l - \tilde{u}_l, u_l - \tilde{u}_l) = 0$  und mit (4.10) folgt  $u_l = \tilde{u}_l$ .

Wir bezeichnen mit  $e_l := u - u_l$  den Fehler zwischen der kontinuierlichen Lösung  $u$  und der Lösung des diskretisierten Problems  $u_l$ .

Das folgende Resultat gibt eine obere Abschätzung für die Konvergenz der diskreten Lösung  $u_l$  gegen die kontinuierliche Lösung für die Dünngitterkonstruktion mittels stückweise konstanter Funktionen in der Zeit. Wir setzen für festes  $l$   $N := 2^l$  und beschränken uns zur Vereinfachung der Darstellung auf  $T = 1$ .

**Satz 4.4.** *Es gebe Konstanten  $C > 0$  und  $c > 0$ , so dass für die Bilinearform  $a(\cdot, \cdot)$  die Bedingungen (3.10) und (3.11) erfüllt sind. Sei die Lösung  $u$  des Problems (3.8) genügend glatt. Sei  $V_l^0$  der aus der hierarchischen Basis im Ort und der stückweise konstanten Zerlegung in der Zeit konstruierte Raum-Zeit Dünngitterraum und  $P_{V_l^0}$  sei der zugehörige Interpolationsoperator. Dann gibt es eine Konstante  $\tilde{c}$  unabhängig von  $l$  und  $u$ , so dass für die Lösung  $u_l \in V_l^0$  von (4.7) mit  $V = V_l^0$*

$$\|\nabla e_l\|_{L^2(\Omega_T)}^2 + D(e_l) \leq \tilde{c} \cdot \left( \|u - P_{V_l^0} u\|_{H_{mix}^{1,0}(\Omega_T)}^2 + 2^l \|u - P_{V_l^0} u\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)}^2 \right)$$

*gilt.*

**Beweis.** Da  $u$  und  $u_l$  die gleiche Anfangsbedingung erfüllen, gilt

$$e_l(x, 0) = 0 \text{ für alle } x \in \Omega$$

und wir können die Identität (4.10) des vorigen Lemmas zusammen mit der Koerzitivität (3.11) der Bilinearform  $a(\cdot, \cdot)$  ausnutzen. Wir erhalten dann

$$\|\nabla e_l\|_{L^2(\Omega_T)}^2 + D(e_l) \leq c \left( \int_0^T a(e_l, e_l) dt + \frac{1}{2} D(e_l) \right) = c \cdot B(e_l, e_l) \quad (4.11)$$

und es genügt,  $B(e_l, e_l)$  nach oben abzuschätzen. Ist  $u$  in den Zeitpunkten  $t_k$ ,  $1 \leq k \leq 2^l$ , stetig, so gilt  $[u_k] = 0$  für alle  $1 \leq k \leq 2^l = N$ , weshalb  $u$  als Lösung des parabolischen Problems trivialerweise (4.7) erfüllt. Hieraus folgt

$$B(e_l, \varphi) = 0 \text{ für alle } \varphi \in V_l^0,$$

und insbesondere gilt  $B(e_l, e_l) = B(e_l, u - P_{V_l^0} u)$ . Damit ergibt sich

$$\begin{aligned} B(e_l, e_l) &= B(e_l, u - P_{V_l^0} u) \\ &= \sum_{k=1}^N \int_{I_k} (\partial_t e_l, u - P_{V_l^0} u)_{L^2(\Omega)} dt + \int_0^T a(e_l, u - P_{V_l^0} u) dt \\ &\quad + \sum_{k=0}^{N-1} ([e_l]_k, (u - P_{V_l^0} u)_k^+)_{L^2(\Omega)}. \end{aligned}$$

Wir schätzen nun diese drei Terme weiter ab.

Für  $\int_0^T a(e_l, u - P_{V_l^0} u) dt$  erhalten wir unter Verwendung der Abkürzung  $\|\cdot\|_a^2 := a(\cdot, \cdot)$

$$\begin{aligned} \int_0^T a(e_l, u - P_{V_l^0} u) dt &\leq C \cdot \int_0^T \|e_l\|_a \|u - P_{V_l^0} u\|_a dt \\ &\leq \frac{1}{2} \int_0^T \|\nabla e_l\|_{L^2(\Omega)}^2 dt + \tilde{c} \cdot \int_0^T \|u - P_{V_l^0} u\|_{H^1(\Omega)}^2 dt, \end{aligned}$$

wobei wir im letzten Schritt die Ungleichung (4.9) mit geeignet gewähltem  $\varepsilon$  sowie die Stetigkeit von  $a(\cdot, \cdot)$  ausgenutzt haben.

Der Term  $\sum_{k=1}^N ([e_l]_{k-1}, (u - P_{V_l^0} u)_{k-1}^+)_{L^2(\Omega)}$  läßt sich wiederum mit (4.9) und geeignetem  $\varepsilon$  durch

$$\begin{aligned} \sum_{k=0}^{N-1} ([e_l]_k, (u - P_{V_l^0} u)_k^+)_{L^2(\Omega)} &\leq \sum_{k=0}^{N-1} \left( \frac{1}{4} \| [e_l]_k \|_{L^2(\Omega)}^2 + c^2 \cdot \| (u - P_{V_l^0} u)_k^+ \|_{L^2(\Omega)}^2 \right) \\ &\leq \frac{1}{4} D(e_l) + c 2^l \|u - P_{V_l^0} u\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)}^2 \end{aligned}$$

abschätzen.

Da  $u_l(x, t)$  auf jedem Intervall  $I_k$  in  $t$  konstant ist, also insbesondere  $\partial_t u_l|_{I_k} = 0$  gilt, folgt

$$\begin{aligned} \sum_{k=1}^N \int_{I_k} (\partial_t e_l, u - P_{V_l^0} u)_{L^2(\Omega)} dt &= \sum_{k=1}^N \int_{I_k} (\partial_t u, u - P_{V_l^0} u)_{L^2(\Omega)} dt \\ &\leq \|\partial_t u\|_{L^2(\Omega_T)} \cdot \|u - P_{V_l^0} u\|_{L^2(\Omega_T)} \end{aligned}$$

und die Behauptung ist bewiesen.  $\square$

Wie wir in dem obigen Beweis zu Satz 4.4 sehen, benötigen wir außer in der Abschätzung des Terms  $\sum_{k=1}^N \int_{I_k} (\partial_t e_l, u - P_{V_l^0} u)_{L^2(\Omega)} dt$  an keiner anderen Stelle, dass die Funktionen in der Zeit stückweise konstant sind. Wir können daher im wesentlichen den obigen Beweis für das folgende Resultat im Fall der stückweisen Linearität der Funktionen in der Zeit verwenden.

**Satz 4.5.** *Es gebe Konstanten  $c > 0$  und  $C > 0$  so dass (3.10) und (3.11) erfüllt sind. Die Lösung  $u$  des Problems (3.8) sei genügend glatt. Desweiteren sei  $V_l^0$  der aus der hierarchischen Basis im Ort und der stückweise linearen Zerlegung in der Zeit konstruierte Raum-Zeit Dünngitterraum mit zugehörigem Interpolationsoperator  $P_{V_l^0}$ . Dann gibt es eine Konstante  $\tilde{c}$  unabhängig von  $l$  und  $u$ , so dass für die Lösung  $u_l \in V_l^0$  (d.h.  $q = 1$ ) von (4.7)*

$$\begin{aligned} \|\nabla e_l\|_{L^2(\Omega_T)}^2 + D(e_l) &\leq \tilde{c} \cdot \left( \|u - P_{V_l^0} u\|_{H_{mix}^{1,0}(\Omega_T)}^2 + 2^l \|u - P_{V_l^0} u\|_{H_{mix}^{(0,2),(0,\infty)}(\Omega_T)}^2 \right. \\ &\quad \left. + \|u - P_{V_l^0} u\|_{L^2(\Omega_T)} \cdot \|\partial_t(u - P_{V_l^0} u)\|_{L^2(\Omega_T)} \right. \\ &\quad \left. + 2^{2l} \|u - P_{V_l^0} u\|_{L^2(\Omega_T)}^2 \right) \end{aligned}$$

*gilt.*

**Beweis.** Es verbleibt die Abschätzung des Terms  $\sum_{k=1}^N \int_{I_k} (\partial_t e_l, u - P_{V_l^0} u)_{L^2(\Omega)} dt$ . Da jedes  $v(x, t) \in V_l^0$  auf dem Intervall  $I_k$  linear in der Zeit ist, existiert eine Konstante  $c > 0$  mit

$$\|\partial_t v\|_{L^2(\Omega \times I_k)} \leq c 2^l \|v\|_{L^2(\Omega \times I_k)}, \quad (4.12)$$

die unabhängig von  $l$  und der gewählten Funktion  $v \in V_l^0$  ist. Dies folgt direkt aus den inversen Abschätzungen für Finite Elemente, vgl. [39]. Wir erhalten dann mit

der Abkürzung  $\Omega_k := \Omega \times I_k$  und unter Verwendung von (4.12)

$$\begin{aligned}
\sum_{k=1}^N \int_{I_k} (\partial_t(u - u_l), u - P_{V_l^0}u)_{L^2(\Omega)} dt &\leq \sum_{k=1}^N \|\partial_t(u - u_l)\|_{L^2(\Omega_k)} \cdot \|u - P_{V_l^0}u\|_{L^2(\Omega_k)} \\
&\leq \sum_{k=1}^N \|u - P_{V_l^0}u\|_{L^2(\Omega_k)} \cdot \left( \|\partial_t(u - P_{V_l^0}u)\|_{L^2(\Omega_k)} \right. \\
&\quad \left. + \|\partial_t(P_{V_l^0}u - u_l)\|_{L^2(\Omega_k)} \right) \\
&\leq \sum_{k=1}^N \|u - P_{V_l^0}u\|_{L^2(\Omega_k)} \cdot \left( \|\partial_t(u - P_{V_l^0}u)\|_{L^2(\Omega_k)} \right. \\
&\quad \left. + c2^l \|P_{V_l^0}u - u_l\|_{L^2(\Omega_k)} \right).
\end{aligned}$$

Die Anwendung der Cauchy-Schwarz Ungleichung auf die Summe liefert weiter

$$\begin{aligned}
\sum_{k=1}^N \int_{I_k} (\partial_t u - u_l, u - P_{V_l^0}u)_{L^2(\Omega)} dt &\leq \|u - P_{V_l^0}u\|_{L^2(\Omega_T)} \cdot \left( \|\partial_t(u - P_{V_l^0}u)\|_{L^2(\Omega_T)} \right. \\
&\quad \left. + c2^l \|P_{V_l^0}u - u_l\|_{L^2(\Omega_T)} \right) \\
&\leq \|u - P_{V_l^0}u\|_{L^2(\Omega_T)} \cdot \left( \|\partial_t(u - P_{V_l^0}u)\|_{L^2(\Omega_T)} \right. \\
&\quad \left. + c2^l (\|P_{V_l^0}u - u\|_{L^2(\Omega_T)} + \|u - u_l\|_{L^2(\Omega_T)}) \right).
\end{aligned}$$

Nach Anwendung der Poincaré-Ungleichung folgt mit Hilfe von (4.9) mit geeignet gewähltem  $\varepsilon$

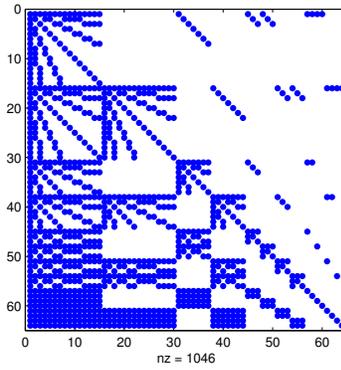
$$2^l \|u - P_{V_l^0}u\|_{L^2(\Omega_T)} \cdot \|u - u_l\|_{L^2(\Omega_T)} \leq c \cdot 2^{2l} \|u - P_{V_l^0}u\|_{L^2(\Omega_T)}^2 + \frac{1}{4} \|\nabla(u - u_l)\|_{L^2(\Omega_T)}^2.$$

Damit gilt dann schließlich

$$\begin{aligned}
\sum_{k=1}^N \int_{I_k} (\partial_t e_l, u - P_{V_l^0}u)_{L^2(\Omega)} dt &\leq c \|u - P_{V_l^0}u\|_{L^2(\Omega_T)} \cdot \left( \|\partial_t(u - P_{V_l^0}u)\|_{L^2(\Omega_T)} \right. \\
&\quad \left. + 2^{2l} \|u - P_{V_l^0}u\|_{L^2(\Omega_T)} \right) + \frac{1}{4} \|\nabla e_l\|_{L^2(\Omega_T)}^2
\end{aligned}$$

woraus die Behauptung folgt.  $\square$

In den obigen Fehlerabschätzungen wurde der Fehler gegen den Interpolationsfehler  $u - P_{V_l^0}u$  in bestimmten Normen abgeschätzt. Setzen wir für  $u$  eine höhere



**Abbildung 4.1.** Besetzungsstruktur der Matrix des Raum-Zeit Dünngitter Crank-Nicolson Verfahrens mit einer Ortsdimension auf Level  $l = 4$ .

Glattheit voraus, so können wir mit Hilfe der Resultate des Abschnitts 2.4 diesen Interpolationsfehler weiter abschätzen.

Betrachten wir zunächst den Fall des Discontinuous-Galerkin Verfahrens mit stückweise konstanten Funktionen in der Zeit, d.h. das Dünngitter  $V_l^0$  bzw.  $\tilde{V}_l^0$  das mit der stückweise konstanten hierarchischen Basis aus Abschnitt 2.3 und einer linearen bzw.  $d$ -linearen hierarchischen Basis im Ort konstruiert wird. Wir erhalten dann mit Satz 4.4 und den Interpolationsfehlerabschätzungen (2.61) und (2.62) die Fehlerabschätzung

$$\|\nabla e_l\|_{L^2(\Omega_T)}^2 + D(e_l) \leq \tilde{c}(2^{-2l}l^2\|u\|_{H_{mix}^{2,1}(\Omega_T)}^2 + 2^{-3l}l^2\|u\|_{H^{(2,2),(1,\infty)}(\Omega_T)}^2) \quad (4.13)$$

bei Verwendung des Raum-Zeit Dünngitterraumes  $\tilde{V}_l^0$  und entsprechend für  $V_l^0$

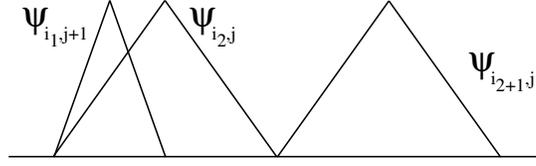
$$\|\nabla e_l\|_{L^2(\Omega_T)}^2 + D(e_l) \leq \tilde{c}(2^{-2l}\|u\|_{H_{mix}^{2,1}(\Omega_T)}^2 + 2^{-3l}l^2\|u\|_{H_{mix}^{(2,2),(1,\infty)}(\Omega_T)}^2). \quad (4.14)$$

Betrachten wir hingegen den Fall von stückweise linearen Funktionen in der Zeit, so erhalten wir mit den Interpolationsfehlerabschätzungen (2.48), (2.49) und (2.50) im Fall von  $\tilde{V}_l^0$

$$\begin{aligned} \|\nabla e_l\|_{L^2(\Omega_T)}^2 + D(e_l) \leq \tilde{c} & \left( 2^{-2l}\|u\|_{H_{mix}^{(2,2),(2,\infty)}(\Omega_T)}^2 + 2^{-3l}l^2\|u\|_{H_{mix}^{2,2}(\Omega_T)}^2 \right. \\ & \left. + 2^{-3l}l\|u\|_{H_{mix}^{2,2}(\Omega_T)}\|u\|_{H_{mix}^{1,2}(\Omega_T)} + 2^{-2l}l^2\|u\|_{H_{mix}^{2,2}(\Omega_T)}^2 \right). \end{aligned}$$

### 4.3 Das Unidirektionale Prinzip

Wie in den beiden vorhergehenden Abschnitten deutlich geworden ist, ist es möglich, klassische Verfahren zur Diskretisierung parabolischer Probleme formal auf die Diskretisierung mit Hilfe der Dünngitterräume zu übertragen. Durch die Verwendung



**Abbildung 4.2.** Eindimensionales Beispiel für die hierarchische Relation " $>$ ". Es gilt  $(j, i_2) > (j + 1, i_1)$  während zwischen  $(j + 1, i_1)$  und  $(i_2 + 1, j)$  keine Relation besteht.

von Multilevelbasen in der Dünngitterkonstruktion sind die durch die Diskretisierung anfallenden linearen Gleichungssysteme allerdings nicht dünn besetzt. Der Träger größerer Basisfunktionen überlappt mit den Trägern von Basisfunktionen auf feineren Leveln, weshalb man eine Art Fingerstruktur als Besetzungsmuster der resultierenden Matrix erhält. In Abbildung 4.1 ist die Besetzungsstruktur einer solchen Matrix für die Crank-Nicolson Diskretisierung auf Level  $l = 4$  dargestellt.

Allerdings ist es möglich, das Matrix-Vektor Produkt effizient, d.h. mit einem Aufwand proportional zu der Anzahl Unbekannter, zu berechnen. In [5, 6] wurde dazu ein Verfahren, das sogenannte Unidirektionale Prinzip, für klassische Dünngitter mit der linearen oder konstanten hierarchischen Basis vorgestellt. Dieses Prinzip wurde dann in [1, 32] auf klassische, aus hierarchischen Polynombasen konstruierten, Dünngitter erweitert.

In diesem Abschnitt stellen wir das Unidirektionale Prinzip etwas genauer vor. Da in der Literatur bisher nur in schwacher Form gegebene Differentialoperatoren behandelt wurden, wir aber durchaus auch allgemeinere Bilinearformen betrachten, wie etwa die Bilinearform der Discontinuous-Galerkin Diskretisierung (4.8), werden wir in einem ersten Schritt das Unidirektionale Prinzip auf beliebige Tensorprodukt-Bilinearformen verallgemeinern. In einem zweiten Schritt diskutieren wir, wie die bei dem Unidirektionalen Prinzip anfallenden niederdimensionalen Terme effizient berechnet werden können.

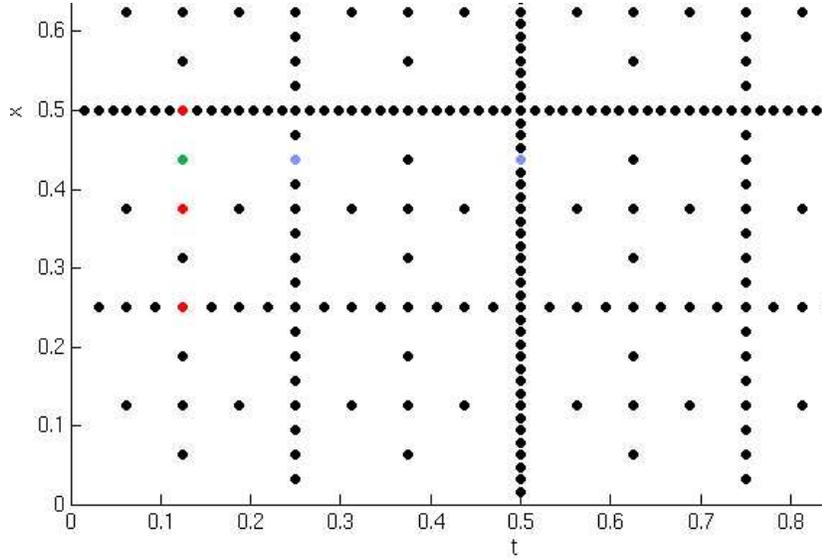
Zur Vereinfachung der Notation betrachten wir hier den Fall der Dünngitterräume  $\tilde{V}_l^0$ . Die Ergebnisse lassen sich allerdings direkt auf  $V_l^0$  übertragen. Wir definieren die Indexmenge, mit der wir im Folgenden das Raum-Zeit Dünngitter beschreiben werden, durch

$$G_l := \{(\mathbf{j}, \mathbf{i}) \mid j_1 + j_2 \leq l, 1 \leq i_1 \leq \dim(W_{j_1}^\Omega), 1 \leq i_2 \leq \dim(W_{j_2}^T)\}.$$

Desweiteren seien wie in Abschnitt 2.4 mit  $\{\psi_{j,i}^\Omega\}$  und  $\{\psi_{j,i}^T\}$  die Basen der Überschussräume  $W_j^\Omega$  bzw.  $W_j^T$  bezeichnet. Damit erhalten wir mit

$$\mathcal{B}_{V_l^0} = \{\psi_{\mathbf{j}, \mathbf{i}} := \psi_{j_1, i_1}^\Omega \cdot \psi_{j_2, i_2}^T \mid (\mathbf{j}, \mathbf{i}) \in G_l\} \quad (4.15)$$

eine Basis des Dünngitterraumes  $\tilde{V}_l^0$  und haben somit insbesondere für jede Funktion  $u_l \in \tilde{V}_l^0$  eine Darstellung  $u_l = \sum_{(\mathbf{j}, \mathbf{i}) \in G_l} u_{\mathbf{j}, \mathbf{i}} \psi_{\mathbf{j}, \mathbf{i}}$  mit geeigneten Koeffizienten  $u_{\mathbf{j}, \mathbf{i}}$ .



**Abbildung 4.3.** Ausschnitt eines Raum-Zeit Dünngitters. Für fest gewählten Index  $(\mathbf{j}, \mathbf{i})$  gehören die rot markierten Punkte zu der Menge  $\mathcal{I}_+^\Omega(\mathbf{j}, \mathbf{i}) \times \{(j_2, i_2)\}$  und die blau markierten Punkte zu  $\{(j_1, i_1)\} \times \mathcal{I}_+^T(\mathbf{j}, \mathbf{i})$  (blau).

Ferner definieren wir für die Indizes  $(j_1, i_1)$  und  $(\tilde{j}_1, \tilde{i}_1)$  der Multilevelbasis im Ort die hierarchischen Relationen

$$\begin{aligned} (j_1, i_1) > (\tilde{j}_1, \tilde{i}_1) &:\Leftrightarrow j_1 < \tilde{j}_1 \wedge \text{supp}(\psi_{j_1, i_1}^\Omega) \cap \text{supp}(\psi_{\tilde{j}_1, \tilde{i}_1}^\Omega) \neq \emptyset, \\ (j_1, i_1) \geq (\tilde{j}_1, \tilde{i}_1) &:\Leftrightarrow j_1 \leq \tilde{j}_1 \wedge \text{supp}(\psi_{j_1, i_1}^\Omega) \cap \text{supp}(\psi_{\tilde{j}_1, \tilde{i}_1}^\Omega) \neq \emptyset, \\ (j_1, i_1) < (\tilde{j}_1, \tilde{i}_1) &:\Leftrightarrow (\tilde{j}_1, \tilde{i}_1) > (j_1, i_1). \end{aligned}$$

Analog definieren wir die hierarchischen Relationen der Indizes für die Multilevelbasen in der Zeit. Abbildung 4.2 zeigt ein eindimensionales Beispiel für die hierarchische Relation. Man beachte, dass bei dieser Definition, im Gegensatz zu der bei den klassischen Dünngittern verwendeten eindimensionalen hierarchischen Basis, die Träger von hierarchisch niedrigeren Basisfunktionen nicht vollständig in den Trägern von hierarchisch höheren Basisfunktionen enthalten sein müssen.

Seien  $a^T(\cdot, \cdot) : V_l^T \times V_l^T \rightarrow \mathbb{R}$  und  $a^\Omega(\cdot, \cdot) : V_l^\Omega \times V_l^\Omega \rightarrow \mathbb{R}$  zwei Bilinearformen, für die

$$\text{supp}(\psi_{j,i}^T) \cap \text{supp}(\psi_{\tilde{j},\tilde{i}}^T) = \emptyset \Rightarrow a^T(\psi_{j,i}^T, \psi_{\tilde{j},\tilde{i}}^T) = 0, \quad (4.16)$$

$$\text{supp}(\psi_{j,i}^\Omega) \cap \text{supp}(\psi_{\tilde{j},\tilde{i}}^\Omega) = \emptyset \Rightarrow a^\Omega(\psi_{j,i}^\Omega, \psi_{\tilde{j},\tilde{i}}^\Omega) = 0 \quad (4.17)$$

gelten. Für  $u_l \in V_l^0$ ,  $u_l = \sum_{(\mathbf{j}, \mathbf{i}) \in G_l} u_{\mathbf{j}, \mathbf{i}} \psi_{\mathbf{j}, \mathbf{i}}$ , möchten wir nun die Größen

$$a(u_l, \psi_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}) := (a^\Omega \otimes a^T)(u_l, \psi_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}) \quad (4.18)$$

für alle Basisfunktionen  $\psi_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}$  des Dünngitters berechnen. Wir erhalten dann

$$\begin{aligned} a(u_l, \psi_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}) &= (a^\Omega \otimes a^T) \left( \sum_{(\mathbf{j}, \mathbf{i}) \in G_l} u_{\mathbf{j}, \mathbf{i}} \psi_{\mathbf{j}, \mathbf{i}}, \psi_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}} \right) \\ &= \sum_{(\mathbf{j}, \mathbf{i}) \in G_l} u_{\mathbf{j}, \mathbf{i}} a^\Omega(\psi_{j_1, i_1}^\Omega, \psi_{\tilde{j}_1, \tilde{i}_1}^\Omega) \cdot a^T(\psi_{j_2, i_2}^T, \psi_{\tilde{j}_2, \tilde{i}_2}^T) \end{aligned}$$

und definieren abhängig von  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})$  die Indexmengen

$$\begin{aligned} \mathcal{I}_+^\Omega(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) &:= \{(j_1, i_1) \mid (j_1, i_1) > (\tilde{j}_1, \tilde{i}_1), ((j_1, \tilde{j}_2), (i_1, \tilde{i}_2)) \in G_l\}, \\ \mathcal{I}_-^\Omega(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) &:= \{(j_1, i_1) \mid (j_1, i_1) \leq (\tilde{j}_1, \tilde{i}_1), ((j_1, \tilde{j}_2), (i_1, \tilde{i}_2)) \in G_l\}, \\ \mathcal{I}_+^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) &:= \{(j_2, i_2) \mid (j_2, i_2) > (\tilde{j}_2, \tilde{i}_2), ((\tilde{j}_1, j_2), (\tilde{i}_1, i_2)) \in G_l\}, \\ \mathcal{I}_-^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) &:= \{(j_2, i_2) \mid (j_2, i_2) \leq (\tilde{j}_2, \tilde{i}_2), ((\tilde{j}_1, j_2), (\tilde{i}_1, i_2)) \in G_l\}, \end{aligned}$$

siehe auch Abbildung 4.3. Da wegen (4.16) und (4.17) jeweils nur dann zwei Faktoren in der Summe nicht verschwinden, wenn für den Multilevelindex  $(\mathbf{j}, \mathbf{i})$

$$(j_1, i_1) \geq (\tilde{j}_1, \tilde{i}_1) \text{ oder } (j_1, i_1) < (\tilde{j}_1, \tilde{i}_1)$$

und gleichzeitig

$$(j_2, i_2) \geq (\tilde{j}_2, \tilde{i}_2) \text{ oder } (j_2, i_2) < (\tilde{j}_2, \tilde{i}_2)$$

gilt, können wir diese Summe auch umschreiben zu

$$\begin{aligned} a(u_l, \psi_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}) &= \sum_{(j_1, i_1) \in \mathcal{I}_+^\Omega(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})} \sum_{\substack{(j_2, i_2) \in \mathcal{I}_-^T((j_1, \tilde{j}_2), (i_1, \tilde{i}_2)) \\ \vee (j_2, i_2) \in \mathcal{I}_+^T((j_1, \tilde{j}_2), (i_1, \tilde{i}_2))}} u_{\mathbf{j}, \mathbf{i}} a^\Omega(\psi_{j_1, i_1}^\Omega, \psi_{\tilde{j}_1, \tilde{i}_1}^\Omega) \cdot a^T(\psi_{j_2, i_2}^T, \psi_{\tilde{j}_2, \tilde{i}_2}^T) + \\ &+ \sum_{\substack{(j_2, i_2) \in \mathcal{I}_+^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \\ \vee (j_2, i_2) \in \mathcal{I}_-^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})}} \sum_{(j_1, i_1) \in \mathcal{I}_-^\Omega((\tilde{j}_1, j_2), (\tilde{i}_1, i_2))} u_{\mathbf{j}, \mathbf{i}} a^\Omega(\psi_{j_1, i_1}^\Omega, \psi_{\tilde{j}_1, \tilde{i}_1}^\Omega) \cdot a^T(\psi_{j_2, i_2}^T, \psi_{\tilde{j}_2, \tilde{i}_2}^T) \\ &= \sum_{(j_1, i_1) \in \mathcal{I}_+^\Omega(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})} a^\Omega(a^T(\sum_{\substack{(j_2, i_2) \in \mathcal{I}_-^T((j_1, \tilde{j}_2), (i_1, \tilde{i}_2)) \\ \vee (j_2, i_2) \in \mathcal{I}_+^T((j_1, \tilde{j}_2), (i_1, \tilde{i}_2))}} u_{\mathbf{j}, \mathbf{i}} \psi_{j_2, i_2}^T, \psi_{\tilde{j}_2, \tilde{i}_2}^T) \psi_{j_1, i_1}^\Omega, \psi_{\tilde{j}_1, \tilde{i}_1}^\Omega) + \\ &+ \sum_{\substack{(j_2, i_2) \in \mathcal{I}_+^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \\ \vee (j_2, i_2) \in \mathcal{I}_-^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})}} a^T(a^\Omega(\sum_{(j_1, i_1) \in \mathcal{I}_-^\Omega((\tilde{j}_1, j_2), (\tilde{i}_1, i_2))} u_{\mathbf{j}, \mathbf{i}} \psi_{j_1, i_1}^\Omega, \psi_{\tilde{j}_1, \tilde{i}_1}^\Omega) \psi_{j_2, i_2}^T, \psi_{\tilde{j}_2, \tilde{i}_2}^T). \end{aligned}$$

Definieren wir nun für gegebene Koeffizienten  $u_{\mathbf{j}, \mathbf{i}} \in \mathbb{R}$  und beliebige  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$

$$\bar{u}_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}^\Omega := a^\Omega\left(\sum_{(j_1, i_1) \in \mathcal{I}_+^\Omega(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})} u_{\mathbf{j}, \mathbf{i}} \psi_{j_1, i_1}^\Omega, \psi_{\tilde{j}_1, \tilde{i}_1}^\Omega\right), \quad \underline{u}_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}^\Omega := a^\Omega\left(\sum_{(j_1, i_1) \in \mathcal{I}_+^\Omega(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})} u_{\mathbf{j}, \mathbf{i}} \psi_{j_1, i_1}^\Omega, \psi_{\tilde{j}_1, \tilde{i}_1}^\Omega\right),$$

sowie

$$\bar{u}_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^T := a^T \left( \sum_{(j_2, i_2) \in \mathcal{I}_+^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})} u_{\mathbf{j},\mathbf{i}} \psi_{j_2, i_2}^T, \psi_{j_2, i_2}^T \right), \quad \underline{u}_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^T := a^T \left( \sum_{(j_2, i_2) \in \mathcal{I}_-^T(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})} u_{\mathbf{j},\mathbf{i}} \psi_{j_2, i_2}^T, \psi_{j_2, i_2}^T \right),$$

so erhalten wir aus den obigen Überlegungen den Algorithmus 4.1 zur Berechnung der Werte  $a(u_l, \psi_{\mathbf{j},\mathbf{i}})$  für alle  $(\mathbf{j}, \mathbf{i}) \in G_l$ .

ALGORITHMUS 4.1 (Unidirektionales Prinzip).

1. Gegeben seien Koeffizienten  $u_{\mathbf{j},\mathbf{i}} \in \mathbb{R}$  der Darstellung  $u_l = \sum_{(\mathbf{j},\mathbf{i}) \in G_l} u_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}}$ .
2. Setze  $v_{\mathbf{j},\mathbf{i}} = u_{\mathbf{j},\mathbf{i}}$  für alle  $(\mathbf{j}, \mathbf{i}) \in G_l$ .
3. Für alle  $\tilde{j}_2, \tilde{i}_2$  so dass es  $\tilde{j}_1, \tilde{i}_1$  mit  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$  gibt:
  - (a) Für alle  $\tilde{j}_1, \tilde{i}_1$  mit  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$ :  $u_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}} = u_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^\Omega$ .
4. Für alle  $\tilde{j}_1, \tilde{i}_1$  so dass es  $\tilde{j}_2, \tilde{i}_2$  mit  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$  gibt:
  - (a) Für alle  $\tilde{j}_2, \tilde{i}_2$  mit  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$ :  $u_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}} = u_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^T + \bar{u}_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^T$ ,  $v_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}} = v_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^T + \bar{v}_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^T$ .
5. Für alle  $\tilde{j}_2, \tilde{i}_2$  so dass es  $\tilde{j}_1, \tilde{i}_1$  mit  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$  gibt:
  - (a) Für alle  $\tilde{j}_1, \tilde{i}_1$  mit  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$ :  $v_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}} = \bar{v}_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}^\Omega$ .
6. Für alle  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}})$ :  $u_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}} = u_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}} + v_{\tilde{\mathbf{j}},\tilde{\mathbf{i}}}$ .

Haben wir nun einen Algorithmus verfügbar, der die inneren Schleifen (a) effizient berechnet, so ermöglicht der Algorithmus 4.1 die effiziente Berechnung des Matrix-Vektor Produktes. Für elliptische Probleme und klassische Dünngitter sind solche Algorithmen bekannt und werden Top-Down bzw. Bottom-Up Algorithmen genannt, vgl. [1, 5, 6, 32, 56]. Bei diesen Ansätzen nutzt man aus, dass die eindimensionalen Multilevelbasisfunktionen durch Translation und Dilatation einer (oder mehrerer) Mutterfunktionen entstehen. Dabei werden die Werte der Bilinearform bezüglich der Mutterfunktion berechnet und dann geeignet (durch Transformation) auf die anderen Multilevelbasisfunktionen übertragen. Bei Anordnung der eindimensionalen Multilevelbasen in einer Baumstruktur können somit die entsprechenden Werte mittels einer Pre- oder Post-Order Traversierung (daher der jeweilige Name Bottom-Up oder Top-Down) durch den Baum berechnet werden.

Dieses Vorgehen ist in unserem Fall nicht möglich, da wir im Ort die aus einer Sequenz von Finiten Element Räumen konstruierte hierarchische Basis verwenden, für die wir keine Mutterfunktion haben, die durch einfache Translation und Dilatation die anderen Funktionen generiert. Zusätzlich besteht aus unserer Sicht bei der in

der Literatur beschriebenen Vorgehensweise das Problem, dass für jede neue Bilinearform oder Multilevelbasis die Algorithmen jeweils neu bestimmt werden müssen. Dies macht eine Verwendung in einem flexiblen Code, der eine möglichst große Problemklasse abdecken soll, nahezu unmöglich.

Wir stellen daher im Folgenden eine Möglichkeit vor, wie analoge Algorithmen aus einer gegebenen Folge von Interpolationsoperatoren zwischen den Gittern und der auf dem feinsten Level diskretisierten Bilinearform gewonnen werden können. Dieses Prinzip ist sehr flexibel und kann direkt auf beliebige Bilinearformen, die nur auf dem feinsten Gitter diskretisiert werden müssen, ohne Modifikationen angewendet werden.

Bevor wir mit der Beschreibung dieser Algorithmen im nächsten Abschnitt fortfahren, möchten wir an dieser Stelle allerdings auf einen weiteren wichtigen Unterschied zwischen klassischen Dünngittern und unseren Raum-Zeit Dünngittern hinweisen. Um das Unidirektionale Prinzip anwenden zu können, benötigen wir die Tensorproduktstruktur der Bilinearform bezüglich Raum und Zeit. Betrachten wir die Discontinuous-Galerkin bzw. die Crank-Nicolson Diskretisierungen der letzten Abschnitte, so sehen wir, dass diese Bilinearformen immer als Summe von solchen Tensorprodukten geschrieben werden können, solange der elliptische Teil  $a(\cdot, \cdot)$  des parabolischen Operators Tensorproduktstruktur bezüglich Raum und Zeit besitzt. Dies ist zum Beispiel immer dann der Fall, wenn  $a(\cdot, \cdot)$  zeitunabhängig ist. Klassische Dünngitter benötigen hingegen, dass  $a(\cdot, \cdot)$  als Tensorprodukt von eindimensionalen Bilinearformen geschrieben werden kann, was eine sehr starke Einschränkung darstellt. In [1, 46, 47, 93] wurde daher vorgeschlagen, im Fall von elliptischen Bilinearformen mit beliebigen variablen Koeffizientenfunktionen, die Koeffizienten durch entsprechende Dünngitterinterpolanten zu ersetzen. Dabei besitzt der resultierende Operator, da der Dünngitterinterpolant Tensorproduktstruktur aufweist, ebenfalls Tensorproduktstruktur und das Unidirektionale Prinzip kann etwas modifiziert angewendet werden. Allerdings schränken bei diesem Vorgehen die Glattheitsanforderungen an die Koeffizientenfunktionen die behandelbaren Probleme sehr ein und es ist bis heute unklar, wie die Konvergenzraten dadurch beeinflusst werden. Wir sehen also, dass mit Hilfe der Raum-Zeit Dünngitter durch Ausnutzung der natürlichen Tensorproduktstruktur zwischen Raum und Zeit eine größere Klasse von Problemen behandelt werden kann, als dies im Fall von klassischen Dünngittern möglich ist.

### 4.3.1 Top-Down und Bottom-Up

Wir beschäftigen uns nun mit der effizienten Berechnung der inneren Schleifen (a) des Algorithmus 4.1.

Da die Algorithmen für die Zeitprobleme bzgl.  $a^T(\cdot, \cdot)$  nur (eindimensionale) Spezialfälle der Algorithmen für die anfallenden Ortsprobleme sind, betrachten wir hier aus Gründen der Übersichtlichkeit unter Verwendung der Notation aus Ab-

schnitt 2.4 nur die Algorithmen für die hierarchische Basis im Ort. Wir möchten also für feste  $\tilde{j}_2, \tilde{i}_2$  und für alle  $\tilde{j}_1, \tilde{i}_1$  mit  $(\tilde{\mathbf{j}}, \tilde{\mathbf{i}}) \in G_l$  die Terme  $\bar{u}_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}^\Omega$  und  $\underline{u}_{\tilde{\mathbf{j}}, \tilde{\mathbf{i}}}^\Omega$  berechnen. Für den gesamten Abschnitt sei daher  $\tilde{j}_2, \tilde{i}_2$  fest gewählt und wir definieren  $G_l^\Omega := \{(j_1, i_1) \mid (j_1, \tilde{j}_2, i_1, \tilde{i}_2) \in G_l\}$ .

Wie in Abschnitt 2.4 sei mit  $\{\varphi_{j,i}^\Omega\}$  die nodale Basis von  $V_j^\Omega$  bezeichnet und  $P_j : C^0(\Omega) \rightarrow V_j^\Omega$  sei der nodale Interpolationsoperator. Für eine fest gegebene Funktion  $u_l = \sum_{j,i} u_{j,i} \psi_{j,i}^\Omega$  möchten wir nun für alle  $(\tilde{j}, \tilde{i})$  den Term

$$\sum_{\substack{(j,i) > (\tilde{j}, \tilde{i}), \\ (j,i) \in G_l^\Omega}} a^\Omega(u_{j,i} \psi_{j,i}^\Omega, \psi_{\tilde{j}, \tilde{i}}^\Omega) \quad (4.19)$$

berechnen. Wir unterteilen dazu die Summe (4.19) in die einzelnen Level und erhalten

$$\sum_{\substack{(j,i) > (\tilde{j}, \tilde{i}), \\ (j,i) \in G_l^\Omega}} a^\Omega(u_{j,i} \psi_{j,i}^\Omega, \psi_{\tilde{j}, \tilde{i}}^\Omega) = \sum_{j < \tilde{j}} a^\Omega\left(\sum_{\substack{i, \\ (j,i) > (\tilde{j}, \tilde{i})}} u_{j,i} \psi_{j,i}^\Omega, \psi_{\tilde{j}, \tilde{i}}^\Omega\right).$$

Wir nutzen nun aus, dass die Räume  $V_j^\Omega$  ineinander geschachtelt sind, so dass wir die Funktion  $\sum_{i, (j,i) > (\tilde{j}, \tilde{i})} u_{j,i} \psi_{j,i}^\Omega$  mit Hilfe des Interpolationsoperators  $P_{\tilde{j}}$  exakt in  $V_{\tilde{j}}^\Omega$  darstellen können und es folgt

$$\sum_{j < \tilde{j}} a^\Omega\left(\sum_{i, (j,i) > (\tilde{j}, \tilde{i})} u_{j,i} \psi_{j,i}^\Omega, \psi_{\tilde{j}, \tilde{i}}^\Omega\right) = \sum_{j < \tilde{j}} a^\Omega(P_{\tilde{j}} \cdot \dots \cdot P_{j+1}\left(\sum_{i, (j,i) > (\tilde{j}, \tilde{i})} u_{j,i} \psi_{j,i}^\Omega\right), \psi_{\tilde{j}, \tilde{i}}^\Omega). \quad (4.20)$$

Wir betrachten nun den zweiten Term

$$\sum_{\substack{(j,i) \leq (\tilde{j}, \tilde{i}), \\ (j,i) \in G_l^\Omega}} a^\Omega(u_{j,i} \psi_{j,i}^\Omega, \psi_{\tilde{j}, \tilde{i}}^\Omega), \quad (4.21)$$

bei dem wir wiederum die Summe nach Leveln sortieren und die Testfunktionen auf dem feinsten Level darstellen

$$\sum_{\substack{(j,i) \leq (\tilde{j}, \tilde{i}), \\ (j,i) \in G_l^\Omega}} a^\Omega(u_{j,i} \psi_{j,i}^\Omega, \psi_{\tilde{j}, \tilde{i}}^\Omega) = \sum_{j \geq \tilde{j}} \sum_{\substack{(j,i) \in G_l^\Omega \\ (j,i) \leq (\tilde{j}, \tilde{i})}} a^\Omega(u_{j,i} \psi_{j,i}^\Omega, P_j \cdot \dots \cdot P_{\tilde{j}} \psi_{\tilde{j}, \tilde{i}}^\Omega). \quad (4.22)$$

Sei  $A^j$  die zu  $a^\Omega(\cdot, \cdot)$  gehörige Systemmatrix bzgl. der nodalen Basis  $\varphi_{j,i}^\Omega$  des Raumes  $V_j^\Omega$ , d.h.  $A^j = (a_{ik}^j)$  mit

$$a_{ik}^j = a^\Omega(\varphi_{j,k}^\Omega, \varphi_{j,i}^\Omega).$$

Weiter sei  $P_j^{j+1} \in \mathbb{R}^{\dim(V_{j+1}^\Omega) \times \dim(V_j^\Omega)}$  die Matrixdarstellung der auf den Raum  $V_j^\Omega$  eingeschränkten Prolongation  $P_{j+1}$  bzgl. der nodalen Basis. Dann erhalten wir mit (4.20) den Algorithmus 4.2 zur Berechnung von (4.19) für fest gegebene Koeffizienten

$u_{j,i}$  und alle  $(\tilde{j}, \tilde{i}) \in G_l^\Omega$ , wobei das Ergebnis in den Koeffizienten  $u_{j,i}$  abgespeichert wird.

ALGORITHMUS 4.2 (Top-Down).

1. Gegeben: Koeffizienten  $u_{j,i} \in \mathbb{R}$  der Darstellung  $u_l = \sum_{(j,i) \in G_l^\Omega} u_{j,i} \psi_{j,i}^\Omega$  und Vektoren  $w^j \in \mathbb{R}^{\dim(V_j^\Omega)}$  für alle  $0 \leq j \leq l$ .
2. Für alle  $0 \leq j \leq l$ : Setze  $w^j$  so, dass  $w_i^j = 0$  falls  $\varphi_{j,i}^\Omega \notin W_j^\Omega$ , und sonst  $w_i^j = u_{k,l}$  mit  $\varphi_{j,i}^\Omega = \psi_{k,l}^\Omega$ .
3. Setze  $j = 0$  und solange  $j \leq l$ :
  - (a) Falls  $j = 0$ :  $w^j = A^j w^j$ ,
  - (b) Sonst:  $w^j = A^j(w^j + P_{j-1}^j w^j)$ .
  - (c)  $j = j + 1$ .
4. Setze für alle  $(j, i) \in G_l^\Omega$   $u_{j,i} = u_k^j$  mit  $\psi_{j,i}^\Omega = \varphi_{j,k}^\Omega$ .

Fällt die Dimension der Räume  $V_j^\Omega$  geometrisch mit fallendem Level  $j$  ab (wie dies bei einer Sequenz von Finite Element Räumen, die durch uniforme Verfeinerung einer groben Triangulierung entstehen, der Fall ist), so sehen wir, dass der Top-Down Algorithmus lediglich  $O(\dim(V_l^\Omega))$  Rechenoperationen und Speicherplatz benötigt.

Analog erhalten wir mit Hilfe von (4.22) zur Berechnung von (4.21) für fest gegebene Koeffizienten  $u_{j,i}$  und alle  $(\tilde{j}, \tilde{i}) \in G_l^\Omega$  den Algorithmus 4.3.

ALGORITHMUS 4.3 (Top-Down).

1. Gegeben: Koeffizienten  $u_{j,i} \in \mathbb{R}$  der Darstellung  $u_l = \sum_{(j,i) \in G_l^\Omega} u_{j,i} \psi_{j,i}^\Omega$  und Vektoren  $w^j \in \mathbb{R}^{\dim(V_j^\Omega)}$  für alle  $0 \leq j \leq l$ .
2. Für alle  $0 \leq j \leq l$ : Setze  $w^j$  so, dass  $w_i^j = 0$  falls  $\varphi_{j,i}^\Omega \notin W_j^\Omega$ , und sonst  $w_i^j = u_{k,l}$  mit  $\varphi_{j,i}^\Omega = \psi_{k,l}^\Omega$ .
3. Setze  $j = l$  und solange  $j \geq 0$ :
  - (a) Falls  $j = l$ :  $w^j = A^j w^j$ ,
  - (b) Sonst:  $w^j = A^j w^j + (P_j^{j+1})^T w^{j+1}$ .
  - (c)  $j = j - 1$ .
4. Setze für alle  $(j, i) \in G_l^\Omega$   $u_{j,i} = u_k^j$  mit  $\psi_{j,i}^\Omega = \varphi_{j,k}^\Omega$ .

Obwohl in der obigen Diskussion Ansatz- und Testraum gleich waren, lassen sich die Algorithmen unmittelbar auf den Petrov-Galerkin Fall übertragen. Hierfür verwendet man in dem Bottom-Up Algorithmus lediglich die Prolongationen zwischen den Ansatzräumen, während im Fall des Top-Down Algorithmus die Prolongationen der Testräume verwendet werden müssen.

#### 4.4 Numerische Ergebnisse

In diesem Abschnitt untersuchen wir die Konvergenzraten des Discontinuous-Galerkin Verfahrens und des Crank-Nicolson Verfahrens anhand einiger Beispiele. Hierbei sind wir an dem Fehlerverhalten bzgl. der Normen

$$\begin{aligned} \|u\|_{L^\infty} &:= \operatorname{ess\,sup}_{t \in (0, T)} \|u(\cdot, t)\|_{L^\infty(\Omega)}, \\ \|u\|_{L^2} &:= \left( \int_{0 < t < T} \|u(\cdot, t)\|_{L^2(\Omega)}^2 dt \right)^{1/2}, \\ \|u\|_{H_{mix}^{(1,2),(0,\infty)}} &:= \operatorname{ess\,sup}_{t \in (0, T)} |u(\cdot, t)|_{H^1(\Omega)}, \\ \|u\|_{H_{mix}^{(1,2),(0,2)}} &:= \left( \int_{0 < t < T} |u(\cdot, t)|_{H^1(\Omega)}^2 dt \right)^{1/2} \end{aligned}$$

interessiert. Wir approximieren dabei den Fehler  $e_l := u_l - u$  zwischen der Dünngitterlösung  $u_l$  der diskretisierten Gleichungen auf Level  $l$  und der kontinuierlichen Lösung  $u$  durch den Fehler  $\hat{e}_l$  zwischen dem Interpolanten  $u_l^\infty \in V_l^\infty$  bzw.  $u_l^\infty \in \tilde{V}_l^\infty$  der Lösung  $u$  und der Lösung des diskretisierten Problems,

$$\hat{e}_l := u_l - u_l^\infty.$$

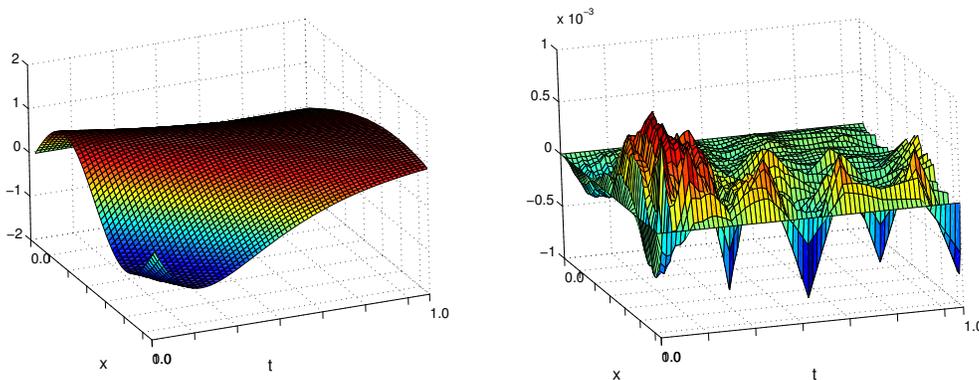
Wir messen im folgenden die relativen Fehler

$$\begin{aligned} e_{L^\infty} &:= \frac{\|\hat{e}_l\|_{L^\infty}}{\|u\|_{L^\infty}}, & e_{L^2} &:= \frac{\|\hat{e}_l\|_{L^2}}{\|u\|_{L^2}}, \\ e_{H_{mix}^{(1,2),(0,\infty)}} &:= \frac{\|\hat{e}_l\|_{H_{mix}^{(1,2),(0,\infty)}}}{\|u\|_{H_{mix}^{(1,2),(0,\infty)}}}, & e_{H_{mix}^{1,0}} &:= \frac{\|\hat{e}_l\|_{H_{mix}^{1,0}}}{\|u\|_{H_{mix}^{1,0}}}. \end{aligned}$$

Zur Bestimmung der Konvergenzordnung untersuchen wir zusätzlich die zugehörigen Fehlerreduktionen

$$\begin{aligned} \rho_{L^\infty}^l &:= \frac{\|\hat{e}_{l+1}\|_{L^\infty}}{\|\hat{e}_l\|_{L^\infty}}, & \rho_{L^2}^l &:= \frac{\|\hat{e}_{l+1}\|_{L^2}}{\|\hat{e}_l\|_{L^2}}, \\ \rho_{H_{mix}^{(1,2),(0,\infty)}}^l &:= \frac{\|\hat{e}_{l+1}\|_{H_{mix}^{(1,2),(0,\infty)}}}{\|\hat{e}_l\|_{H_{mix}^{(1,2),(0,\infty)}}}, & \rho_{H_{mix}^{1,0}}^l &:= \frac{\|\hat{e}_{l+1}\|_{H_{mix}^{1,0}}}{\|\hat{e}_l\|_{H_{mix}^{1,0}}}. \end{aligned}$$

Während wir für das Discontinuous-Galerkin Verfahren in Abschnitt 4.2 den Fehler bezüglich der  $H_{mix}^{1,0}$ -Norm gegen den Dünngitter-Interpolationsfehler abgeschätzt



**Abbildung 4.4.** Lösung  $u(x, t) = \sin(\pi \cdot \frac{x}{t+0.5})$  (links) des eindimensionalen Problems (4.25) und Fehler (rechts) der Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung auf Level  $l = 6$ .

haben, stehen uns solche Abschätzungen für das Crank-Nicolson Verfahren sowie für die anderen Normen im Falle des Discontinuous-Galerkin Verfahrens nicht zur Verfügung. Bestenfalls können wir Fehlerreduktionsraten erwarten, die den Fehlerreduktionsraten des Interpolationsfehlers der zugrunde liegenden Ansatzräume entsprechen. Nach den Sätzen 2.12 und 2.13 erwarten wir daher im günstigsten Fall für das Crank-Nicolson Verfahren bzgl.  $\tilde{V}_l^0$ , abgekürzt CN, und das Discontinuous Galerkin Verfahren mit in der Zeit stückweise linearen Funktionen, abgekürzt DG1,

$$\begin{aligned} \rho_{L^\infty}^l &\approx 0.25 \frac{l+1}{l}, & \rho_{L^2}^l &\approx 0.25 \frac{l+1}{l}, \\ \rho_{H_{mix}^{(1,2),(0,\infty)}}^l &\approx 0.5, & \rho_{H_{mix}^{1,0}}^l &\approx 0.5. \end{aligned} \quad (4.23)$$

Im Fall des Euler Verfahrens mit  $\tilde{V}_l^0$  erwarten wir nach Satz 2.14

$$\begin{aligned} \rho_{L^\infty}^l &\approx 0.5 \frac{l+1}{l}, & \rho_{L^2}^l &\approx 0.5 \frac{l+1}{l}, \\ \rho_{H_{mix}^{(1,2),(0,\infty)}}^l &\approx 0.5 \frac{l+1}{l}, & \rho_{H_{mix}^{1,0}}^l &\approx 0.5 \frac{l+1}{l}. \end{aligned} \quad (4.24)$$

während die Interpolationsfehlerabschätzungen für das Euler Verfahren bzgl.  $V_l^0$ , das wir auch als anisotropes Eulerverfahren bezeichnen, die gleichen oberen Schranken für die Fehlerreduktionen wie im Fall des CN und des DG1 Verfahrens liefern.

Sofern nicht anders erwähnt verwenden wir im Ort die aus einer Sequenz von uniformen Triangulierungen konstruierte  $d$ -lineare hierarchische Basis. Alle für diese Arbeit benötigten Ortsdiskretisierungen wurden mit Hilfe des frei verfügbaren Finite Element Paketes deal.II [7] berechnet.

*Beispiel 4.1.* Wir untersuchen zunächst ein eindimensionales Problem. Gesucht

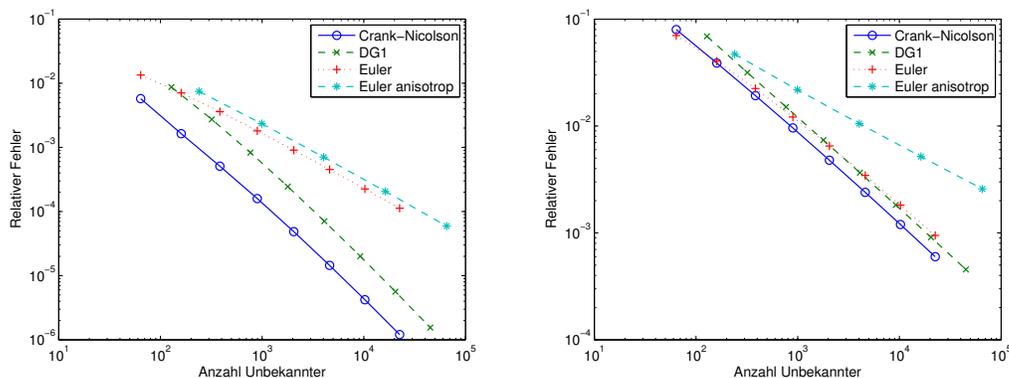
**Tabelle 4.1.** Konvergenz der verschiedenen Diskretisierungsverfahren für das eindimensionale Problem (4.25) mit Lösung  $u(x, t) = \sin(\pi \cdot \frac{x}{t+0.5})$ .

l	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H_{mix}^{(1,2),(0,\infty)}}$	$\rho_{H_{mix}^{(1,2),(0,\infty)}}$	$e_{H_{mix}^{1,0}}$	$\rho_{H_{mix}^{1,0}}$
Crank-Nicolson								
4	1.079 <sub>-2</sub>	-	5.750 <sub>-3</sub>	-	1.755 <sub>-1</sub>	-	7.975 <sub>-2</sub>	-
5	2.856 <sub>-3</sub>	0.26	1.637 <sub>-3</sub>	0.28	9.012 <sub>-2</sub>	0.51	3.880 <sub>-2</sub>	0.49
6	8.636 <sub>-4</sub>	0.30	5.091 <sub>-4</sub>	0.31	4.536 <sub>-2</sub>	0.50	1.922 <sub>-2</sub>	0.50
7	2.518 <sub>-4</sub>	0.29	1.590 <sub>-4</sub>	0.31	2.272 <sub>-2</sub>	0.50	9.582 <sub>-3</sub>	0.50
8	7.303 <sub>-5</sub>	0.29	4.856 <sub>-5</sub>	0.31	1.136 <sub>-2</sub>	0.50	4.788 <sub>-3</sub>	0.50
9	2.076 <sub>-5</sub>	0.28	1.446 <sub>-5</sub>	0.30	5.680 <sub>-3</sub>	0.50	2.394 <sub>-3</sub>	0.50
10	5.817 <sub>-6</sub>	0.28	4.211 <sub>-6</sub>	0.29	2.840 <sub>-3</sub>	0.50	1.197 <sub>-3</sub>	0.50
11	1.611 <sub>-6</sub>	0.28	1.205 <sub>-6</sub>	0.29	1.420 <sub>-3</sub>	0.50	5.986 <sub>-4</sub>	0.50
DG1								
4	2.292 <sub>-2</sub>	-	8.730 <sub>-3</sub>	-	1.394 <sub>-1</sub>	-	6.907 <sub>-2</sub>	-
5	7.546 <sub>-3</sub>	0.33	2.736 <sub>-3</sub>	0.31	7.213 <sub>-2</sub>	0.52	3.158 <sub>-2</sub>	0.46
6	2.333 <sub>-3</sub>	0.31	8.292 <sub>-4</sub>	0.30	3.639 <sub>-2</sub>	0.50	1.507 <sub>-2</sub>	0.48
7	6.904 <sub>-4</sub>	0.30	2.446 <sub>-4</sub>	0.29	1.824 <sub>-2</sub>	0.50	7.378 <sub>-3</sub>	0.49
8	1.988 <sub>-4</sub>	0.29	7.065 <sub>-5</sub>	0.29	9.126 <sub>-3</sub>	0.50	3.657 <sub>-3</sub>	0.50
9	5.537 <sub>-5</sub>	0.28	7.527 <sub>-5</sub>	0.28	8.006 <sub>-3</sub>	0.50	4.925 <sub>-3</sub>	0.50
10	1.544 <sub>-5</sub>	0.28	2.112 <sub>-5</sub>	0.28	4.003 <sub>-3</sub>	0.50	2.459 <sub>-3</sub>	0.50
11	4.281 <sub>-6</sub>	0.28	5.820 <sub>-6</sub>	0.28	2.002 <sub>-3</sub>	0.50	1.229 <sub>-3</sub>	0.50
Euler								
4	3.296 <sub>-2</sub>	-	1.343 <sub>-2</sub>	-	1.159 <sub>-1</sub>	-	7.013 <sub>-2</sub>	-
5	1.879 <sub>-2</sub>	0.57	7.078 <sub>-3</sub>	0.53	7.709 <sub>-2</sub>	0.67	4.042 <sub>-2</sub>	0.58
6	1.117 <sub>-2</sub>	0.59	3.602 <sub>-3</sub>	0.51	4.706 <sub>-2</sub>	0.61	2.240 <sub>-2</sub>	0.55
7	6.180 <sub>-3</sub>	0.55	1.809 <sub>-3</sub>	0.50	2.726 <sub>-2</sub>	0.58	1.215 <sub>-2</sub>	0.54
8	3.276 <sub>-3</sub>	0.53	9.043 <sub>-4</sub>	0.50	1.527 <sub>-2</sub>	0.56	6.497 <sub>-3</sub>	0.53
9	1.693 <sub>-3</sub>	0.52	4.515 <sub>-4</sub>	0.50	8.374 <sub>-3</sub>	0.55	3.445 <sub>-3</sub>	0.53
10	8.624 <sub>-4</sub>	0.51	2.254 <sub>-4</sub>	0.50	4.524 <sub>-3</sub>	0.54	1.814 <sub>-3</sub>	0.53
11	4.407 <sub>-4</sub>	0.51	4.227 <sub>-4</sub>	0.50	4.259 <sub>-3</sub>	0.54	2.556 <sub>-3</sub>	0.52
Euler anisotrop								
4	2.250 <sub>-2</sub>	-	7.453 <sub>-3</sub>	-	1.042 <sub>-1</sub>	-	4.693 <sub>-2</sub>	-
5	8.005 <sub>-3</sub>	0.36	2.343 <sub>-3</sub>	0.31	5.414 <sub>-2</sub>	0.52	2.183 <sub>-2</sub>	0.47
6	2.606 <sub>-3</sub>	0.33	7.058 <sub>-4</sub>	0.30	2.731 <sub>-2</sub>	0.50	1.053 <sub>-2</sub>	0.48
7	8.027 <sub>-4</sub>	0.31	2.066 <sub>-4</sub>	0.29	1.367 <sub>-2</sub>	0.50	5.187 <sub>-3</sub>	0.49
8	2.384 <sub>-4</sub>	0.30	2.222 <sub>-4</sub>	0.29	1.199 <sub>-2</sub>	0.50	6.972 <sub>-3</sub>	0.50

ist hierbei eine Funktion  $u$ , die die Gleichungen

$$\begin{aligned}
 \partial_t u - \partial_{xx} u &= f \text{ für alle } x \in \Omega = (0, 1), t \in (0, 1], \\
 u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega, \\
 u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega
 \end{aligned}
 \tag{4.25}$$

erfüllt. Dabei setzen wir die rechte Seite  $f$ , die Randbedingung  $g$  und die Anfangs-



**Abbildung 4.5.** Konvergenzhistorie für das Problem (4.25) bzgl. der  $L^2$ - (links) und der  $H_{mix}^{1,0}$ -Norm (rechts) für das CN, das DG1, das Euler und das anisotrope Euler Verfahren.

bedingung  $u_0(\cdot)$  derart, dass wir

$$u(x, t) = \sin\left(\pi \cdot \frac{x}{t + 0.5}\right) \quad (4.26)$$

als exakte Lösung des Problems erhalten, vgl. Abbildung 4.4.

Tabelle 4.1 zeigt die entsprechenden Resultate der unterschiedlichen Diskretisierungsverfahren. Wir betrachten zunächst das DG1 Verfahren. Die Fehlerreduktionsraten bzgl. der  $H_{mix}^{1,0}$ -Norm und der  $H_{mix}^{(1,2),(0,\infty)}$ -Norm betragen  $\rho = 0.5$ , was dem in (4.23) beschriebenen Verhalten des Interpolationsfehlers des zugrunde liegenden Dünngitters entspricht. So ist der Fehler  $9.126_{-3}$  bzgl. der  $H_{mix}^{(1,2),(0,\infty)}$  auf Level  $l = 8$  nur noch halb so groß wie der Fehler  $1.824_{-2}$  auf Level  $l = 7$ .

Für die Fehler in der  $L^\infty$ - und der  $L^2$ -Norm beobachten wir ein leichtes Abnehmen der Fehlerreduktionsrate mit wachsendem Level  $l$ , wobei sich die Rate dem Wert  $\rho = 0.25$  zu nähern scheint. So haben wir etwa im Fall der  $L^\infty$ -Norm auf Level  $l = 5$  einen Fehler  $e_{L^\infty} = 7.546_{-3}$  und damit eine Reduktion um den Faktor 0.33 verglichen mit dem Fehler  $e_{L^\infty} = 2.292_{-2}$  auf Level  $l = 4$ , während die Reduktion von Level  $l = 7$  auf Level  $l = 8$   $\rho_{L^\infty} = 0.28$  beträgt. Dieses Verhalten deutet auf den logarithmischen Faktor  $(l+1)/l$  hin, der auch für die Reduktion des Interpolationsfehlers in Erscheinung tritt, vgl. (4.23). Mit wachsendem Level nähert sich dieser Faktor dem Wert 1 und die Reduktionsrate nähert sich somit dem Wert  $\rho = 0.25$ .

Betrachten wir nun die Ergebnisse des Crank-Nicolson Verfahrens, so ist das gleiche Verhalten wie bei dem DG1 Verfahren beobachtbar. Insbesondere verhalten sich auch hier die Diskretisierungsfehler wie die Interpolationsfehler. Wir erhalten Fehlerreduktionsraten von 0.5 für die Normen, die den Gradienten im Ort enthalten, während sich für die beiden anderen Normen die Raten dem Wert  $\rho = 0.25$  nähern. Allerdings fällt auf, dass die Fehler gemessen in der  $L^\infty$ - und der  $L^2$ -Norm

kleiner sind als die entsprechenden Fehler für das DG1 Verfahren. So ist auf Level  $l = 8$  der Fehler gemessen in der  $L^\infty$ -Norm mit  $1.988_{-4}$  für das DG1 Verfahren mehr als doppelt so groß wie der entsprechende Fehler des CN Verfahrens, der  $7.303_{-5}$  beträgt. Dies ist vor allem insofern von Bedeutung, da der Ansatzraum der stückweise linearen, unstetigen Funktionen doppelt so viele Freiheitsgrade enthält wie der Raum der stückweise linearen, stetigen Ansatzfunktionen. Damit besitzt der entsprechende Dünngitterraum im DG1 Fall doppelt so viele Freiheitsgrade wie der Dünngitterraum des Crank-Nicolson Verfahrens, führt aber dennoch zu einem wesentlich größeren relativen Fehler.

Für die Eulerdiskretisierung beobachten wir, dass sich die Reduktionsraten für alle betrachteten Normen mit wachsendem Level dem Wert 0.5 nähern. Hierbei dominiert offensichtlich die Fehlerreduktionsrate der stückweise konstanten Ansatzfunktionen in der Zeit den Gesamtfehler, so dass wir für die  $L^2$ - und  $L^\infty$ -Normen keine besseren Raten als im Fall der  $H_{mix}^{1,0}$ - bzw. der  $H_{mix}^{(1,2),(0,\infty)}$ -Normen erhalten. Dies deckt sich mit den Abschätzungen (4.24) für die Interpolationsfehler, die einen Abfall von  $2^{-l}$  sowohl für die  $L^2$ -Norm als auch für die  $H_{mix}^{1,0}$ -Norm zeigen. Dieser Effekt wird durch eine feinere Zeitauflösung (auf Kosten der Anzahl an Unbekannten) bei dem Dünngitterraum  $V_l^0$  ausgeglichen. Somit erhalten wir in der Tat für das anisotrope Euler Verfahren eine Fehlerreduktion von 0.5, unabhängig von dem Level  $l$ , in der  $H_{mix}^{1,0}$ - und  $H_{mix}^{(1,2),(0,\infty)}$ -Norm, während sich die Fehlerreduktionsrate bezüglich der beiden anderen Normen dem Wert  $\rho = 0.25$  nähert.

Wir vergleichen nun die Verfahren im Hinblick auf die Kosten-Nutzen Relation. Dabei verstehen wir unter dem Kosten-Nutzen Verhältnis das Verhältnis der benötigten Anzahl von Freiheitsgraden zu der erzielten Genauigkeit. Die Abbildung 4.5 zeigt dazu die Anzahl der Freiheitsgrade in Relation zu den Fehlern in der  $L^2$ - und der  $H_{mix}^{1,0}$ -Norm. Im Falle der  $L^2$ -Norm sehen wir deutlich, dass das Crank-Nicolson Verfahren zu dem kleinsten Fehler pro Unbekannter führt und damit das beste Kosten-Nutzen Verhältnis liefert. Wie wir auch aus der obigen Diskussion erwarten würden, verläuft der zum DG1 Verfahren mit stückweise linearen Funktionen in der Zeit gehörige Graph nahezu parallel zu dem Graphen des Crank-Nicolson Verfahrens. Interessant ist der Vergleich zwischen der Euler und der anisotropen Euler Diskretisierung. Während für die  $L^2$ -Norm die Euler Diskretisierung und das anisotrope Euler Verfahren nahezu das gleiche Kosten-Nutzen Verhältnis aufweisen (die entsprechenden Graphen sind fast identisch), so ist der Fehler bzgl. der  $H_{mix}^{1,0}$ -Norm im Verhältnis zu der Anzahl der Unbekannten bei dem Euler Verfahren wesentlich geringer als bei dem anisotropen Euler Verfahren. Hier wird also im Fall der  $L^2$ -Norm die wesentlich größere Dimension des Raumes  $V_l^0$  im Vergleich zu  $\tilde{V}_l^0$  durch eine bessere Fehlerreduktion kompensiert, während wir im Fall der  $H_{mix}^{1,0}$ -Norm die gleichen Reduktionsraten erhalten und somit das gesamte Kosten-Nutzen Verhältnis für  $V_l^0$  schlechter ist als für  $\tilde{V}_l^0$ .

*Beispiel 4.2.* Analog zu dem vorigen Beispiel 4.1 betrachten wir nun das zwei-

**Tabelle 4.2.** Konvergenz der verschiedenen Diskretisierungsverfahren für das zweidimensionale Problem (4.27) mit Lösung  $u(x, t) = \sin(\pi \cdot \frac{x_1}{t+0.5}) \sin(\pi \cdot \frac{x_2}{t+0.5})$ .

l	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H_{mix}^{(1,2),(0,\infty)}}$	$\rho_{H_{mix}^{(1,2),(0,\infty)}}$	$e_{H_{mix}^{1,0}}$	$\rho_{H_{mix}^{1,0}}$
Crank-Nicolson								
4	3.100 <sub>-3</sub>	-	7.043 <sub>-4</sub>	-	1.715 <sub>-2</sub>	-	8.187 <sub>-3</sub>	-
5	1.086 <sub>-3</sub>	0.35	2.526 <sub>-4</sub>	0.36	8.959 <sub>-3</sub>	0.52	3.973 <sub>-3</sub>	0.49
6	3.652 <sub>-4</sub>	0.34	8.482 <sub>-5</sub>	0.34	4.531 <sub>-3</sub>	0.51	1.947 <sub>-3</sub>	0.49
7	1.129 <sub>-4</sub>	0.31	2.689 <sub>-5</sub>	0.32	2.272 <sub>-3</sub>	0.50	9.636 <sub>-4</sub>	0.49
8	3.323 <sub>-5</sub>	0.29	8.175 <sub>-6</sub>	0.30	1.136 <sub>-3</sub>	0.50	4.798 <sub>-4</sub>	0.50
9	9.508 <sub>-6</sub>	0.29	2.411 <sub>-6</sub>	0.29	5.680 <sub>-4</sub>	0.50	2.396 <sub>-4</sub>	0.50
10	2.703 <sub>-6</sub>	0.28	6.954 <sub>-7</sub>	0.29	2.840 <sub>-4</sub>	0.50	1.197 <sub>-4</sub>	0.50
DG1								
4	3.652 <sub>-3</sub>	-	9.779 <sub>-4</sub>	-	1.352 <sub>-2</sub>	-	7.540 <sub>-3</sub>	-
5	1.246 <sub>-3</sub>	0.34	3.384 <sub>-4</sub>	0.35	7.164 <sub>-3</sub>	0.53	3.367 <sub>-3</sub>	0.45
6	4.056 <sub>-4</sub>	0.33	1.083 <sub>-4</sub>	0.32	3.637 <sub>-3</sub>	0.51	1.557 <sub>-3</sub>	0.46
7	1.239 <sub>-4</sub>	0.31	3.301 <sub>-5</sub>	0.30	1.825 <sub>-3</sub>	0.50	7.478 <sub>-4</sub>	0.48
8	3.646 <sub>-5</sub>	0.29	9.737 <sub>-6</sub>	0.29	9.128 <sub>-4</sub>	0.50	3.675 <sub>-4</sub>	0.49
9	1.047 <sub>-5</sub>	0.29	2.806 <sub>-6</sub>	0.29	4.564 <sub>-4</sub>	0.50	1.825 <sub>-4</sub>	0.50
10	2.956 <sub>-6</sub>	0.28	7.946 <sub>-7</sub>	0.28	2.282 <sub>-4</sub>	0.50	9.102 <sub>-5</sub>	0.50
Euler								
4	5.152 <sub>-2</sub>	-	1.399 <sub>-2</sub>	-	1.395 <sub>-1</sub>	-	8.176 <sub>-2</sub>	-
5	2.956 <sub>-2</sub>	0.57	7.417 <sub>-3</sub>	0.53	7.927 <sub>-2</sub>	0.57	4.652 <sub>-2</sub>	0.57
6	1.843 <sub>-2</sub>	0.62	3.787 <sub>-3</sub>	0.51	4.682 <sub>-2</sub>	0.59	2.544 <sub>-2</sub>	0.55
7	1.059 <sub>-2</sub>	0.57	1.907 <sub>-3</sub>	0.50	2.770 <sub>-2</sub>	0.59	1.362 <sub>-2</sub>	0.54
8	5.762 <sub>-3</sub>	0.54	9.551 <sub>-4</sub>	0.50	1.569 <sub>-2</sub>	0.57	7.208 <sub>-3</sub>	0.53
9	3.026 <sub>-3</sub>	0.53	4.774 <sub>-4</sub>	0.50	8.637 <sub>-3</sub>	0.55	3.785 <sub>-3</sub>	0.53
10	1.556 <sub>-3</sub>	0.51	2.385 <sub>-4</sub>	0.50	4.668 <sub>-3</sub>	0.54	1.977 <sub>-3</sub>	0.52
Euler anisotrop								
4	3.543 <sub>-2</sub>	-	8.461 <sub>-3</sub>	-	1.018 <sub>-1</sub>	-	5.504 <sub>-2</sub>	-
5	1.353 <sub>-2</sub>	0.38	2.793 <sub>-3</sub>	0.33	5.417 <sub>-2</sub>	0.53	2.397 <sub>-2</sub>	0.44
6	4.603 <sub>-3</sub>	0.34	8.680 <sub>-4</sub>	0.31	2.741 <sub>-2</sub>	0.51	1.099 <sub>-2</sub>	0.46
7	1.455 <sub>-3</sub>	0.32	2.597 <sub>-4</sub>	0.30	1.371 <sub>-2</sub>	0.50	5.274 <sub>-3</sub>	0.48
8	4.393 <sub>-4</sub>	0.30	7.563 <sub>-5</sub>	0.29	6.844 <sub>-3</sub>	0.50	2.595 <sub>-3</sub>	0.49

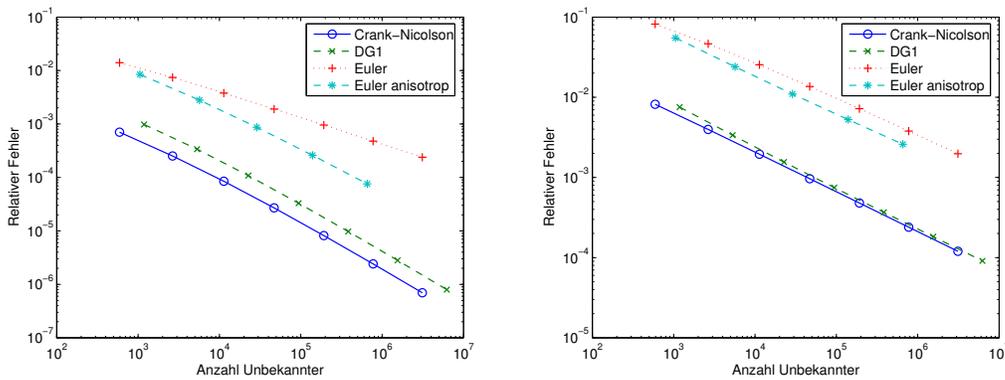
dimensionale Problem

$$\begin{aligned}
 \partial_t u - \Delta u &= f \text{ für alle } x \in \Omega = (0, 1)^2, \\
 u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega, \\
 u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega.
 \end{aligned} \tag{4.27}$$

Wir setzen die rechte Seite  $f$  und die Anfangsbedingung  $u_0(\cdot)$  so, dass das Problem die exakte Lösung

$$u(x, t) = \sin(\pi \cdot \frac{x_1}{t+0.5}) \sin(\pi \cdot \frac{x_2}{t+0.5}) \tag{4.28}$$

besitzt.



**Abbildung 4.6.** Konvergenzhistorie von Problem (4.27) bzgl. der  $L^2$ - (links) und der  $H_{mix}^{1,0}$ -Norm (rechts) für das CN, das DG1, das Euler und das anisotrope Euler Verfahren.

Die entsprechenden Ergebnisse sind in Tabelle 4.2 zusammengefasst. Wie in dem vorherigen Beispiel mit nur einer Ortsdimension beobachten wir für das DG1 Verfahren und das Crank-Nicolson Verfahren Fehlerreduktionen von  $\rho = 0.5$  im Falle der  $H_{mix}^{1,0}$ - und der  $H_{mix}^{(1,2),(0,\infty)}$ -Normen. Bezüglich der  $L^2$ - und der  $L^\infty$ -Normen sehen wir hier erneut einen Abfall der Reduktionsraten mit steigenden Levelzahlen. Im Fall des Crank-Nicolson Verfahrens erhalten wir beispielsweise für Level  $l = 5$  den Fehler  $e_{L^2} = 2.526_{-4}$  und damit eine Reduktion von  $\rho_{L^2} = 0.36$  des Fehlers  $e_{L^2} = 7.043_{-4}$  auf Level  $l = 4$ , während der Fehler  $e_{L^2} = 2.411_{-6}$  von Level  $l = 9$  um einen Faktor 0.29 auf einen Fehler  $e_{L^2} = 6.954_{-7}$  auf Level  $l = 10$  reduziert wird.

Wie auch im eindimensionalen Fall nähert sich die Fehlerreduktion des Euler Verfahrens mit wachsendem Level einer Fehlerreduktion von  $\rho = 0.5$  in allen Normen an. Verwenden wir hingegen das anisotrope Euler Verfahren, bei dem die Zeit mit doppelt so großer Levelzahl wie der Ort aufgelöst wird, so sehen wir, dass sich die Fehlerreduktionen  $\rho_{L^2}$  und  $\rho_{L^\infty}$  mit wachsendem Level auch im zweidimensionalen Fall dem asymptotischen Wert 0.25 annähern.

In Abbildung 4.6 sind die Fehler der Diskretisierungsverfahren bzgl. der  $L^2$ - und der  $H_{mix}^{1,0}$ -Normen in Abhängigkeit von der Anzahl der Unbekannten dargestellt. Die Crank-Nicolson und die DG1 Diskretisierung verhalten sich dabei sehr ähnlich, d.h. die Kurven besitzen nahezu die gleiche Steigung. Hierbei liefert das Crank-Nicolson Verfahren im Falle der  $L^2$ -Norm einen etwas geringeren Fehler pro Unbekannter als das DG1 Verfahren. Man beachte, dass bereits im zweidimensionalen Fall das Euler Verfahren mit  $V_l^0$  ein besseres Kosten-Nutzen Verhältnis als das Euler Verfahren mit  $\tilde{V}_l^0$  aufweist. Dies gilt sowohl bzgl. der  $L^2$ -Norm als auch bzgl. der  $H_{mix}^{1,0}$ -Norm und ist darauf zurückzuführen, dass die Dimension des Raumes  $V_l^0$  in zwei Dimensionen lediglich um einen logarithmischen Faktor schlechter ist als die Dimension des

**Tabelle 4.3.** Konvergenz der verschiedenen Diskretisierungsverfahren für das dreidimensionale Problem (4.29) mit Lösung  $u(x, t) = \prod_{i=1}^3 \sin(\pi \cdot \frac{x_i}{t+0.5})$ .

l	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H_{mix}^{(1,2),(0,\infty)}}$	$\rho_{H_{mix}^{(1,2),(0,\infty)}}$	$e_{H_{mix}^{1,0}}$	$\rho_{H_{mix}^{1,0}}$
Crank-Nicolson								
3	1.220 <sub>-2</sub>	-	1.869 <sub>-3</sub>	-	2.389 <sub>-2</sub>	-	1.559 <sub>-2</sub>	-
4	4.901 <sub>-3</sub>	0.40	7.752 <sub>-4</sub>	0.41	1.452 <sub>-2</sub>	0.61	7.652 <sub>-3</sub>	0.49
5	1.635 <sub>-3</sub>	0.33	2.867 <sub>-4</sub>	0.37	7.714 <sub>-3</sub>	0.53	3.623 <sub>-3</sub>	0.47
6	5.506 <sub>-4</sub>	0.34	9.619 <sub>-5</sub>	0.34	3.921 <sub>-3</sub>	0.51	1.730 <sub>-3</sub>	0.48
7	1.704 <sub>-4</sub>	0.31	3.032 <sub>-5</sub>	0.32	1.967 <sub>-3</sub>	0.50	8.437 <sub>-4</sub>	0.49
DG1								
3	1.034 <sub>-2</sub>	-	2.349 <sub>-3</sub>	-	2.214 <sub>-2</sub>	-	1.608 <sub>-2</sub>	-
4	4.898 <sub>-3</sub>	0.47	9.670 <sub>-4</sub>	0.41	1.134 <sub>-2</sub>	0.51	7.383 <sub>-3</sub>	0.46
5	1.721 <sub>-3</sub>	0.35	3.460 <sub>-4</sub>	0.36	6.154 <sub>-3</sub>	0.54	3.197 <sub>-3</sub>	0.43
6	5.606 <sub>-4</sub>	0.33	1.126 <sub>-4</sub>	0.33	3.147 <sub>-3</sub>	0.51	1.418 <sub>-3</sub>	0.44
7	-	-	-	-	-	-	-	-
Euler								
3	1.105 <sub>-1</sub>	-	2.198 <sub>-2</sub>	-	2.455 <sub>-1</sub>	-	1.348 <sub>-1</sub>	-
4	6.705 <sub>-2</sub>	0.61	1.321 <sub>-2</sub>	0.60	1.516 <sub>-1</sub>	0.62	8.413 <sub>-2</sub>	0.62
5	3.965 <sub>-2</sub>	0.59	7.043 <sub>-3</sub>	0.53	8.307 <sub>-2</sub>	0.55	4.722 <sub>-2</sub>	0.56
6	2.534 <sub>-2</sub>	0.64	3.599 <sub>-3</sub>	0.51	4.482 <sub>-2</sub>	0.54	2.545 <sub>-2</sub>	0.54
7	1.495 <sub>-2</sub>	0.59	1.813 <sub>-3</sub>	0.50	2.438 <sub>-2</sub>	0.54	1.347 <sub>-2</sub>	0.53
Euler anisotrop								
3	1.084 <sub>-1</sub>	-	2.122 <sub>-2</sub>	-	2.353 <sub>-1</sub>	-	1.303 <sub>-1</sub>	-
4	4.783 <sub>-2</sub>	0.44	8.162 <sub>-3</sub>	0.38	8.902 <sub>-2</sub>	0.38	5.615 <sub>-2</sub>	0.43
5	1.889 <sub>-2</sub>	0.39	2.757 <sub>-3</sub>	0.34	4.688 <sub>-2</sub>	0.53	2.321 <sub>-2</sub>	0.41
6	6.567 <sub>-3</sub>	0.35	8.692 <sub>-4</sub>	0.32	2.384 <sub>-2</sub>	0.51	1.008 <sub>-2</sub>	0.43

Raumes  $\tilde{V}_l^0$ , während wir bessere Approximationseigenschaften erhalten. Insbesondere weist die Kurve des anisotropen Euler Verfahrens bzgl. der  $L^2$ -Norm nahezu die gleiche Steigung auf, wie die Kurven des CN und des DG1 Verfahrens, obwohl der Fehler im Vergleich zu diesen beiden Verfahren immer noch größer ist.

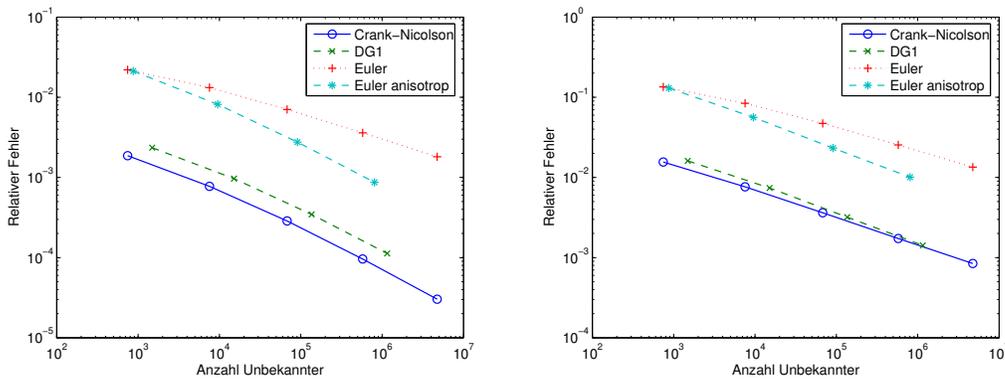
*Beispiel 4.3.* Wir betrachten nun das dreidimensionale Problem

$$\begin{aligned}
 \partial_t u - \Delta u &= f \text{ für alle } x \in \Omega = (0, 1)^3, \\
 u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega, \\
 u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega.
 \end{aligned} \tag{4.29}$$

Wir setzen die rechte Seite  $f$  und die Anfangsbedingung  $u_0(\cdot)$  so, dass wir

$$u(x, t) = \prod_{i=1}^3 \sin\left(\pi \cdot \frac{x_i}{t+0.5}\right)$$

als Lösung erhalten. Wie in den zwei vorherigen Beispielen so ist auch hier feststellbar, dass sich die Fehlerreduktionsraten  $\rho_{L^2}$  und  $\rho_{L^\infty}$  des Crank-Nicolson und des



**Abbildung 4.7.** Konvergenzhistorie für die CN, die DG1, die Euler und die anisotrope Euler Diskretisierung des Problems (4.29) bzgl. der  $L^2$ - (links) und der  $H_{mix}^{1,0}$ -Norm (rechts).

DG1 Verfahrens mit wachsendem Level verbessern, während wir Reduktionsraten  $\rho \approx 0.5$  bzgl. der  $H_{mix}^{(1,2),(0,\infty)}$ - und  $H_{mix}^{1,0}$ -Normen erhalten. Das gleiche Verhalten beobachten wir für das anisotrope Euler Verfahren, allerdings sind hier die Fehler im Vergleich zu dem Crank-Nicolson und dem DG1 Verfahren um eine Größenordnung schlechter.

Besonders interessant ist dabei die Untersuchung der Kosten-Nutzen Relation in Abbildung 4.7. Wie wir in Lemma 2.3 gezeigt haben, besitzt die Dimension der Räume  $V_l^0$  und  $V_l^0$  in drei Ortsdimensionen die gleiche Komplexität bzgl. des gewählten Levels. Da wir jedoch ähnliche Reduktionsraten für das anisotrope Euler Verfahren wie für die Crank-Nicolson oder die DG1 Diskretisierung erhalten, stimmen die Steigungen der in Abbildung 4.7 dargestellten Kurven dieser drei Verfahren nahezu überein. Damit liefert das anisotrope Euler Verfahren sowohl in der  $L^2$  als auch in der  $H_{mix}^{1,0}$ -Norm mit wesentlich weniger Freiheitsgraden den gleichen Fehler wie die Euler Diskretisierung und der Betrag der Steigung der entsprechenden Kurve ist wesentlich größer.

*Beispiel 4.4.* Wie bereits in Kapitel 2 diskutiert wurde, besteht ein wesentlicher Vorteil der Raum-Zeit Dünngitter gegenüber den klassischen Dünngitterräumen in der einfacheren Handhabung komplizierter Geometrien. Da wir im Ort die ( $d$ -) lineare hierarchische Basis verwenden, benötigen wir zur Konstruktion der Raum-Zeit Dünngitter lediglich eine Sequenz von ineinander geschachtelten Finite Element Räumen, die heute von vielen Finite Element Paketen im Zusammenhang mit geometrischen Mehrgitterlösern bereitgestellt wird.

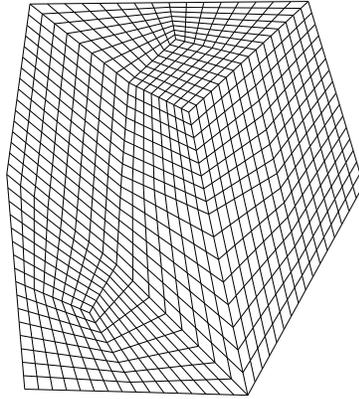


Abbildung 4.8. In Beispiel 4.4 verwendete Geometrie mit einem Beispiel-Gitter.

Tabelle 4.4. Ergebnisse für Beispiel 4.4 mit der exakten Lösung  $u(x, t) = \sin(\pi \cdot \frac{x_1}{t+0.5}) \sin(\pi \cdot \frac{x_2}{t+0.5})$  auf dem in Abb. 4.8 dargestellten Gebiet für verschiedene Anzahlen an Unbekannten (Levels).

dof	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H_{mix}^{(1,2),(0,\infty)}}$	$\rho_{H_{mix}^{(1,2),(0,\infty)}}$	$e_{H_{mix}^{1,0}}$	$\rho_{H_{mix}^{1,0}}$
Crank-Nicolson								
5613	3.829 <sub>-3</sub>	-	7.469 <sub>-4</sub>	-	2.469 <sub>-2</sub>	-	1.333 <sub>-2</sub>	-
23532	9.401 <sub>-4</sub>	0.25	1.829 <sub>-4</sub>	0.24	1.317 <sub>-2</sub>	0.53	6.090 <sub>-3</sub>	0.46
95595	2.564 <sub>-4</sub>	0.27	5.010 <sub>-5</sub>	0.27	6.713 <sub>-3</sub>	0.51	2.924 <sub>-3</sub>	0.48
383850	7.057 <sub>-5</sub>	0.28	1.471 <sub>-5</sub>	0.29	3.376 <sub>-3</sub>	0.50	1.439 <sub>-3</sub>	0.49
1535337	2.192 <sub>-5</sub>	0.31	4.398 <sub>-6</sub>	0.30	1.691 <sub>-3</sub>	0.50	7.157 <sub>-4</sub>	0.50
DG1								
8313	4.682 <sub>-3</sub>	-	1.039 <sub>-3</sub>	-	1.945 <sub>-2</sub>	-	1.218 <sub>-2</sub>	-
35526	1.344 <sub>-3</sub>	0.29	3.690 <sub>-4</sub>	0.36	1.054 <sub>-2</sub>	0.54	5.288 <sub>-3</sub>	0.43
145650	4.126 <sub>-4</sub>	0.31	1.186 <sub>-4</sub>	0.32	5.395 <sub>-3</sub>	0.51	2.388 <sub>-3</sub>	0.45
587229	1.198 <sub>-4</sub>	0.29	3.543 <sub>-5</sub>	0.30	2.714 <sub>-3</sub>	0.50	1.130 <sub>-3</sub>	0.47
2352711	3.449 <sub>-5</sub>	0.29	1.013 <sub>-5</sub>	0.29	1.359 <sub>-3</sub>	0.50	5.513 <sub>-4</sub>	0.49
Euler								
5613	3.056 <sub>-2</sub>	-	6.703 <sub>-3</sub>	-	1.242 <sub>-1</sub>	-	7.455 <sub>-2</sub>	-
23532	1.729 <sub>-2</sub>	0.57	3.654 <sub>-3</sub>	0.55	8.656 <sub>-2</sub>	0.70	4.616 <sub>-2</sub>	0.62
95595	8.659 <sub>-3</sub>	0.50	1.849 <sub>-3</sub>	0.51	5.607 <sub>-2</sub>	0.65	2.694 <sub>-2</sub>	0.58
383850	4.204 <sub>-3</sub>	0.49	9.166 <sub>-4</sub>	0.50	3.400 <sub>-2</sub>	0.61	1.520 <sub>-2</sub>	0.56
1535337	2.051 <sub>-3</sub>	0.49	4.543 <sub>-4</sub>	0.50	1.974 <sub>-2</sub>	0.58	8.380 <sub>-3</sub>	0.55

Wir betrachten das zu dem Beispiel 4.2 analoge Problem

$$\begin{aligned}
 \partial_t u - \Delta u &= f \text{ für alle } x \in \Omega, \\
 u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega, \\
 u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega,
 \end{aligned} \tag{4.30}$$

wobei wir hier nicht  $\Omega = (0, 1)^2$  wählen, sondern das in Abbildung 4.8 mit einem

Beispielgitter dargestellte kompliziertere Gebiet  $\Omega$  verwenden. Hierbei konstruieren wir durch sukzessive uniforme Verfeinerung eines groben Ausgangsgitters die feineren Gitter. Wir setzen die rechte Seite  $f$  und die Anfangsbedingung  $u_0(\cdot)$  so, dass wir als exakte Lösung erneut

$$u(x, t) = \sin\left(\pi \cdot \frac{x_1}{t + 0.5}\right) \sin\left(\pi \cdot \frac{x_2}{t + 0.5}\right) \quad (4.31)$$

erhalten.

Da wir in den letzten Beispielen bereits ausführlich die Euler Diskretisierung mit der anisotropen Euler Diskretisierung verglichen haben, verzichten wir im Folgenden auf die Präsentation der Ergebnisse der anisotropen Euler Diskretisierung.

In Tabelle 4.4 sind die Ergebnisse für die verschiedenen Diskretisierungsverfahren zusammen mit der dazugehörigen Zahl an Freiheitsgraden aufgeführt. Erneut beobachten wir für das DG1 und das Crank-Nicolson Verfahren sowohl bzgl. der  $H_{mix}^{(1,2),(0,\infty)}$ - als auch bzgl. der  $H_{mix}^{1,0}$ -Norm eine Fehlerreduktion von 0.5. Im Fall der  $L^\infty$ - und  $L^2$ -Normen beobachten wir für das Crank-Nicolson Verfahren, dass die Konvergenzraten auf den kleineren Leveln zunächst bereits  $\rho = 0.25$  betragen, sich dann aber mit wachsendem Level etwas verschlechtern. So erhalten wir von Level  $l = 4$  auf Level  $l = 5$  eine Fehlerreduktion von  $\rho_{L^\infty} = 0.31$  während von Level  $l = 1$  auf Level  $l = 2$  eine Reduktion von  $\rho_{L^\infty} = 0.25$  feststellbar ist. Für das DG1 Verfahren beobachten wir das gleiche Verhalten wie in den vorherigen Beispielen: Die Konvergenzrate bzgl. der  $L^2$ - und der  $L^\infty$ -Norm nähert sich mit zunehmendem Level dem Wert 0.25 an.

Im Fall des Euler Verfahrens erhalten wir ähnliche Resultate wie in den früheren Beispielen. Während der Fehler bzgl. der  $L^\infty$ - und der  $L^2$ -Norm mit 0.5 unabhängig vom Level abfällt, verbessern sich die Fehlerreduktionen für die anderen beiden Normen mit zunehmenden Leveln. So ist beispielsweise eine Fehlerreduktion  $\rho_{H_{mix}^{(1,2),(0,\infty)}} = 0.70$  von Level  $l = 1$  auf Level  $l = 2$  und  $\rho_{H_{mix}^{(1,2),(0,\infty)}} = 0.58$  von  $l = 3$  auf Level  $l = 4$  zu konstatieren.

*Beispiel 4.5.* Um das Unidirektionale Prinzip im Fall klassischer Dünngitter effizient anwenden zu können, ohne dass dabei die Konvergenzordnung der Diskretisierung verloren geht, muss der zugrunde liegende Differentialoperator im Ort Tensorproduktstruktur aufweisen. Damit ist die Diskretisierung beliebiger Differentialoperatoren mit variablen Koeffizienten im klassischen Dünngitterfall sehr schwierig. Wie in Abschnitt 4.3 diskutiert, benötigen wir zur Anwendung des Unidirektionalen Prinzips auf die Raum-Zeit Dünngitterdiskretisierungen lediglich eine Tensorproduktstruktur des Differentialoperators bezüglich Raum und Zeit. Wir können somit ohne weiteres parabolische Probleme behandeln, bei denen der elliptische Anteil des parabolischen Operators beliebige in der Zeit konstante aber im Ort variable Koeffizienten enthält. Als Beispiel hierfür betrachten wir nun das zweidimensionale

**Tabelle 4.5.** Numerische Ergebnisse für das zweidimensionale Problem (4.32) mit variablen Koeffizienten und der exakten Lösung  $u(x, t) = \sin(\pi \cdot \frac{x_1}{t+0.5}) \sin(\pi \cdot \frac{x_2}{t+0.5})$  auf  $\Omega = (0, 1)^2$ .

1	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{a(\cdot, \cdot) \otimes L^\infty}$	$\rho_{a(\cdot, \cdot) \otimes L^\infty}$	$e_{a(\cdot, \cdot) \otimes L^2}$	$\rho_{a(\cdot, \cdot) \otimes L^2}$
Crank-Nicolson								
4	3.175 <sub>-3</sub>	-	7.191 <sub>-4</sub>	-	1.684 <sub>-2</sub>	-	8.069 <sub>-3</sub>	-
5	1.122 <sub>-3</sub>	0.35	2.555 <sub>-4</sub>	0.36	8.814 <sub>-3</sub>	0.52	3.908 <sub>-3</sub>	0.48
6	3.837 <sub>-4</sub>	0.34	8.652 <sub>-5</sub>	0.34	4.460 <sub>-3</sub>	0.51	1.916 <sub>-3</sub>	0.49
7	1.184 <sub>-4</sub>	0.31	2.767 <sub>-5</sub>	0.32	2.236 <sub>-3</sub>	0.50	9.486 <sub>-4</sub>	0.50
8	3.490 <sub>-5</sub>	0.29	8.469 <sub>-6</sub>	0.31	1.118 <sub>-3</sub>	0.50	4.724 <sub>-4</sub>	0.50
9	1.002 <sub>-5</sub>	0.29	2.511 <sub>-6</sub>	0.30	5.592 <sub>-4</sub>	0.50	2.358 <sub>-4</sub>	0.50
10	2.806 <sub>-6</sub>	0.28	7.391 <sub>-7</sub>	0.29	2.796 <sub>-4</sub>	0.50	1.179 <sub>-4</sub>	0.50
DG1								
4	3.763 <sub>-3</sub>	-	9.838 <sub>-4</sub>	-	1.326 <sub>-2</sub>	-	7.382 <sub>-3</sub>	-
5	1.286 <sub>-3</sub>	0.34	3.417 <sub>-4</sub>	0.35	7.048 <sub>-3</sub>	0.53	3.305 <sub>-3</sub>	0.45
6	4.194 <sub>-4</sub>	0.33	1.103 <sub>-4</sub>	0.32	3.580 <sub>-3</sub>	0.51	1.531 <sub>-3</sub>	0.46
7	1.286 <sub>-4</sub>	0.31	3.387 <sub>-5</sub>	0.31	1.796 <sub>-3</sub>	0.50	7.362 <sub>-4</sub>	0.48
8	3.795 <sub>-5</sub>	0.30	1.005 <sub>-5</sub>	0.30	8.986 <sub>-4</sub>	0.50	3.618 <sub>-4</sub>	0.49
9	1.092 <sub>-5</sub>	0.29	2.910 <sub>-6</sub>	0.29	4.493 <sub>-4</sub>	0.50	1.797 <sub>-4</sub>	0.50
10	3.200 <sub>-6</sub>	0.29	8.278 <sub>-7</sub>	0.28	2.212 <sub>-4</sub>	0.50	8.985 <sub>-5</sub>	0.50

Problem

$$\begin{aligned} \partial_t u - \operatorname{div}(A \nabla u) &= f \text{ für alle } x \in \Omega = (0, 1)^2, \\ u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega, \\ u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega, \end{aligned} \quad (4.32)$$

mit

$$A = \begin{pmatrix} \exp(-(x_1 - 0.5)^2 - (x_2 - 0.5)^2) & 0 \\ 0 & 1 \end{pmatrix}.$$

Hierbei wählen wir die rechte Seite  $f$ , die Anfangs- und die Randbedingung so, dass wir erneut

$$u(x, t) = \sin\left(\pi \cdot \frac{x_1}{t+0.5}\right) \sin\left(\pi \cdot \frac{x_2}{t+0.5}\right)$$

als die Lösung des Problems erhalten. Wir ersetzen in der Fehlermessung die  $H^1$ -Norm durch die von dem Differentialoperator induzierte Energienorm

$$\|v\|_a^2 := \int_{\Omega} (\nabla v)^T A \nabla v \, dx$$

und erhalten dementsprechend

$$\rho_{a(\cdot, \cdot) \otimes L^\infty} := \frac{\|\hat{e}_{t+1}\|_{a(\cdot, \cdot) \otimes L^\infty}}{\|\hat{e}_t\|_{a(\cdot, \cdot) \otimes L^\infty}}, \quad \rho_{a(\cdot, \cdot) \otimes L^2} := \frac{\|\hat{e}_{t+1}\|_{a(\cdot, \cdot) \otimes L^2}}{\|\hat{e}_t\|_{a(\cdot, \cdot) \otimes L^2}}.$$

**Tabelle 4.6.** Numerische Ergebnisse für die Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung des zweidimensionalen Problems (4.2) mit exakter Lösung  $u(x, t) = \sin(\pi \cdot \frac{x_1}{t+0.5}) \sin(\pi \cdot \frac{x_2}{t+0.5})$  auf  $\Omega = (0, 1)^2$  bei Verwendung einer biquadratischen hierarchischen Basis im Ort.

l	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H_{mix}^{(1,2),(0,\infty)}}$	$\rho_{H_{mix}^{(1,2),(0,\infty)}}$	$e_{H_{mix}^{1,0}}$	$\rho_{H_{mix}^{1,0}}$
Crank-Nicolson, Ansatzraum $V_l^0$								
3	4.144 <sub>-3</sub>	-	5.522 <sub>-3</sub>	-	2.523 <sub>-2</sub>	-	2.395 <sub>-2</sub>	-
4	1.037 <sub>-3</sub>	0.25	1.389 <sub>-3</sub>	0.25	6.566 <sub>-3</sub>	0.26	6.081 <sub>-3</sub>	0.25
5	2.584 <sub>-4</sub>	0.25	3.475 <sub>-4</sub>	0.25	1.676 <sub>-3</sub>	0.26	1.539 <sub>-3</sub>	0.25
6	6.459 <sub>-5</sub>	0.25	8.661 <sub>-5</sub>	0.25	4.229 <sub>-4</sub>	0.25	3.880 <sub>-4</sub>	0.25
7	1.623 <sub>-5</sub>	0.25	2.091 <sub>-5</sub>	0.24	1.043 <sub>-4</sub>	0.25	9.548 <sub>-5</sub>	0.25
Crank-Nicolson, Ansatzraum $V_l^0$								
3	5.577 <sub>-4</sub>	-	6.086 <sub>-4</sub>	-	2.240 <sub>-2</sub>	-	1.599 <sub>-2</sub>	-
4	7.167 <sub>-5</sub>	0.13	6.517 <sub>-5</sub>	0.11	5.600 <sub>-3</sub>	0.25	3.995 <sub>-3</sub>	0.25
5	9.566 <sub>-6</sub>	0.13	7.583 <sub>-6</sub>	0.12	1.400 <sub>-3</sub>	0.25	9.991 <sub>-4</sub>	0.25
6	1.212 <sub>-6</sub>	0.13	9.495 <sub>-7</sub>	0.13	3.493 <sub>-4</sub>	0.25	2.494 <sub>-4</sub>	0.25
7	1.495 <sub>-7</sub>	0.12	1.183 <sub>-7</sub>	0.12	8.473 <sub>-5</sub>	0.24	6.064 <sub>-5</sub>	0.24

In Tabelle 4.5 sind die Ergebnisse für das Crank-Nicolson und das DG1 Verfahren dargestellt.

Wir erhalten das gleiche Konvergenzverhalten wie in den vorherigen Beispielen. Der Fehler gemessen in der  $\|\cdot\|_{a(\cdot,\cdot)\otimes L^\infty}$ - und der  $\|\cdot\|_{a(\cdot,\cdot)\otimes L^2}$ -Norm wird von Level zu Level halbiert, während sich die Reduktionsraten bzgl. der  $L^2$ - und der  $L^\infty$ -Normen mit wachsenden Leveln dem Wert 0.25 nähern.

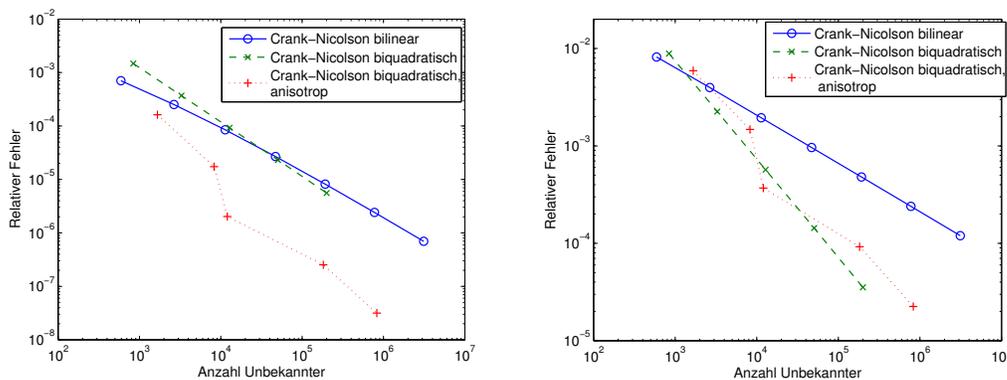
*Beispiel 4.6.* Wie wir in Abschnitt 4.3 ausführlich diskutiert haben, können wir für jede Folge ineinander geschachtelter Räume  $V_j^\Omega$  bzw.  $V_j^T$ , sofern wir Interpolationsoperatoren zwischen den Räumen gegeben haben, das Matrix-Vektor Produkt der Dünngitterdiskretisierungen effizient berechnen. Dabei ist der Algorithmus unabhängig von der konkreten Wahl der Räume. Wir sind somit in der Lage, ohne weitere Modifikationen Ansatzfunktionen höherer Ordnung im Ort oder in der Zeit zur Diskretisierung zu verwenden. Dies ist ein grundlegender Vorteil der in Abschnitt 4.3.1 vorgestellten Algorithmen Top-Down und Bottom-Up gegenüber früheren Versionen, bei denen die Algorithmik den zugrunde liegenden diskreten Funktionenräumen jeweils angepasst werden mußte, vgl. [32].

Als Beispiel betrachten wir nun die Crank-Nicolson Diskretisierung des Problems (4.2) in zwei Dimensionen mit der Lösung

$$u(x, t) = \sin\left(\pi \cdot \frac{x_1}{t+0.5}\right) \sin\left(\pi \cdot \frac{x_2}{t+0.5}\right), \quad (4.33)$$

wobei wir im Ort stückweise biquadratische Lagrange-Elemente verwenden.

Da wir in der Zeit weiterhin nur lineare Ansatzfunktionen benutzen, erwarten wir, dass die Crank-Nicolson Diskretisierung auf  $\tilde{V}_l^0$  mit biquadratischen Lagrange-



**Abbildung 4.9.** Konvergenzhistorie für das Problem (4.27) bzgl. der  $L^2$ - (links) und der  $H_{mix}^{1,0}$ -Norm (rechts) und das Crank-Nicolson Verfahren mit bilinearer Basis im Ort (Crank-Nicolson), mit biquadratischer Basis im Ort (Crank-Nicolson biquadratisch) und dem Crank-Nicolson Verfahren auf  $V_l^0$  (Crank-Nicolson biquadratisch, anisotrop).

Elementen bzgl. der  $L^2$ - und der  $L^\infty$ -Normen ähnliche Resultate wie die Crank-Nicolson Diskretisierung mit bilinearen Finiten Elementen in Beispiel 4.2 liefert, da der Fehler bzgl. der Zeitdiskretisierung den Gesamt-Diskretisierungsfehler dominiert. Im Gegensatz dazu sollte dieser Effekt, vergleichbar mit der Euler Diskretisierung, bei Verwendung von  $V_l^0$  durch eine feinere Zeit- als Ortsauflösung ausgeglichen werden und die Ergebnisse sollten sich dementsprechend verbessern.

Die Ergebnisse für die Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung bzgl. der Räume  $\tilde{V}_l^0$  und  $V_l^0$  bei Verwendung der biquadratischen hierarchischen Basis im Ort sind in Tabelle 4.6 aufgeführt. Wie erwartet beobachten wir für  $\tilde{V}_l^0$  eine Fehlerreduktion des Fehlers gemessen in der  $L^2$ - bzw. in der  $L^\infty$ -Norm um einen Faktor  $\rho = 0.25$ . Die Verwendung einer Diskretisierung die von höherer Ordnung im Ort ist, scheint den logarithmischen Faktor  $(l+1)/l$ , den wir noch in Beispiel 4.2 beobachten konnten, zu eliminieren. Für die Normen, die die Ableitungen bzgl. des Raumes enthalten, beobachten wir unabhängig davon, ob wir  $V_l^0$  oder  $\tilde{V}_l^0$  verwenden, eine Reduktionsrate von 0.25. Hier wird also die höhere Ableitung bzgl. der Ortsvariablen in der Norm bereits durch die Verwendung der biquadratischen Elemente kompensiert.

Verwenden wir hingegen das in der Zeit mit doppelt so großem Level wie im Ort konstruierte Raum-Zeit Dünngitter  $V_l^0$ , so erhalten wir damit, wie in Tabelle 4.6 zu sehen, schließlich die Reduktionsraten  $\rho_{L^2} \approx 0.12$  und  $\rho_{L^\infty} \approx 0.12$ .

In Abbildung 4.9 sind die Zahl der benötigten Freiheitsgrade zu dem Fehler gemessen in der  $L^2$ - und  $H_{mix}^{1,0}$ -Norm für die Crank-Nicolson Diskretisierung bzgl.  $\tilde{V}_l^0$  mit bilinearen Elementen sowie die Crank-Nicolson Diskretisierung mit biquadrati-

schen Elementen bzgl.  $V_l^0$  und  $\tilde{V}_l^0$  dargestellt. Obwohl wir bei der Verwendung von biquadratischen Elementen im Ort bei gleichem Level mehr Freiheitsgrade benötigen und somit die resultierenden Raum-Zeit Dünngitter mehr Freiheitsgrade als bei Verwendung der bilinearen hierarchischen Basis besitzen, erreichen wir bei Verwendung einer biquadratischen Basis bedingt durch die geringere Fehlerreduktion ein besseres Kosten-Nutzen Verhältnis.

In diesem Kapitel haben wir die Diskretisierung parabolischer Probleme mit Hilfe der Raum-Zeit Dünngitter diskutiert. Zur Erweiterung des klassischen Crank-Nicolson Verfahrens auf eine Crank-Nicolson Raum-Zeit Dünngitter Diskretisierung haben wir das Crank-Nicolson Verfahren als Continuous Galerkin oder auch Petrov-Galerkin Verfahren interpretiert. Des Weiteren haben wir das Discontinuous-Galerkin Verfahren untersucht, bei dem wir als Ansatz- und Testraum den aus der hierarchischen Basis im Ort und der stückweise linearen bzw. stückweise konstanten hierarchischen Basis in der Zeit konstruierten Raum-Zeit Dünngitterraum verwendet haben. Hierbei konnten wir eine a priori Fehlerabschätzung angeben, die den Diskretisierungsfehler durch entsprechende Interpolationsfehler abschätzt.

Zur effizienten Berechnung des Matrix-Vektor Produktes haben wir das von den klassischen Dünngittern bekannte Unidirektionale Prinzip auf den Raum-Zeit Dünngitterfall übertragen und dabei insbesondere die erforderlichen Teilalgorithmen Bottom Up und Top Down entsprechend modifiziert. Hiermit ist es ohne weitere Schwierigkeiten direkt möglich, Differentialoperatoren mit beliebigen variablen (in der Zeit konstanten) Koeffizienten zu behandeln, was bei klassischen Dünngittern zu besonderen Problemen führt.

Schließlich haben wir anhand einiger numerischer Beispiele die Konvergenzeigenschaften der Raum-Zeit Dünngitter Crank-Nicolson und der Discontinuous-Galerkin Verfahren mit stückweise konstanten bzw. linearen Funktionen in der Zeit und der  $d$ -linearen hierarchischen Basis im Ort untersucht. Die Ergebnisse deuten daraufhin, dass sich der Diskretisierungsfehler wie der Interpolationsfehler des Raum-Zeit Dünngitters verhält.

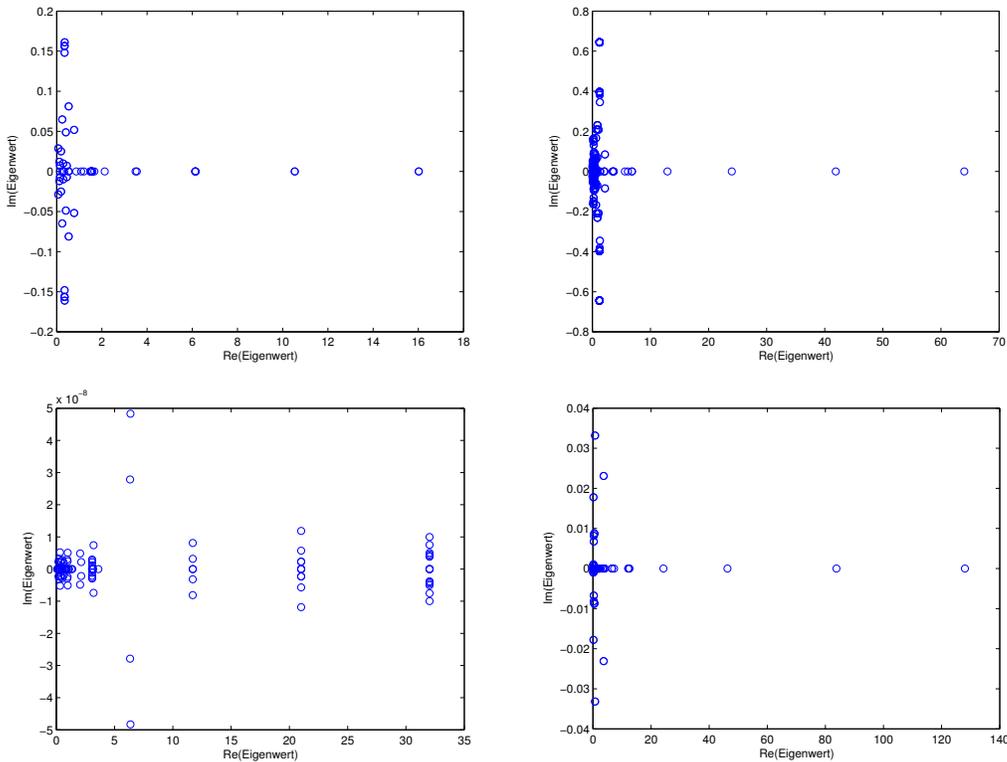
## Kapitel 5

# Lösung des linearen Gleichungssystems

Die im vorigen Kapitel beschriebenen Raum-Zeit Dünngitter Diskretisierungen des parabolischen Problems führen zu großen linearen Gleichungssystemen. Dabei sind die zugehörigen Matrizen nicht dünnbesetzt, sondern besitzen eine für die Diskretisierung mit Multilevel-Basen typische Fingerstruktur, vgl. Abbildung 4.1.

Wie wir in Abschnitt 4.3 gezeigt haben, kann das Matrix-Vektorprodukt mit einem Aufwand proportional zu der Anzahl der Unbekannten mit Hilfe des Unidirektionalen Prinzips berechnet werden. Aus diesem Grund scheinen sich zur Lösung der linearen Gleichungssysteme iterative Löser anzubieten, die nur die Berechnung des Matrix-Vektor Produktes benötigen. Hierbei haben wir insbesondere zu beachten, dass die zugehörigen Matrizen weder positiv definit noch symmetrisch sind. Die Eigenwerte sind komplex und liegen in der rechten Halbebene wie in Abbildung 5.1 am Beispiel der Crank-Nicolson und der Euler Diskretisierung in einer Ortsdimension für  $\tilde{V}_5^0$  und  $\tilde{V}_7^0$  zu sehen ist. Wir benötigen daher einen iterativen Löser für unsymmetrische Probleme.

Bei einer ersten Betrachtung bieten sich so genannte Krylov-Unterraum-Verfahren für unsymmetrische Probleme, wie etwa das CGNE-, das QMR- das GMRES oder das BiCGStab-Verfahren [9] an. Allerdings verschlechtern sich in der Praxis die Konvergenzraten dieser Verfahren mit wachsender Kondition. Wie bei der Diskretisierung von elliptischen Problemen ist die Kondition der Matrizen der Raum-Zeit Dünngitter Diskretisierungsverfahren jedoch von der Levelzahl des verwendeten Dünngitters abhängig und wächst mit steigender Levelzahl. In Abbildung 5.1 ist dieses Verhalten gut nachzuvollziehen. So erkennt man, dass der betragsmäßig größte Eigenwert von Level zu Level wächst. Bei Verwendung dieser Krylov-Unterraum-Verfahren würde deshalb die Anzahl der benötigten Iterationen zur Lösung des linearen Gleichungssystems mit steigender Diskretisierungsgenauigkeit zunehmen, da die Verfahren immer langsamer konvergieren. Wir möchten daher in diesem Abschnitt die Entwicklung von Lösern diskutieren, deren Konvergenzrate unabhängig von der Feinheit der Diskretisierung ist.



**Abbildung 5.1.** Die zur Crank-Nicolson Diskretisierung (oben) und der Euler Diskretisierung (unten) auf  $\Omega_T = (0, 1) \times (0, 1)$  mit  $\tilde{V}_5^0$  (links) und  $\tilde{V}_7^0$  (rechts) gehörigen Eigenwerte.

Hierbei spielen Mehrgitter- [72, 108] und Multilevelmethoden [61] eine bedeutende Rolle, die in den letzten Jahrzehnten zur effizienten Lösung großer linearer Gleichungssysteme entwickelt worden sind. Diese Verfahren besitzen pro Iteration einen Rechenaufwand, der proportional zu der Zahl der Unbekannten ist, während sich deren Konvergenzraten durch eine von der Feinheit der Diskretisierung unabhängigen Konstanten kleiner 1 nach oben abschätzen lassen. Wir beschäftigen uns daher in diesem Kapitel mit der Entwicklung und Untersuchung entsprechender Multilevelmethoden für unsere Raum-Zeit Dünngitterdiskretisierung.

Für die Lösung parabolischer Probleme auf vollen Gittern sind einige Multilevelverfahren entwickelt und untersucht worden. Zur Lösung der auf vollen Gittern diskretisierten parabolischen Probleme ist im Regelfall eine Sequenz von stationären Ortsproblemen zu lösen. In diesem Fall können die Gleichungssysteme der stationären Probleme mit Hilfe von klassischen Mehrgitterverfahren für elliptische Probleme [28, 71, 76] gelöst werden. Eine weitere Möglichkeit besteht darin, das auf vollen Gittern diskretisierte Raum-Zeit Problem gleichzeitig in Raum und Zeit mit-

tels Mehrgittermethoden zu behandeln, vgl. [71, 110]. Da hier formal nicht zwischen Raum und Zeit unterschieden wird, (obwohl eigentlich ein Zeitschrittverfahren vorliegt) lassen sich diese Verfahren nicht nur bzgl. des Ortes sondern auch bzgl. der Zeit effizienter parallelisieren. Zusätzlich ermöglichen Mehrgitterverfahren auf vollen Raum-Zeit Gittern die effiziente Lösung von Systemen parabolischer Gleichungen, bei denen einige Unbekannte vorwärts- und andere rückwärts in der Zeit laufen, wie dies beispielsweise bei instationären Kontrollproblemen häufig anzutreffen ist [24]. Für die Lösung mit Hilfe der Waveform-Relaxation wurden Raum-Zeit Mehrgitterverfahren beispielsweise in [109, 111, 112] untersucht. Allerdings werden diese Raum-Zeit Mehrgittermethoden aufgrund der hohen Komplexität und des damit verbundenen hohen Speicherverbrauchs von vollen Gittern bislang nur auf Probleme mit maximal zwei Ortsdimensionen angewendet.

Im Fall von klassischen Dünngittern gibt es für parabolische Probleme hingegen keine genaueren Untersuchungen zur Lösung der anfallenden linearen Gleichungssysteme. In [5] wird zwar ein Iterationsverfahren über die Zeit vorgeschlagen, bei dem die anfallenden Ortsprobleme mittels eines vorkonditionierten CG Verfahrens gelöst werden, jedoch bleibt eine theoretische oder numerische Untersuchung des Konvergenzverhaltens der Methode aus.

Für elliptische Probleme existieren mehrere Arbeiten zur Lösung der bei der Diskretisierung mit klassischen Dünngittern anfallenden Gleichungssysteme. In [60, 92] werden Mehrgitterverfahren für die Dünngitterdiskretisierung mittels der eindimensionalen hierarchischen Basis im elliptischen Fall vorgestellt, die als Folge von  $V$ -Zyklen mit Semi-Vergrößerung interpretiert werden können. Die numerischen Experimente lassen dabei auf eine levelunabhängige Konvergenzrate schließen. Für Dünngitterdiskretisierungen mit Multilevelbasen höherer Ordnung wird in [31] ein Mehrgitterverfahren entwickelt. In [61] werden schließlich klassische Multilevelöser als Iterationsverfahren über einem mit Hilfe eines Erzeugendensystems diskretisierten System interpretiert. Dies ermöglicht die einfache Beschreibung verschiedenster Multilevel-Verfahren, die sich leicht auf den Dünngitterfall übertragen lassen. Ordnet man beispielsweise das Erzeugendensystem levelweise an, so entspricht im Vollgitterfall ein symmetrisches Gauß-Seidel Verfahren über dem Erzeugendensystem dem klassischen  $V$ -Mehrgitterzyklus und kann auf Dünngitter direkt übertragen werden. Die numerischen Experimente in [61] zeigen dabei, dass somit auch im Dünngitterfall levelunabhängige Konvergenzraten erzielt werden können.

Die hohe Flexibilität dieser Sichtweise und die einfache Beschreibung der Verfahren sind der Grund dafür, dass wir in diesem Kapitel das Erzeugendensystem zur Entwicklung eines Multilevel-Verfahrens zur Lösung der linearen Gleichungssysteme verwenden werden.

In Abschnitt 5.1 führen wir zunächst das Erzeugendensystem ein. Ferner untersuchen wir die Konvergenzeigenschaften des (Block-) SOR Verfahrens im Fall von unsymmetrischen, semidefiniten Matrizen. Es wird sich zeigen, dass der Relaxati-

onsparameter stets so gewählt werden kann, dass das (Block-) SOR Verfahren für die aus der Diskretisierung mit dem Discontinuous-Galerkin Verfahrens bzgl. des Erzeugendensystems resultierenden linearen Gleichungssysteme konvergiert. Dadurch motiviert konstruieren wir in den Abschnitten 5.2 und 5.3 zwei unterschiedliche Multilevelverfahren als (Block-) Gauß-Seidel Verfahren über dem Erzeugendensystem. Abschließend untersuchen wir in Abschnitt 5.4 die Konvergenzeigenschaften der beiden Verfahren exemplarisch anhand einiger numerischer Beispiele.

## 5.1 Iterationsverfahren über dem Erzeugendensystem

Wie bereits in Kapitel 2 sei  $V_j^\Omega$  ein Finite Element Raum bzgl. Level  $j$  mit nodaler Basis  $\varphi_{j,i}^\Omega$ ,  $1 \leq i \leq \dim(V_j^\Omega)$  im Ort und  $\psi_{j,i}^\Omega$  bezeichne die dazugehörige hierarchische Basis. Sei  $V_j^T$  einer der Räume aus Abschnitt 2.3 mit nodaler Basis  $\varphi_{j,i}^T$ ,  $1 \leq i \leq \dim(V_j^T)$  und der hierarchischen Basis  $\psi_{j,i}^T$ . Dann erhalten wir wie in (4.15) mit Hilfe der hierarchischen Basen eine Basis  $\mathcal{B}_{V_l^0}$  für den Raum-Zeit Dünngitterraum.

Wir bezeichnen nun den Dünngitterraum, den wir als Ansatzraum der Raum-Zeit Diskretisierung verwenden, mit  $V_l^{0,an}$  und den Dünngitterraum, den wir als Testraum verwenden wollen als  $V_l^{0,te}$ . Dabei ist  $\psi_{j,i}^{an} \in \mathcal{B}_{V_l^{0,an}}$  die entsprechende hierarchische Basis des Ansatzraumes und  $\psi_{j,i}^{te} \in \mathcal{B}_{V_l^{0,te}}$  die hierarchische Basis des Testraumes. Sei  $n$  die Dimension des Ansatzraumes und nach (eventueller Elimination von Anfangs- und Randbedingungen) entspricht diese Dimension der Dimension des Testraumes, d.h.

$$n := \dim(V_l^{0,an}) = \dim(V_l^{0,te}).$$

Sei  $B(\cdot, \cdot)$  nun die zu einer wie in Kapitel 4 beschriebenen Diskretisierung gehörige Bilinearform, so erhalten wir nach Anordnung der Basisfunktionen des Ansatz- und des Testraumes den Koeffizientenvektor  $\mathbf{u}_l = (u_i) \in \mathbb{R}^{\dim(V_l^{0,an})}$  der eindeutigen Darstellung  $u_l = \sum_{i=1}^n u_i \psi_i$  der diskreten Lösung  $u_l$  über die Lösung des linearen Gleichungssystems

$$L\mathbf{u}_l = \mathbf{f}$$

mit  $L \in \mathbb{R}^{n \times n}$ ,  $\mathbf{f} \in \mathbb{R}^n$ ,

$$(L)_{ij} := B(\psi_j^{an}, \psi_i^{te}), \psi_j^{an} \in \mathcal{B}_{V_l^{0,an}}, \psi_i^{te} \in \mathcal{B}_{V_l^{0,te}},$$

$$(\mathbf{f})_i := \int_0^T (f, \psi_i^{te})_0 dt, \psi_i^{te} \in \mathcal{B}_{V_l^{0,te}}.$$

Hierbei wissen wir aus Lemma 4.1 bzw. 4.3 dass die bei der Crank-Nicolson bzw. Discontinuous-Galerkin Diskretisierung anfallenden linearen Gleichungssysteme eine eindeutige Lösung besitzen und jede Funktion des Dünngitterraumes durch genau einen Koeffizientenvektor repräsentiert wird. Wir lassen nun eine nicht-eindeutige Darstellung von Funktionen in unserem Dünngitterraum zu und erweitern daher wie

in [61] beschrieben unsere Basis  $\mathcal{B}_{V_l^0}$  für den Dünngitterraum zu einem Erzeugendensystem  $\mathcal{E}_{V_l^0}$ ,

$$\mathcal{E}_{V_l^0} = \{\varphi_{\mathbf{j},\mathbf{i}} := \varphi_{j_1,i_1}^\Omega \cdot \varphi_{j_2,i_2}^T \mid j_1+j_2 \leq l, 1 \leq i_1 \leq \dim(V_{j_1}^\Omega), 1 \leq i_2 \leq \dim(V_{j_2}^T)\}. \quad (5.1)$$

Wir bezeichnen mit  $\mathcal{E}_{V_l^{0,an}}$  das auf diese Weise konstruierte Erzeugendensystem für den Ansatzraum  $V_l^{0,an}$  und mit  $\mathcal{E}_{V_l^{0,te}}$  das Erzeugendensystem des Testraumes  $V_l^{0,te}$ . Wir können ohne Einschränkung annehmen (nach eventueller Elimination der Rand- und Anfangsbedingungen), dass  $\mathcal{E}_{V_l^{0,an}}$  und  $\mathcal{E}_{V_l^{0,te}}$  die gleiche Anzahl von Elementen enthalten und definieren

$$n_E := |\mathcal{E}_{V_l^{0,an}}| = |\mathcal{E}_{V_l^{0,te}}|.$$

Ordnen wir nun die Elemente des Erzeugendensystems an, so erhalten wir für die Koeffizienten  $\mathbf{u}_l^E = (u_i^E) \in \mathbb{R}^{n_E}$  einer Darstellung der Lösung  $u_l$  bzgl. des Erzeugendensystems das Gleichungssystem

$$L^E \mathbf{u}_l^E = \mathbf{f}^E \quad (5.2)$$

mit  $L^E \in \mathbb{R}^{n_E \times n_E}$ ,  $\mathbf{f}^E \in \mathbb{R}^{n_E}$ ,

$$\begin{aligned} (L^E)_{ij} &:= B(\varphi_j^{an}, \varphi_i^{te}), \quad \varphi_j^{an} \in \mathcal{E}_{V_l^{0,an}}, \varphi_i^{te} \in \mathcal{E}_{V_l^{0,te}}, \\ (\mathbf{f}^E)_i &:= \int_0^T (f, \varphi_i^{te})_0 dt, \quad \varphi_i^{te} \in \mathcal{E}_{V_l^{0,te}}. \end{aligned}$$

Die derart definierte Matrix  $L^E$  ist semidefinit und besitzt den gleichen Rang wie die Matrix  $L$ . Zusätzlich ist das lineare Gleichungssystem lösbar, da die rechte Seite konsistent konstruiert ist, vgl. [61].

Obwohl die Darstellung einer Funktion  $u_l \in V_l^{0,an}$  bzgl. des Erzeugendensystems nicht eindeutig ist, können wir durch Interpolation und Summation ohne Schwierigkeiten die eindeutige Darstellung bzgl. der Basis berechnen. Damit erhalten wir nach Anordnung der Elemente der Basis  $\mathcal{B}_{V_l^{0,an}}$  und der Funktionen des Erzeugendensystems  $\mathcal{E}_{V_l^{0,an}}$  eine surjektive lineare Abbildung

$$S^{E,B} : \mathbb{R}^{n_E} \rightarrow \mathbb{R}^n, \quad (5.3)$$

die einen Koeffizientenvektor bzgl. des Erzeugendensystems  $\mathcal{E}_{V_l^{0,an}}$  einer Funktion  $u_l \in V_l^{0,an}$  in den entsprechenden Koeffizientenvektor bzgl. der Basis  $\mathcal{B}_{V_l^{0,an}}$  transformiert. Mit Hilfe dieser Abbildung können wir die Matrix  $L^E$  durch

$$L^E = (S^{E,B})^T L S^{E,B} \quad (5.4)$$

darstellen. Nun können, wie in [61] beschrieben, herkömmliche lineare stationäre Iterationsverfahren der Art

$$\mathbf{u}_l^{E,it+1} = \mathbf{u}_l^{E,it} + C^E (\mathbf{f}^E - L^E \mathbf{u}_l^{E,it}), \quad (5.5)$$

wobei  $C^E$  regulär ist, zur Lösung der Gleichung (5.2) betrachtet werden. Das Iterationsverfahren (5.5) ist dabei genau dann konvergent, wenn

$$\tilde{\rho}(I - C^E L^E) < 1. \quad (5.6)$$

Hierbei ist  $\tilde{\rho}(A)$  der verallgemeinerte Spektralradius der Matrix  $A$

$$\tilde{\rho}(A) := \max\{|\lambda| \mid \lambda \in \sigma(A), \lambda \neq 1\}, \quad (5.7)$$

wobei  $\sigma(A)$  den Spektralradius der Matrix  $A$  bezeichnet. Wir erhalten nun unterschiedlichste Multilevelverfahren indem wir klassische Iterationsverfahren zur Lösung des semidefiniten Systems (5.2) anwenden. Im Folgenden betrachten wir zwei Methoden, die auf dem (Block-) Gauß-Seidel Verfahren basieren. Sei dazu die Matrix  $L^E = D^E + U^E + R^E$  in eine (Block-) Diagonalmatrix  $D^E$ , eine untere (Block-) Dreiecksmatrix  $U^E$  und eine obere (Block-) Dreiecksmatrix  $R^E$  aufgespalten. Wir erhalten das (Block-) SOR Verfahren für das Problem (5.2) durch die Wahl von

$$C^E := \omega(D^E + \omega U^E)^{-1}$$

in der Iterationsvorschrift (5.5) und damit als Spezialfall für  $\omega = 1.0$  das (Block-) Gauß-Seidel Verfahren.

Im Fall unsymmetrischer Matrizen ist nur sehr wenig Konvergenztheorie für klassische Iterationsverfahren vorhanden, was theoretische Betrachtungen und eine systematische Untersuchung von Verfahren zur Lösung von (5.2) sehr schwierig macht. Für den Fall, dass die zugrunde liegende Matrix regulär ist, werden in [117] Konvergenzuntersuchungen für das SOR Verfahren unter der Voraussetzung angestellt, dass der nicht-symmetrische Anteil der Matrix genügend klein ist. Hierbei erhält man ähnliche Bedingungen an den Relaxationsparameter wie im symmetrischen, positiv definiten Fall. Ist der symmetrische Anteil der Matrix positiv definit ist, wird in [88] ebenfalls ein Wertebereich für den Relaxationsparameter  $\omega$  angegeben, für den das SOR Verfahren konvergent ist. Dieses Resultat verallgemeinern wir nun auf den Fall singulärer Matrizen. Wir nehmen dazu an, dass  $D^E = I$  gilt und spalten  $L^E$  additiv in

$$L^E = I + B$$

auf, wobei  $I$  die Identität bezeichnet und die Matrix  $B = U^E + R^E$  in eine untere Dreiecksmatrix  $U^E$  und eine obere Dreiecksmatrix  $R^E$  aufgeteilt ist. Analog dazu definieren wir

$$\tilde{B} := U^E - R^E.$$

Für eine Matrix  $L^E$  bezeichne  $\mathcal{G}(L^E)$  den symmetrischen und  $\mathcal{S}(L^E)$  den schief-symmetrischen Anteil

$$\begin{aligned} \mathcal{G}(L^E) &:= \frac{1}{2}(L^E + L^{E^T}), \\ \mathcal{S}(L^E) &:= \frac{1}{2}(L^E - L^{E^T}). \end{aligned}$$

Sei  $P$  die bzgl. des euklidischen Skalarproduktes orthogonale Projektion auf den Kern von  $L^E$ ,

$$P : \mathbb{R}^{n_E} \rightarrow \ker(L^E)$$

und  $P^\perp$  dementsprechend die orthogonale Projektion auf das orthogonale Komplement des Kernes,

$$P^\perp : \mathbb{R}^{n_E} \rightarrow (\ker(L^E))^\perp.$$

Wir verwenden abkürzend  $\tilde{n} = \dim(\ker(L^E)^\perp)$ . Es sei  $\hat{B} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$  definiert durch

$$\hat{B} := P^\perp B P^\perp \big|_{(\ker(L^E))^\perp}.$$

Es seien  $\hat{g}_1 \leq \hat{g}_2 \leq \dots \leq \hat{g}_{\tilde{n}}$ , die Eigenwerte der symmetrischen Matrix  $\mathcal{G}(\hat{B})$ . Analog seien  $\tilde{g}_1 \leq \dots \leq \tilde{g}_n$  die Eigenwerte des symmetrischen Teils  $\mathcal{G}(\tilde{B})$  von  $\tilde{B}$ . Als Verallgemeinerung des Resultates in [88] erhalten wir nun für das SOR Verfahren die folgende Abschätzung für den Relaxationsparameter  $\omega$ .

**Satz 5.1.** *Der symmetrische Anteil  $\mathcal{G}(L^E)$  der Matrix  $L^E$  sei eingeschränkt auf  $(\ker(L^E))^\perp$  positiv definit. Seien  $\hat{\sigma}$  und  $\tilde{\sigma}$  die Spektralradien von  $\mathcal{S}(\hat{B})$  bzw.  $\mathcal{S}(\tilde{B})$ . Weiter gelte*

$$\ker(L^E) = \ker((L^E)^T).$$

Dann konvergiert das SOR-Verfahren für alle  $\omega$  mit

$$0 < \omega < 2 \left( 1 - \tilde{g}_1 + \frac{\hat{\sigma} \tilde{\sigma}}{1 + \hat{g}_1} \right)^{-1}.$$

Der Beweis ist sehr technisch und verläuft ähnlich wie der Beweis in [88] für den Fall von regulären Matrizen. Aus diesem Grund verzichten wir an dieser Stelle auf eine Beweisführung und diskutieren den Beweis im Anhang.

Die Ergebnisse des Satzes übertragen sich analog auf das Block-SOR Verfahren. Wie Satz 5.1 zeigt, ist das SOR-Verfahren bei geeigneter Wahl des Relaxationsparameters konvergent, falls die Matrix  $L^E$  die obigen Bedingungen erfüllt. Das nächste Resultat zeigt, dass das diagonalskalierte System

$$\tilde{L}^E := (D^E)^{-1/2} L^E (D^E)^{-1/2},$$

welches aus der Dünngitter Discontinuous-Galerkin Diskretisierung stammt, die für die Anwendung des Satzes notwendigen Bedingungen erfüllt.

**Satz 5.2.** *Sei  $L^E$  die aus der Raum-Zeit Dünngitter Discontinuous-Galerkin Diskretisierung, vgl. (4.7), resultierende Matrix eines parabolischen Problems mit in der*

Zeit konstantem elliptischen Ortsoperator  $a(\cdot, \cdot)$  und  $u_0 = 0$ . Dann erfüllt die diagonalskalierte Matrix  $\tilde{L}^E$  des diskretisierten Problems die notwendigen Bedingungen des Satzes 5.1, d.h. es gilt

$$\ker(\tilde{L}^E) = \ker((\tilde{L}^E)^T)$$

und  $\mathcal{G}(\tilde{L}^E)$  eingeschränkt auf  $(\ker(\tilde{L}^E))^\perp$  ist positiv definit.

**Beweis.** Aus (4.10) folgt unmittelbar, dass  $\mathcal{G}(L^B)$  positiv definit ist. Da nun per Konstruktion

$$\ker(L^E) = \ker(S^{E,B})$$

gilt, folgt aus der Darstellung (5.4), dass für alle  $x \notin \ker(L^E)$

$$x^T \mathcal{G}(L^E)x = (S^{E,B}x)^T \mathcal{G}(L^B)(S^{E,B}x) > 0 \quad (5.8)$$

gilt. Des Weiteren haben wir

$$\ker(\tilde{L}^E) = \ker(S^{E,B}(D^E)^{-1/2}),$$

und für  $x \in \ker(S^{E,B}(D^E)^{-1/2})^\perp$  gilt  $(D^E)^{-1/2}x \notin \ker(S^{E,B})$ . Zusammen mit (5.8) folgt daraus, dass  $\tilde{L}^E$  auf  $(\ker(\tilde{L}^E))^\perp$  positiv definit ist. Weiterhin gilt

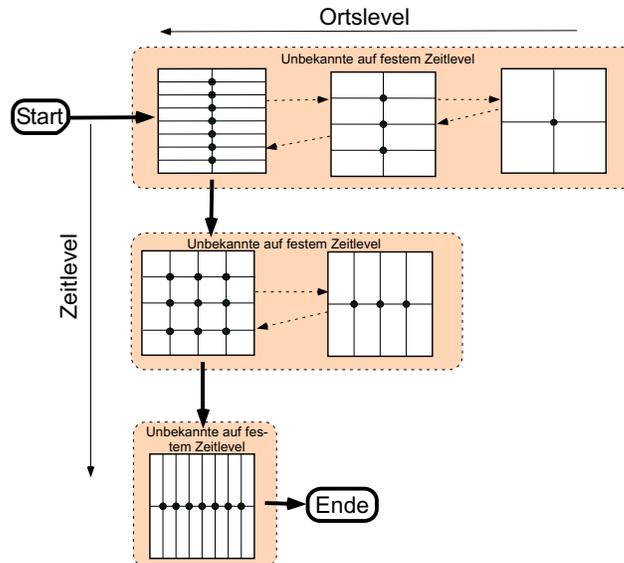
$$\ker(\tilde{L}^E) = \ker(S^{E,B}) = \ker((\tilde{L}^E)^T)$$

und die Behauptung ist gezeigt.  $\square$

Verschwundet die Anfangsbedingung des parabolischen Problems nicht, so kann das aus der Elimination der Anfangsbedingung aus dem ursprünglichen Gleichungssystem resultierende System als die Diskretisierung eines Problems mit Nullrandbedingungen und modifizierter rechter Seite aufgefasst werden. Aus diesem Grund überträgt sich das Ergebnis auch auf den Fall von inhomogenen Anfangsbedingungen.

Für die Crank-Nicolson Diskretisierung ist der symmetrische Anteil eingeschränkt auf das orthogonale Komplement des Kerns nicht positiv definit, weshalb wir Satz 5.1 nicht anwenden können, um die Konvergenz des SOR Verfahrens zu beweisen.

Man beachte, dass die Konvergenzgeschwindigkeit in aller Regel von der gewählten Anordnung der Unbekannten abhängt. So ist anzunehmen, dass bei parabolischen Gleichungen die zugehörigen Unbekannten nicht rückwärts in der Zeit durchlaufen werden sollten. Wie schon in [61] beschrieben, erhält man durch die Wahl von unterschiedlichen Durchlaufreihenfolgen, verschiedenste Partitionierung in Blöcke und entsprechenden Block-Varianten des SOR Verfahrens unterschiedlichste Multilevelverfahren. Motiviert durch das obige Resultat stellen wir nun auf Basis des (Block-)SOR Verfahrens zwei Multilevellöser vor, die sich durch die Block-Partitionierung und damit verbunden in der Durchlaufreihenfolge unterscheiden. Wir beschränken



**Abbildung 5.2.** Durchlaufreihenfolge des nach Zeitleveln angeordneten Block-Gauß-Seidel Verfahrens.

uns bei der Untersuchung auf den Fall  $\omega = 1$  und verwenden für die eine Methode das Gauß-Seidel Verfahren, während die andere Methode auf dem Block-Gauß-Seidel Verfahren beruht.

Zur Vereinfachung der Darstellung betrachten wir in den nächsten beiden Abschnitten den Fall, dass der Ansatz- und der Testraum gleich sind, d.h.  $V_l^{0,an} = V_l^{0,te} = \tilde{V}_l^0$ . Die Überlegungen lassen sich jedoch direkt auf den Petrov-Galerkin Fall übertragen.

## 5.2 Die Zeitlevelweise angeordnete Block-Gauß-Seidel Iteration

Für das Block-Gauß-Seidel Verfahren, das wir in diesem Abschnitt betrachten, ordnen wir die Funktionen des Erzeugendensystems nach dem Level der Zeitfunktionen  $\varphi_{j,i}^T$  an, d.h. wir zerlegen die Menge  $\mathcal{E}_{\tilde{V}_l^0}$  in Teilmengen  $\mathcal{E}_{\tilde{V}_l^0}^j$ ,  $1 \leq j \leq l$ , so dass

$$\mathcal{E}_{\tilde{V}_l^0} = \bigcup_{j=1}^l \mathcal{E}_{\tilde{V}_l^0}^j \text{ mit } \mathcal{E}_{\tilde{V}_l^0}^j := \{\varphi_{\tilde{j},i}^T \mid \tilde{j}_1 = j\}$$

gilt. Wir erhalten damit das Blocksystem

$$\begin{pmatrix} L_{00}^E & \dots & L_{0l}^E \\ \vdots & \dots & \vdots \\ L_{l0}^E & \dots & L_{ll}^E \end{pmatrix} \begin{pmatrix} \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_l \end{pmatrix} = \begin{pmatrix} \mathbf{f}_0 \\ \vdots \\ \mathbf{f}_l \end{pmatrix}, \quad (5.9)$$

wobei

$$L_{ij}^E := (B(\varphi_{\mathbf{p}_1, \mathbf{q}_1}, \varphi_{\mathbf{p}_2, \mathbf{q}_2}))_{\varphi_{\mathbf{p}_1, \mathbf{q}_1} \in \mathcal{E}_{\tilde{V}_l^i}, \varphi_{\mathbf{p}_2, \mathbf{q}_2} \in \mathcal{E}_{\tilde{V}_l^j}}.$$

Bei dem zugehörigen Block-Gauß-Seidel Verfahren treten in dem  $it$ -ten Iterationsschritt zur Bestimmung der  $it+1$ -ten Unbekannten nun die zu lösenden Teilprobleme

$$L_{kk} \mathbf{u}_k^{it+1} = \mathbf{f}_k - \sum_{j=0}^{k-1} L_{kj} \mathbf{u}_j^{it+1} - \sum_{j=k+1}^l L_{kj} \mathbf{u}_j^{it} \quad (5.10)$$

auf. Dabei ist  $L_{kk}$  in der Regel immer noch zu groß, um mittels eines direkten Verfahrens, wie beispielsweise der Gauß-Elimination, gelöst zu werden. Da allerdings alle in jedem Teilproblem vorkommenden Funktionen bzgl. der Zeit gleiches Level haben, reduziert sich die Lösung jedes Teilproblems auf das der Raum-Zeit Dünngitterdiskretisierung zugrunde liegende klassische Zeitschrittverfahren, bei dem eine Folge von elliptischen Problemen zu lösen ist. Für die Behandlung der elliptischen Teilprobleme bieten sich zwei Möglichkeiten an.

Zunächst können wir jedes Teilproblem inexakt mit Hilfe von wenigen Schritten eines stationären linearen Iterationsverfahrens lösen. In unserem Kontext kann dieses Vorgehen als Glättung des Fehlers interpretiert werden. Hierbei übernimmt das verwendete Iterationsverfahren die Aufgabe eines Glätters.

Da jedes anfallende elliptische Problem wiederum Multilevelstruktur besitzt, können die elliptischen Teilprobleme jedoch auch effizient mittels eines Multilevelverfahrens exakt (bis auf Maschinengenauigkeit) gelöst werden. Numerische Experimente haben allerdings gezeigt, dass wir auch bei inexakter Lösung der Block-Systeme die gleichen Konvergenzraten wie bei exakter Lösung erhalten. Hierbei haben wir die Lösung jedes elliptischen Problems durch jeweils einen Schritt des symmetrischen Gauß-Seidel Verfahrens über dem Erzeugendensystem im Ort approximiert, was einem symmetrischen V-Mehrgitterzyklus, vgl. [61], entspricht. Dieses Vorgehen führt zu nahezu den gleichen Konvergenzraten wie die bis auf Maschinengenauigkeit exakte Lösung jedes Teilproblems mit Hilfe eines symmetrischen V-Zyklus. Aus diesem Grund werden wir in den numerischen Ergebnissen nur die Resultate für das Block-Gauß-Seidel Verfahren mit inexakter Lösung diskutieren. Für das Block-Gauß-Seidel Verfahren wählen wir die Blockdurchlaufreihenfolge  $k = 0, \dots, l-1, l, l-1, \dots, 0$ , d.h. wir verwenden ein symmetrisches Block-Gauß-Seidel Verfahren das mit dem größten Zeitlevel beginnt. Abbildung 5.2 zeigt am Beispiel von  $\tilde{V}_3^0$  die Durchlaufreihenfolge dieses Verfahrens, das wir abkürzend auch einfach als Gauß-Seidel Verfahren bezeichnen werden.

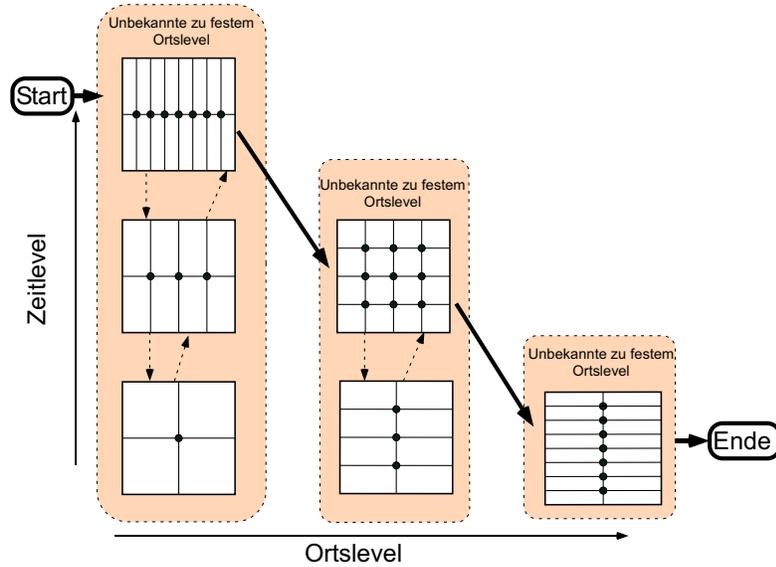


Abbildung 5.3. Durchlaufreihenfolge des nach Ortsleveln angeordneten Waveform-Gauß-Seidel Verfahrens.

### 5.3 Die Ortslevelweise angeordnete Gauß-Seidel Iteration: Das Waveform-Gauß-Seidel Verfahren

In Anlehnung an die Mehrgitter-Waveform-Relaxation [109] stellen wir nun einen auf dem Gauß-Seidel Verfahren über dem Erzeugendensystem basierenden Multi-levellöser vor. Hierbei ist wiederum die Wahl der Durchlaufenreihenfolge bzw. die Anordnung der Unbekannten für das Verfahren von Bedeutung.

Betrachten wir hierfür zunächst eine Semidiskretisierung des parabolischen Problems bezüglich des Ortes mit Finiten Differenzen auf einem Gitter der Maschenweite  $h$ . Wir erhalten dann ein großes System gewöhnlicher Differentialgleichungen

$$\partial_t u_h(t) = L_h u_h(t) + f_h(t) \quad (5.11)$$

für die semidiskrete Lösung  $u_h(t) = (u_h^1(t), \dots, u_h^2(t))^T$ , wobei  $L_h$  die zu der Ortsdiskretisierung gehörende Finite Differenzen Matrix bezeichnet. Ist nun zum Beispiel  $d = 1$  und  $L_h$  die Finite Differenzen Diskretisierung des Laplace Operators, so erhält man die lexikographisch angeordnete Gauß-Seidel Waveform Relaxation zur Bestimmung der  $\nu$ -ten Iterierten  $u_h^\nu = (u_{h,0}^\nu, \dots, u_{h,N}^\nu)$  als Lösung von

$$\partial_t u_{h,i}^\nu = \frac{u_{h,i-1}^\nu - 2u_{h,i}^\nu + u_{h,i+1}^{\nu-1}}{h^2} + f_{h,i}. \quad (5.12)$$

Allgemein erhalten wir für eine gegebene Matrix  $L_h = (a_{ij})$  die  $\nu$ -te Iterierte der Gauß-Seidel Waveform Relaxation (in lexikographischer Durchlaufreihenfolge) durch

$$\partial_t u_{h,i}^\nu = \sum_{j \leq i} a_{ij} u_{h,j}^\nu + \sum_{j > i} a_{ij} u_{h,j}^{\nu-1} + f_{h,i}. \quad (5.13)$$

Somit wird für jeden Gitterpunkt im Ort eine gewöhnliche Differentialgleichung gelöst, die wiederum mittels geeigneter Zeitschrittverfahren diskretisiert wird. Hierbei ist die Konvergenzgeschwindigkeit des Verfahrens von der Feinheit der Diskretisierung abhängig. Um ein von dem Diskretisierungslevel unabhängig konvergierendes Verfahren zu erhalten, wird nun in [109] in Analogie zu herkömmlichen Mehrgittermethoden die so genannte Multigrid-Waveform Methode vorgeschlagen. Wie bei elliptischen Problemen wird hierbei das Residuum  $r_h^\nu$

$$r_h^\nu = \partial_t u_h^\nu - L_h u_h - f_h$$

der  $\nu$ -ten Iterierten zunächst auf ein gröberes Gitter (mit der Maschenweite  $H$ ) mit Hilfe eines Restriktionsoperators  $R_h^H$  transformiert und dort die so genannte Defektgleichung

$$\begin{aligned} \partial_t u_H(t) &= L_H u_H(t) + R_h^H r_h, \\ u_H(0) &= 0 \end{aligned}$$

gelöst. Ist dieses grobe Problem noch zu groß, um exakt gelöst zu werden, so wird das Verfahren rekursiv angewendet. Anschließend wird die mit Hilfe eines Interpolationsoperators  $I_H^h$  in den Feingitterraum interpolierte Grobgitterlösung  $u_H$  als Korrektur auf die momentane Feingitteriterierte addiert

$$u_{h,i}^{\nu+1} = I_H^h u_H,$$

weshalb  $u_H$  auch Grobgitterkorrektur genannt wird. Um ein ähnliches Verfahren für die Raum-Zeit Dünngitterdiskretisierung zu erhalten, zerlegen wir zunächst das Erzeugendensystem  $\mathcal{E}_{\tilde{V}_l^0}$  in Teilmengen  $\mathcal{E}_{\tilde{V}_l^0}^j$  so dass

$$\mathcal{E}_{\tilde{V}_l^0} = \bigcup_{j=1}^l \mathcal{E}_{\tilde{V}_l^0}^j \quad \text{mit} \quad \mathcal{E}_{\tilde{V}_l^0}^j := \{\varphi_{\tilde{\mathbf{j}},i} \mid \tilde{j}_2 = j\}$$

gilt. Daraus resultiert ein Blocksystem

$$\begin{pmatrix} L_{00}^E & \cdots & L_{0l}^E \\ \vdots & \cdots & \vdots \\ L_{l0}^E & \cdots & L_{ll}^E \end{pmatrix} \begin{pmatrix} \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_l \end{pmatrix} = \begin{pmatrix} \mathbf{f}_0 \\ \vdots \\ \mathbf{f}_l \end{pmatrix}, \quad (5.14)$$

wobei hier in jedem Block

$$L_{ij}^E := (B(\varphi_{\mathbf{p}_1, \mathbf{q}_1}, \varphi_{\mathbf{p}_2, \mathbf{q}_2}))_{\varphi_{\mathbf{p}_1, \mathbf{q}_1} \in \mathcal{E}_{\tilde{V}_l^0}^i, \varphi_{\mathbf{p}_2, \mathbf{q}_2} \in \mathcal{E}_{\tilde{V}_l^0}^j}$$

alle Funktionen des Erzeugendensystems, die zu dem gleichem Ortslevel gehören, enthalten sind. Zunächst unterteilen wir die zu den einzelnen Blöcken gehörigen Unbekannten weiter, so dass alle Unbekannten, die zu der gleichen Basisfunktion im Ort gehören, in der gleichen Partition enthalten sind. Dies führt für die Diagonalblöcke  $L_{kk}^E$  auf die Blockstruktur

$$L_{kk}^E = \begin{pmatrix} T_{11}^k & \cdots & T_{1N}^k \\ \vdots & \cdots & \vdots \\ T_{N1}^k & \cdots & T_{NN}^k \end{pmatrix},$$

wobei  $N := \dim(V_k^\Omega)$ . Ein Schritt des Block-Gauß-Seidel Verfahrens bzgl. dieser Blockstruktur für die Lösung von  $L_{kk}^E x = b$  entspricht der Gauß-Seidel Waveform Relaxation (5.13), wobei die Ortsdiskretisierung bzgl.  $V_k^\Omega$  durchgeführt ist und das System gewöhnlicher Differentialgleichungen mit dem der Raum-Zeit Dünngitterdiskretisierung zugrunde liegendem Zeitschrittverfahren diskretisiert wurde. Insgesamt kann damit ein Schritt des symmetrischen Gauß-Seidel Verfahrens über das derartig angeordnete System als Raum-Zeit Dünngitter Version der Multigrid-Waveform Relaxation betrachtet werden. Abbildung 5.3 veranschaulicht die Durchlaufreihenfolge des Verfahrens für  $d = 1$  an dem Beispiel von  $\tilde{V}_3^0$ .

## 5.4 Numerische Ergebnisse

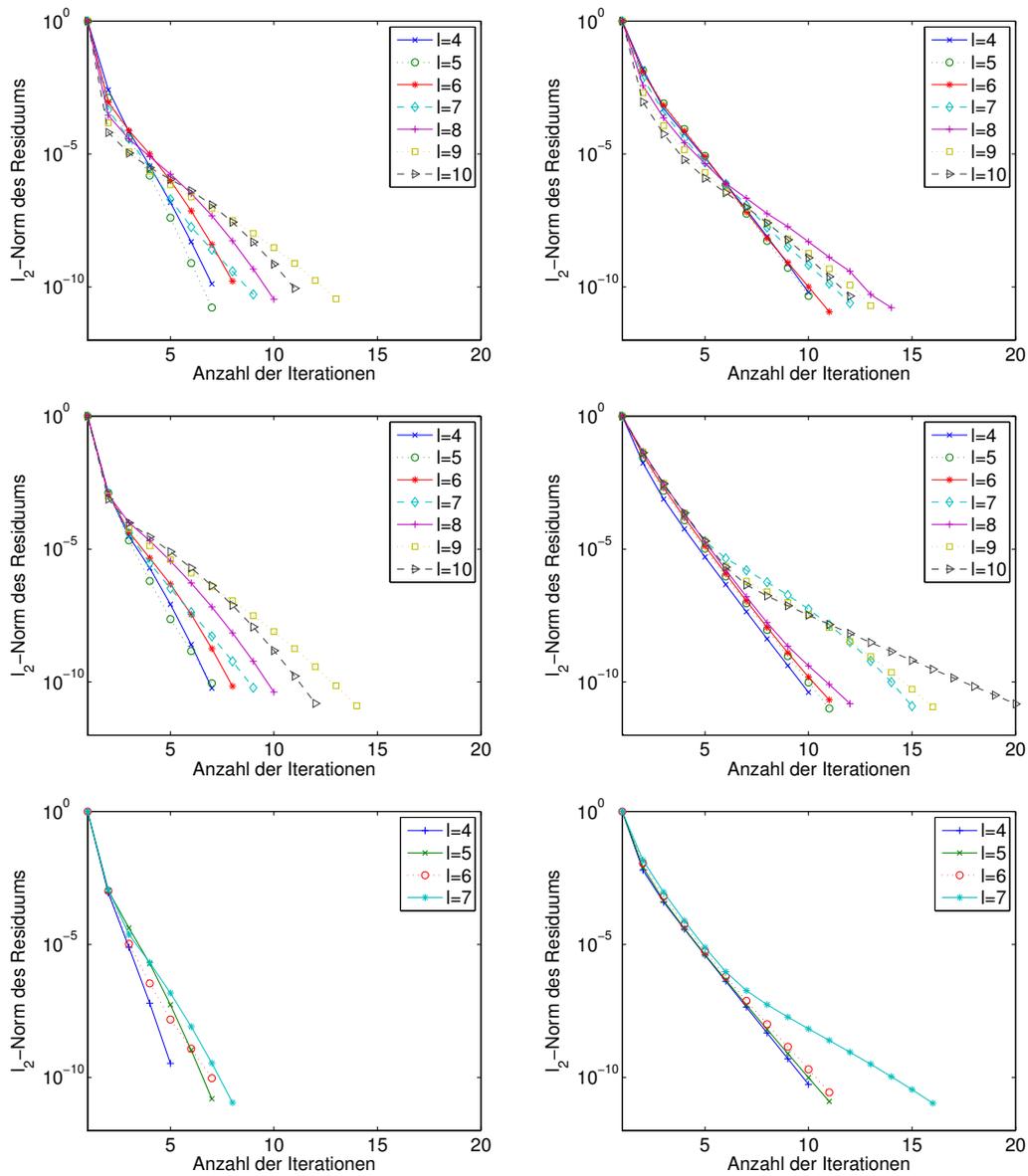
In diesem Abschnitt untersuchen wir die Konvergenzeigenschaften der vorgestellten Verfahren. Dazu setzen wir bei allen betrachteten Problemen die rechte Seite, sowie die Anfangsbedingung und die Randbedingungen auf Null, so dass der Nullvektor die Lösung des linearen Gleichungssystems ist. Wir bezeichnen mit  $\|\cdot\|_2$  die  $l_2$ -Norm, d.h.  $\|x\|_2^2 = \sum_{i=1}^n x_i^2$  für jedes  $x \in \mathbb{R}^n$ . Als Startwert wählen wir einen Zufallsvektor  $x_0$  der derart normiert ist, dass

$$\|L^E x_0\|_2 = 1$$

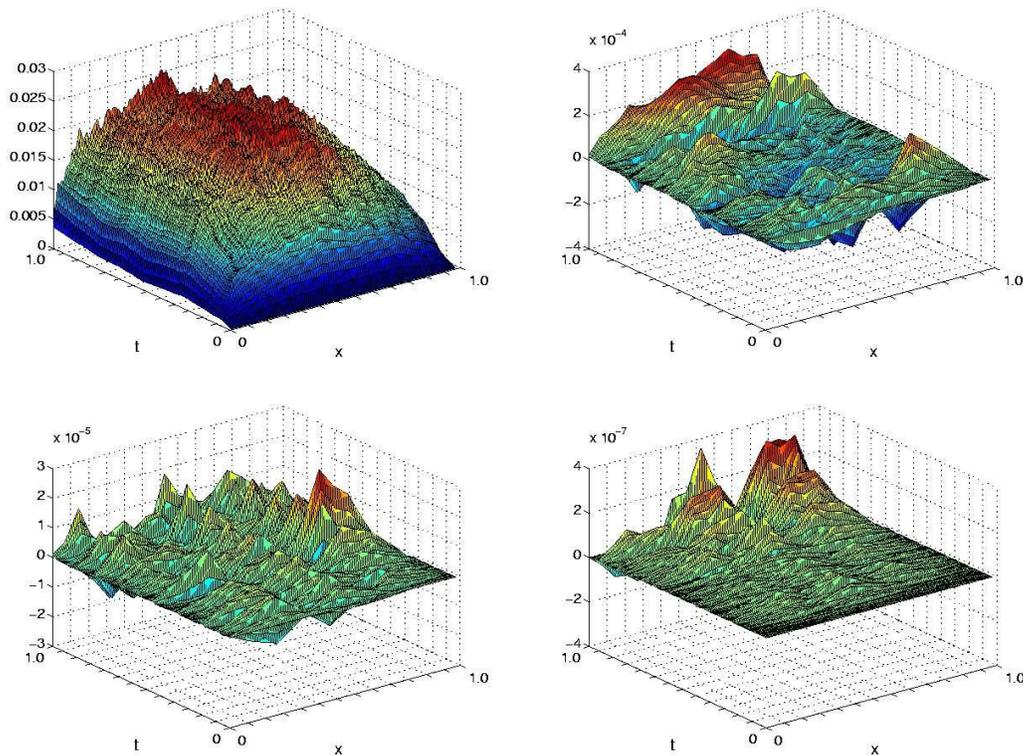
gilt. Die Iteration wird gestoppt, wenn das  $l_2$ -Residuum  $\|L^E x^i\|_2$  der  $i$ -ten Iterierten  $x_i$  kleiner als  $10^{-11}$  ist. Wir messen im folgenden die Fehlerreduktion

$$\rho = \left( \frac{\|L^E x_2\|_2}{\|L^E x_i\|_2} \right)^{1/(i-2)}$$

wobei  $x_i$  die Iterierte der letzten Iteration ist, bevor das Abbruchkriterium erfüllt ist. Wir beziehen die ersten zwei Iterierten nicht in die Messung mit ein, da dort die Reduktionsraten je nach Problem und Verfahren übermäßig gut sein können und somit das Ergebnis verfälschen würden. Da wir in den Experimenten keinen Unterschied in dem Löserverhalten für die Diskretisierung mit  $V_l^0$  und  $\tilde{V}_l^0$  beobachtet haben, beschränken wir uns auf die Diskussion der Ergebnisse für  $\tilde{V}_l^0$ .



**Abbildung 5.4.** Konvergenzhistorie des Gauß-Seidel (links) und des Waveform-Gauß-Seidel Verfahrens (rechts) für das mit dem Crank-Nicolson Verfahren diskretisierte Problem (5.15) und  $d = 1$  (oben),  $d = 2$  (mitte) und  $d = 3$  (unten).



**Abbildung 5.5.** Startfehler (links oben), Fehler nach ein (rechts oben), zwei (links unten) und vier (rechts unten) Iterationen des Gauß-Seidel Verfahrens für die Crank-Nicolson Diskretisierung des Problems (5.15) auf Level  $l = 7$ .

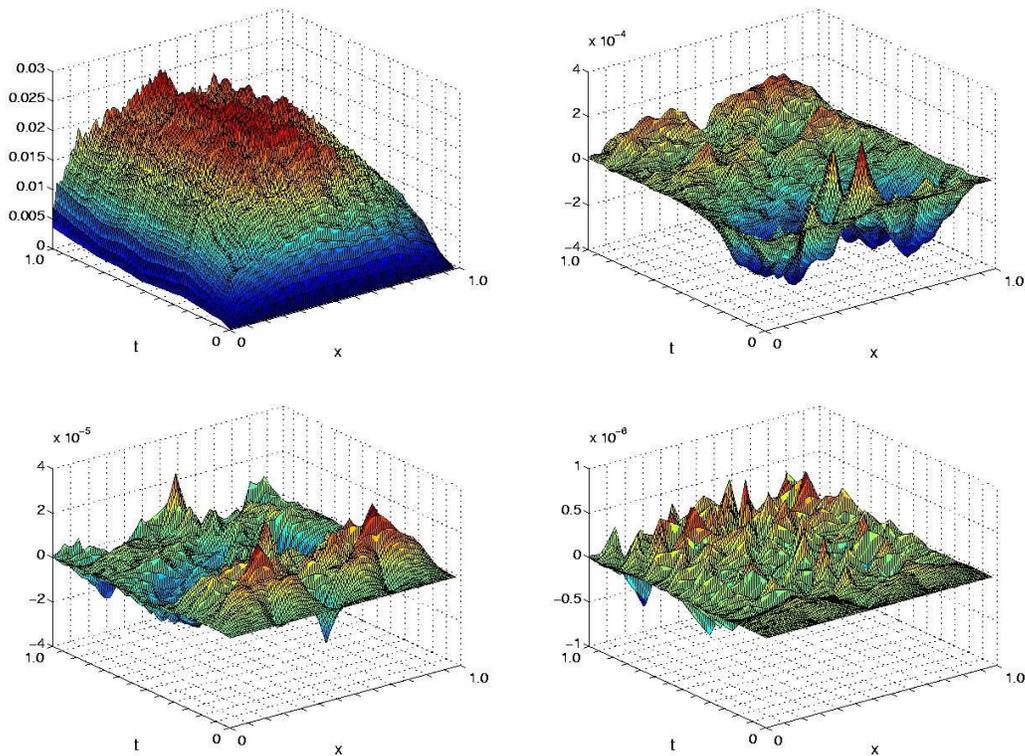
### 5.4.1 Erste numerische Ergebnisse und das kombinierte Gauß-Seidel Verfahren

In diesem Abschnitt untersuchen wir die Konvergenzeigenschaften des Gauß-Seidel und des Waveform-Gauß-Seidel Verfahrens für das Problem

$$\partial_t u - \Delta u = 0 \text{ in } \Omega_T = (0, 1)^d \times (0, 1] \quad (5.15)$$

mit  $d = 1, 2, 3$ . Abbildung 5.4 zeigt die Konvergenzhistorie des Gauß-Seidel und des Waveform-Gauß-Seidel Verfahrens für das mit dem Crank-Nicolson Verfahren auf unterschiedlichen Leveln diskretisierte Problem (5.15).

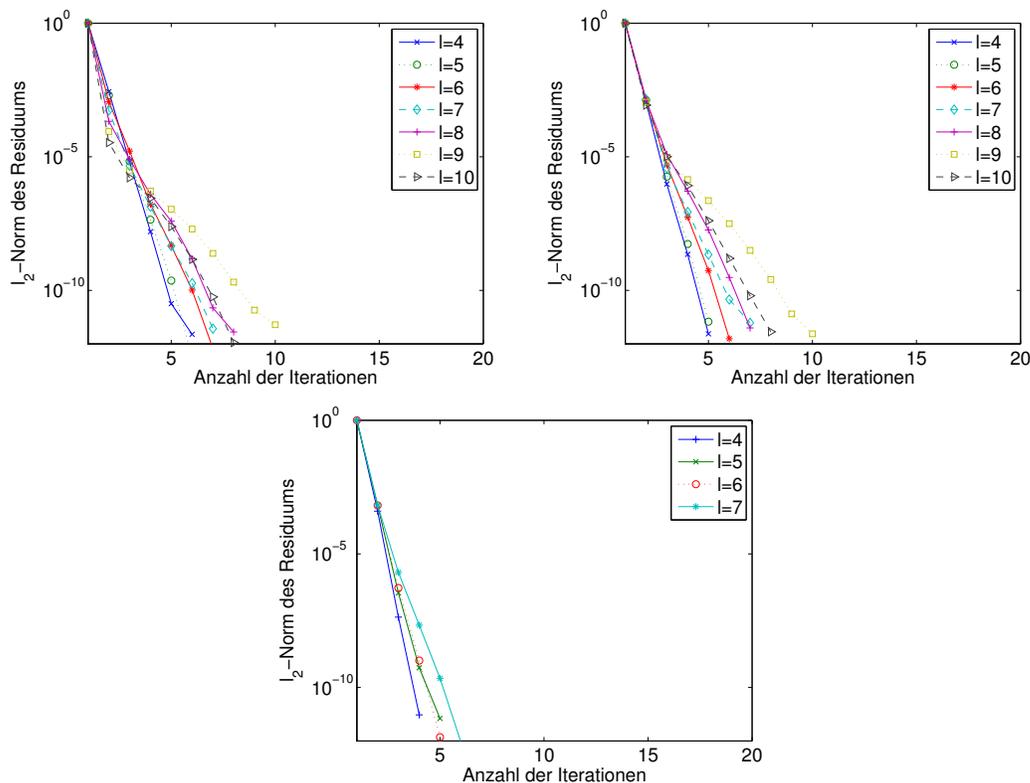
Betrachten wir zunächst die Ergebnisse für das Gauß-Seidel Verfahren. Hierbei ist eine besonders große Reduktion des Residuums durch die ersten zwei Iterationen insbesondere für  $d = 1, 2$  auffällig. Das Gauß-Seidel Verfahren liefert insgesamt sehr gute Konvergenzraten. Für den eindimensionalen Fall erhalten wir beispielsweise auf Level  $l = 10$  eine Reduktionsrate  $\rho = 0.21$ , während  $d = 2$  und Level  $l = 10$



**Abbildung 5.6.** Startfehler (links oben), Fehler nach ein (rechts oben), zwei (links unten) und vier (rechts unten) Iterationen des Waveform-Gauß-Seidel Verfahrens für die Crank-Nicolson Diskretisierung des Problems (5.15) auf Level  $l = 7$ .

zu  $\rho = 0.16$  und  $d = 3$  mit  $l = 7$  zu  $\rho = 0.05$  führen. Allerdings beobachten wir mit wachsendem Level einen leichten Anstieg der Anzahl der benötigten Iterationen. Dabei ist die Konvergenz der ersten zwei bis drei Iterationen und der letzten Iterationen, bevor das Abbruchkriterium erreicht wird, nahezu levelunabhängig. Zum Beispiel erhalten wir im zweidimensionalen Fall  $d = 2$  eine Reduktion des Residuums durch die letzte Iteration von 0.05 für  $l = 5$  und 0.08 für  $l = 10$ . Dieses Verhalten deutet daraufhin, dass einige Fehlerkomponenten, die nicht optimal reduziert werden können, nach den ersten Iterationen dominieren, bis nach entsprechend vielen weiteren Iterationen wiederum die Fehlerkomponenten auftreten, die mit besseren Raten reduziert werden. Hierzu sind für den eindimensionalen Fall  $d = 1$  in Abbildung 5.5 der Startfehler und die Fehler nach ein, zwei und vier Iterationen des Gauß-Seidel Verfahrens dargestellt.

Dabei ist beachtenswert, dass bereits nach wenigen Iterationen der Fehler zum Endzeitpunkt dominiert. Obwohl wir dieses Fehlerverhalten durchaus erwarten dürfen,

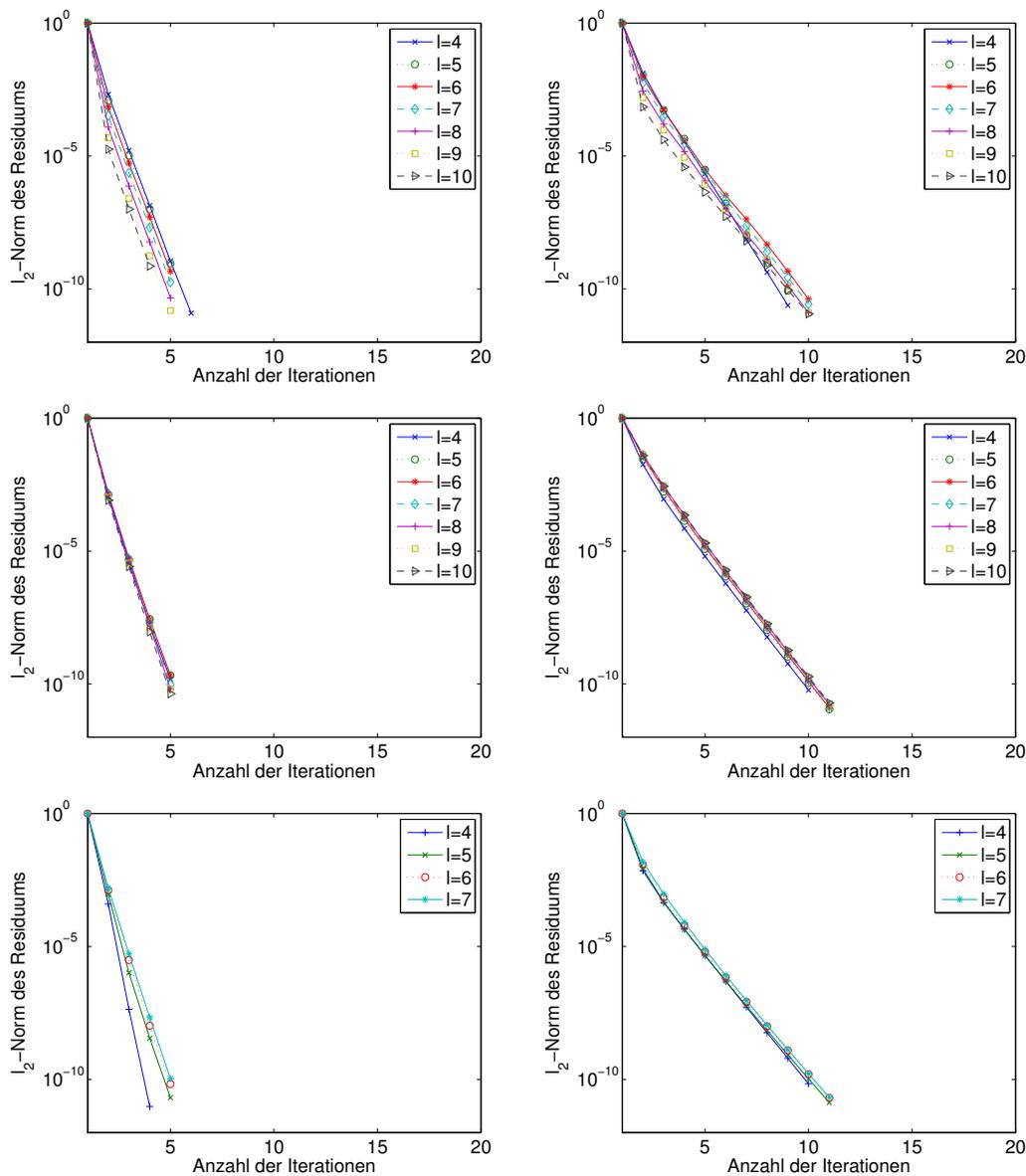


**Abbildung 5.7.** Konvergenzhistorie des aus dem Waveform-Gauß-Seidel und des Gauß-Seidel Verfahrens konstruierten Verfahrens für den mit dem Crank-Nicolson Verfahren diskretisierten Operator (5.15) und  $d = 1$  (oben links),  $d = 2$  (oben rechts) und  $d = 3$  (unten).

schließlich entspricht der Endzeitpunkt dem freien Rand und ein ähnliches Verhalten beobachtet man auch für Mehrgitterlöser im Fall von elliptischen Problemen am Neumann-Rand, ist dieser Effekt überraschend stark ausgeprägt und möglicherweise die Ursache für das beobachtete Verhalten des Löser.

Wir diskutieren nun die Ergebnisse für das Waveform-Gauß-Seidel Verfahren in Abbildung 5.4. Während die Konvergenzrate im eindimensionalen Fall  $d = 1$  unabhängig von der Levelzahl beschränkt ist, beobachten wir für  $d = 2$  und  $d = 3$  eine Verschlechterung der Konvergenz des Waveform-Gauß-Seidel Verfahrens mit wachsendem Level.

Vergleichen wir die in Abbildung 5.6 dargestellte Fehlerentwicklung des Waveform-Gauß-Seidel Verfahrens nach ein, zwei und vier Iterationen mit der Fehlerentwicklung des Gauß-Seidel Verfahrens, so sehen wir deutlich, dass beide Verfahren bereits nach wenigen Iterationen zu sehr unterschiedlichen Fehlern führen. Während der Fehler des Waveform-Gauß-Seidel Verfahrens relativ gleichmäßig über den ge-



**Abbildung 5.8.** Konvergenzhistorie des Gauß-Seidel (links) und des Waveform-Gauß-Seidel Verfahrens (rechts) für den mit dem Euler Verfahren diskretisierten Operator (5.15) und  $d = 1$  (oben),  $d = 2$  (mitte) und  $d = 3$  (unten).

samten Raum-Zeit Zylinder verteilt ist, dominiert bereits nach wenigen Iterationen des Gauß-Seidel Verfahrens der Fehler zum Endzeitpunkt. Dabei ist der Fehler des Gauß-Seidel Verfahrens kleiner als der Fehler des Waveform-Gauß-Seidel Verfahrens.

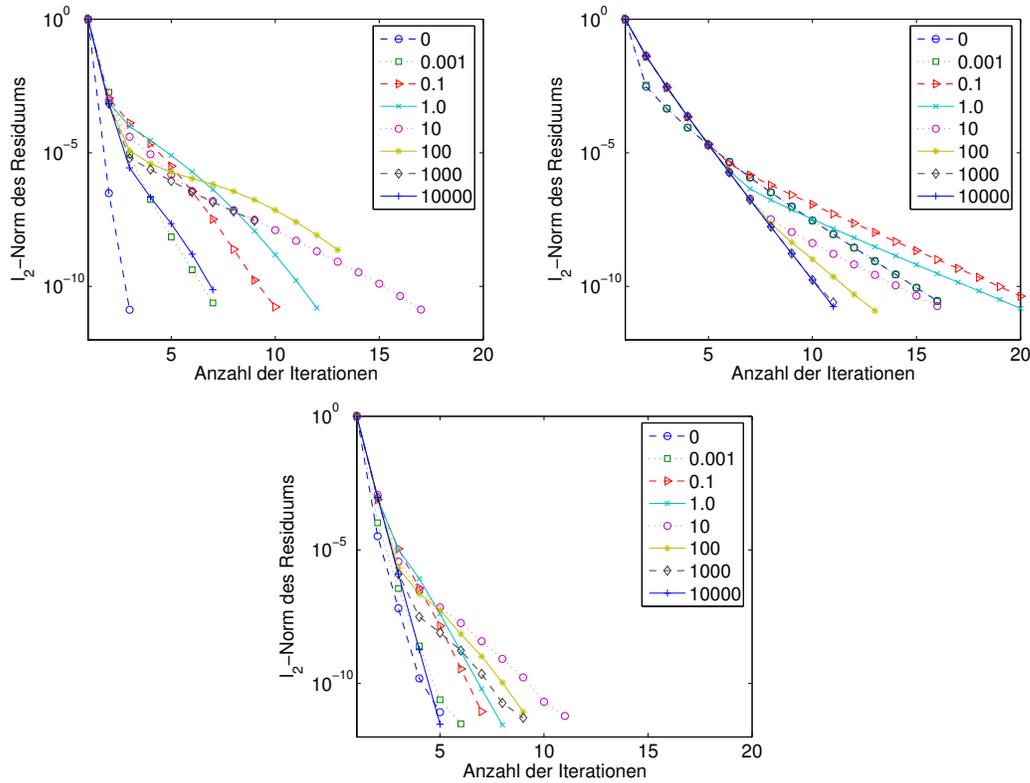
Das unterschiedliche Fehlerverhalten beider Verfahren motiviert dazu, das Gauß-

Seidel Verfahren mit dem Waveform-Gauß-Seidel Verfahren zu kombinieren. Dabei hoffen wir, dass das Waveform-Gauß-Seidel Verfahren die Fehleranteile effizient reduziert, die bei dem Gauß-Seidel Verfahren zu schlechteren Reduktionsraten für einige Iterationen führen. Umgekehrt sollte gleichzeitig das Gauß-Seidel Verfahren die Fehlerkomponenten stark dämpfen, für die das Waveform-Gauß-Seidel Verfahren nur langsam konvergiert.

Die Kombination verschiedener Durchlaufreihenfolgen ist bereits in [92] mit dem so genannten  $Q$ -Zyklus im Dünngitterkontext für elliptische Probleme vorgeschlagen worden. Mit dem  $Q$ -Zyklus wurden dabei in numerischen Experimenten gute Konvergenzresultate erzielt. Wir definieren daher eine neue Iteration, bei der in jedem Iterationsschritt zunächst ein Vorwärtsschritt des Gauß-Seidel Verfahrens und danach ein Rückwärtsschritt des Waveform-Gauß-Seidel Verfahrens ausgeführt werden. Die Konvergenzhistorien dieses kombinierten Gauß-Seidel Verfahrens sind in Abbildung 5.7 dargestellt. Die Ergebnisse zeigen deutlich, dass das resultierende Verfahren schneller als das Gauß-Seidel Verfahren konvergiert. Wir erhalten so zum Beispiel für  $l = 10$  und  $d = 2$  eine Reduktionsrate  $\rho = 0.05$ . Zusätzlich scheint die Residuenreduktion unabhängig von der Anzahl der Iterationen zu sein, d.h. wir erhalten während der Iteration keine Fehlerkomponenten, die langsamer als andere reduziert werden.

Wenden wir uns jetzt den Ergebnissen für die Euler Diskretisierung des Problems (5.15) zu. Man beachte, dass der symmetrische Anteil des aus der Euler Diskretisierung resultierenden linearen Gleichungssystems positiv semidefinit ist. Scheinbar ist dies eine wichtige Eigenschaft für die Konvergenz der Iterationsverfahren bei unsymmetrischen Problemen, wie wir dies bereits an Satz 5.1 über die Konvergenz der SOR-Iteration gesehen haben. Wir erwarten daher, dass die Verfahren im Fall der Euler Diskretisierung mindestens so gute Resultate liefern wie für die Crank-Nicolson Diskretisierung, deren symmetrischer Anteil nicht positiv definit ist.

In der Tat zeigen die Ergebnisse in Abbildung 5.8, dass sowohl das Gauß-Seidel als auch das Waveform-Gauß-Seidel Verfahren sehr schnell mit levelunabhängigen Reduktionsraten konvergieren. Im Fall des Gauß-Seidel Verfahrens verkleinern sich die Raten sogar mit zunehmendem Level. So liefert das Gauß-Seidel Verfahren in zwei Ortsdimensionen eine Reduktionsrate von  $\rho = 0.007$  auf Level  $l = 5$  und auf Level  $l = 10$  den Wert  $\rho = 0.004$ . Die Kombination aus dem Gauß-Seidel und dem Waveform-Gauß-Seidel Verfahren liefert hingegen noch etwas bessere Resultate als das Gauß-Seidel Verfahren. Allerdings sind die Konvergenzraten des Gauß-Seidel Verfahrens und des Waveform-Gauß-Seidel Verfahrens bereits so gut, dass wir an dieser Stelle auf die Präsentation der Ergebnisse für das kombinierte Verfahren verzichten möchten.



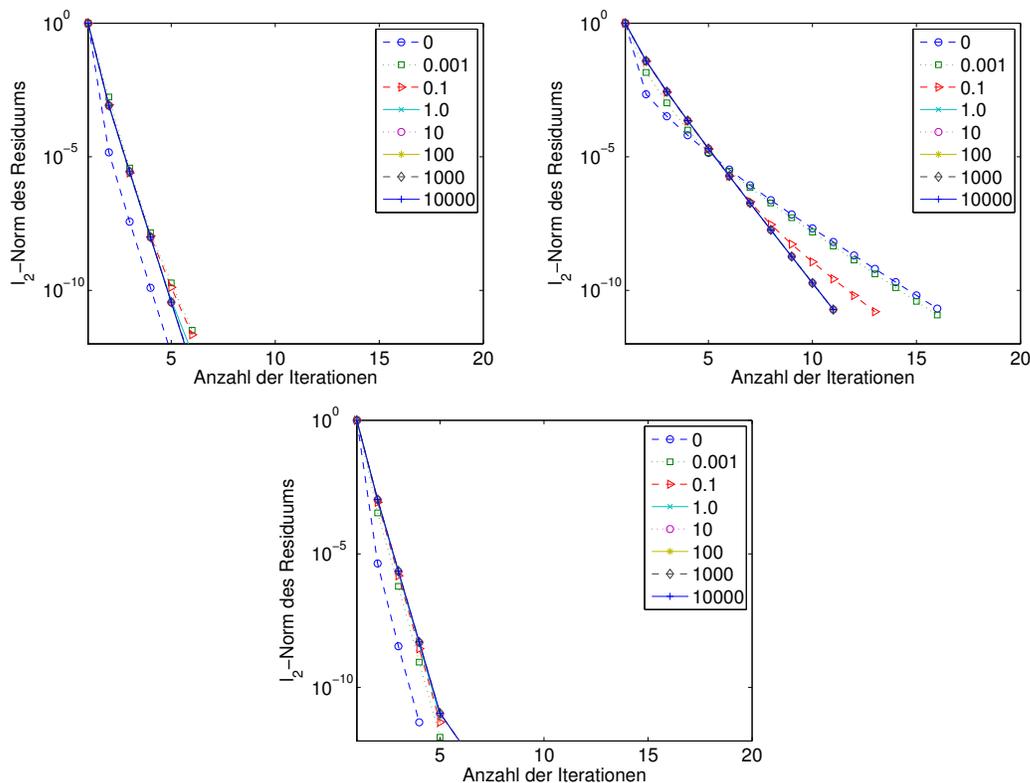
**Abbildung 5.9.** Konvergenzhistorie des Gauß-Seidel Verfahrens (oben links), des Waveform-Gauß-Seidel Verfahrens (oben rechts) und der Kombination aus beiden Verfahren (unten) für das mit dem Crank-Nicolson Verfahren auf Level  $l = 10$  diskretisierte Problem (5.16) und verschiedenen  $\varepsilon$ .

### 5.4.2 Numerische Ergebnisse für ein zeitlich anisotropes Problem

Die Länge des Zeitintervalls  $(0, T]$ , für das eine Lösung gesucht ist, wird in realen Anwendungen stark variieren. In unserer Formulierung bedeutet das, dass wir nach Transformation auf das Einheitsintervall  $(0, 1]$  einen entsprechenden Faktor vor dem elliptischen Teil des parabolischen Operators erhalten. Wir möchten nun die Robustheit der Verfahren gegenüber einem solchen Parameter  $\varepsilon \geq 0$  untersuchen. Dazu betrachten wir das Problem

$$\partial_t u - \varepsilon \cdot \Delta u = 0 \quad (5.16)$$

auf  $\Omega = (0, 1)^2$ . Dieses Beispiel ist nicht nur interessant weil wir einen solchen Parameter durch Transformation auf das Einheitsintervall erhalten, sondern auch die Verwendung von unterschiedlichen Materialien mit unterschiedlichen Wärmeleitkoeffizienten führt zu einer Reihe von verschiedenen Werten für  $\varepsilon$ .



**Abbildung 5.10.** Konvergenzhistorie des Gauß-Seidel Verfahrens (oben links), des Waveform-Gauß-Seidel Verfahrens (oben rechts) und der Kombination aus beiden Verfahren (unten) für das mit dem Euler Verfahren auf Level  $l = 10$  diskretisierte Problem (5.16) und verschiedenen  $\varepsilon$ .

Abbildung 5.9 zeigt die Konvergenzhistorie der Verfahren für die Crank-Nicolson Diskretisierung auf Level  $l = 10$  des Problems für verschiedene Werte  $\varepsilon$ . Wie in Abschnitt 5.4.1 konvergiert die Kombination aus dem Waveform-Gauß-Seidel und dem Gauß-Seidel Verfahren am schnellsten. Auffällig ist, dass für den Grenzfall  $\varepsilon = 0$  die Konvergenz der drei Verfahren sehr schnell ist, dann für wachsendes  $\varepsilon$  etwas schlechter wird, und ab einem gewissen Schwellwert die Konvergenz wieder beschleunigt.

Interessant ist hierbei das unterschiedliche Verhalten des Waveform-Gauß-Seidel und des Gauß-Seidel Verfahrens. Während das Gauß-Seidel Verfahren für den Grenzfall  $\varepsilon = 0$  schneller konvergiert als für  $\varepsilon = 10^4$ , liefert das Waveform-Gauß-Seidel Verfahren für  $\varepsilon = 10^4$  bessere Ergebnisse als im Fall  $\varepsilon = 0$ . Dementsprechend erhalten wir für die Kombination aus beiden Verfahren nahezu die gleichen Reduktionsraten für  $\varepsilon = 0$  und  $\varepsilon = 10^4$ .

Die Ergebnisse für die Euler Diskretisierung auf Level  $l = 10$  sind in Abbil-

**Tabelle 5.1.** *Ergebnisse des Gauß-Seidel Verfahrens (GS), des Waveform-Gauß-Seidel Verfahrens (WGS) und der Kombination aus beiden Verfahren (KGS) für die Crank-Nicolson Diskretisierung von (5.17) auf  $\tilde{V}_l^0$ .*

1	$\gamma = 0$			$\gamma = 1$			$\gamma = 2$			$\gamma = 3$		
	GS	WGS	KGS									
6	0.05	0.1	0.01	0.13	0.25	0.02	0.08	0.37	0.07	0.16	0.38	0.09
7	0.07	0.22	0.03	0.12	0.27	0.05	0.18	0.43	0.10	0.12	0.41	0.13
8	0.13	0.13	0.02	0.28	0.30	0.10	0.35	0.48	0.14	0.14	0.50	0.16
9	0.23	0.29	0.11	0.32	0.40	0.16	0.30	0.52	0.16	0.25	0.55	0.20
10	0.16	0.46	0.05	0.30	0.38	0.14	0.56	0.44	0.26	0.49	0.60	0.24

**Tabelle 5.2.** *Ergebnisse des Gauß-Seidel Verfahrens (GS), des Waveform-Gauß-Seidel Verfahrens (WGS) und der Kombination aus beiden Verfahren (KGS) für die Euler Diskretisierung von (5.17) auf  $\tilde{V}_l^0$ .*

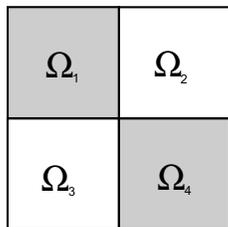
1	$\gamma = 0$			$\gamma = 1$			$\gamma = 2$			$\gamma = 3$		
	GS	WGS	KGS	GS	WGS	KGS	GS	WGS	KGS	GS	WGS	KGS
6	0.007	0.09	0.007	0.04	0.25	0.04	0.08	0.37	0.09	0.09	0.38	0.10
7	0.005	0.09	0.005	0.05	0.28	0.04	0.12	0.43	0.13	0.13	0.44	0.12
8	0.004	0.09	0.004	0.06	0.31	0.05	0.16	0.48	0.17	0.18	0.50	0.17
9	0.006	0.10	0.006	0.07	0.33	0.08	0.20	0.52	0.19	0.22	0.55	0.23
10	0.006	0.10	0.006	0.08	0.35	0.10	0.23	0.53	0.23	0.26	0.59	0.26

dung 5.9 festgehalten. Hierbei ist die Konvergenz des Gauß-Seidel und des aus dem Gauß-Seidel und dem Waveform-Gauß-Seidel Verfahren kombinierten Verfahrens unabhängig von  $\varepsilon$ . Im Falle des Waveform-Gauß-Seidel Verfahrens beobachten wir zudem eine Verbesserung mit wachsendem  $\varepsilon$ . So konvergiert das Waveform-Gauß-Seidel Verfahren für  $\varepsilon = 0$  mit einer Reduktionsrate von  $\rho = 0.3$ , wogegen wir für  $\varepsilon = 10$  eine Reduktion von  $\rho = 0.09$  erhalten, die sich mit wachsendem  $\varepsilon$  auch nicht weiter verändert.

Insgesamt verhalten sich alle drei Verfahren robust gegenüber dem Parameter  $\varepsilon$ , wobei die Kombination aus dem Gauß-Seidel und dem Waveform-Gauß-Seidel Verfahren sowohl für die Crank-Nicolson als auch für die Euler Diskretisierung in diesem Zusammenhang zu den besten Resultaten führt.

### 5.4.3 Numerische Ergebnisse für Probleme mit springenden Koeffizienten

In der Praxis werden häufig Probleme betrachtet, bei denen unterschiedliche Materialien aneinander grenzen. Diese unterschiedlichen Materialien besitzen verschiedene Materialparameter, die direkt als Koeffizienten in den elliptischen Teil des parabolischen Operators eingehen. Dies führt zu unstetigen Koeffizientenfunktionen



**Abbildung 5.11.** Zerlegung des zu dem Problem (5.17) gehörigen Gebietes  $\Omega$ .

im Raum-Zeit Operator.

Im Fall von stationären elliptischen Problemen verschlechtert sich die Konvergenz von klassischen geometrischen Mehrgittermethoden mit zunehmender Sprunggröße. Somit sind die Verfahren im Sinn von [120] nicht robust. Aus diesem Grund können wir im parabolischen Fall für die in diesem Kapitel vorgestellten Löser keineswegs ein von der Sprunggröße unabhängiges Konvergenzverhalten erwarten.

An dieser Stelle wenden wir uns den Konvergenzeigenschaften der Verfahren bezüglich solcher Sprünge zu. Dazu unterteilen wir das Gebiet  $\Omega = (0, 1)^2$  wie in Abbildung 5.11 dargestellt in die vier Teilgebiete  $\Omega_1$ ,  $\Omega_2$ ,  $\Omega_3$  und  $\Omega_4$ . Wir betrachten für festes  $\gamma$  das Problem

$$\partial_t u - \operatorname{div}(A(x)\nabla u) = 0 \quad (5.17)$$

mit

$$A(x) = \begin{cases} 1.0 & \text{falls } x \in \Omega_1 \cup \Omega_4 \\ 10^\gamma & \text{falls } x \in \Omega_2 \cup \Omega_3. \end{cases}$$

Für  $\gamma = 0$  erhalten wir damit das Problem (5.15).

Die Ergebnisse für die Crank-Nicolson Diskretisierung des Problems (5.17) mit  $\gamma \in \{0, 1, 2, 3\}$  sind in Tabelle 5.1 wiedergegeben. Die Reduktionsrate des Residuums nimmt für alle drei Verfahren mit wachsendem  $\gamma$  zu. Festzustellen ist im Fall der Kombination aus dem Gauß-Seidel und dem Waveform-Gauß-Seidel Verfahren eine Reduktionsrate von  $\rho = 0.05$  für  $l = 10$  und  $\gamma = 0$ , während wir für  $\gamma = 3$  eine Reduktionsrate von  $\rho = 0.24$  erhalten. Im Ergebnis liefert die Kombination des Gauß-Seidel und des Waveform-Gauß-Seidel Verfahrens allerdings gute Resultate, wobei wir selbst für den größten betrachteten Sprung mit  $\gamma = 3$  eine Konvergenzrate  $\rho = 0.24$  auf Level  $l = 10$  erhalten.

Wir nehmen nun die Ergebnisse für die Euler Diskretisierung in Tabelle 5.2 in Augenschein. Wie schon bei der Crank-Nicolson Diskretisierung verschlechtert sich auch hier die Konvergenz mit wachsendem  $\gamma$ , so dass wir beispielsweise für  $\gamma = 3$  eine Reduktionsrate  $\rho = 0.26$  für das Gauß-Seidel Verfahren auf Level  $l = 10$  messen, während  $\gamma = 0$  zu der Rate  $\rho = 0.006$  führt. Insbesondere ist die Rate  $\rho = 0.26$  für  $l = 10$  und  $\gamma = 3$  für das aus dem Gauß-Seidel und dem Waveform-Gauß-Seidel Verfahren kombinierte Verfahren nahezu identisch zu der Rate  $\rho = 0.24$  im Falle der

Crank-Nicolson Diskretisierung, obwohl wir für  $\gamma = 0$  bei der Euler Diskretisierung wesentlich bessere Ergebnisse erzielt haben.

Insgesamt beobachten wir unabhängig von der gewählten Diskretisierung jeweils eine Verschlechterung der Konvergenz mit wachsendem Sprung, wie wir dies auch für geometrische Mehrgittermethoden bei elliptischen Problemen mit Sprüngen erwarten. Dennoch liefert die Kombination aus dem Waveform-Gauß-Seidel und dem Gauß-Seidel Verfahren auch für den Fall  $\gamma = 3$  noch akzeptable Resultate, so dass wir dennoch eine recht große Problemklasse effizient lösen können. Um die Resultate weiter zu verbessern und auch noch größere Sprünge behandeln zu können, sind Modifikationen der Raum-Zeit Multilevellöser wie sie auch bei geometrischen Mehrgitterverfahren durchgeführt wurden, denkbar. So könnten die Prolongationsoperatoren für den Ortsanteil wie im elliptischen Fall matrixabhängig (bzgl. des elliptischen Teils des parabolischen Operators) berechnet werden [45, 118, 119]. Ebenso ist eine matrixabhängige Vergrößerung wie bei den algebraischen Mehrgitterverfahren [26, 96] denkbar. Dabei sollten die algebraisch konstruierten Prolongationsoperatoren und Gitterhierarchien nur zur Lösung der Gleichungssysteme, nicht aber zur Konstruktion der Multilevelbasis im Ort verwendet werden, da sich sonst u. U. die Approximationsraten der Raum-Zeit Dünngitter verschlechtern. Der Einsatz matrixabhängiger Prolongationen und Vergrößerungsalgorithmen soll an dieser Stelle nicht weiter vertieft werden, da das den Umfang dieser Arbeit sprengen würde, sondern Gegenstand zukünftiger Forschung sein.

In diesem Kapitel haben wir zwei verschiedene Multilevelverfahren zur Lösung der mit dem Raum-Zeit Dünngitter Crank-Nicolson bzw. Euler Verfahren diskretisierten linearen Gleichungssysteme vorgestellt. Hierbei haben wir Multilevelverfahren, wie in [61] beschrieben, als klassische Iterationsverfahren über einem mittels eines Erzeugendensystems diskretisierten System interpretiert. Wir haben gezeigt, dass das SOR-Verfahren mit geeignet gewähltem Relaxationsparameter für das semidefinite, unsymmetrische Gleichungssystem der Raum-Zeit Eulerdiskretisierung mittels des Erzeugendensystems konvergiert. Durch dieses Ergebnis motiviert haben wir dann auf Basis des (Block-) Gauß-Seidel Verfahrens zwei verschiedene Multilevellöser vorgestellt und deren Konvergenzeigenschaften an zahlreichen Beispielen untersucht. Hierbei hat vor allem die Kombination aus beiden Verfahren sowohl für die Crank-Nicolson als auch für die Euler Diskretisierung zu von der Levelzahl unabhängigen Konvergenzraten geführt. Des Weiteren verhält sich das Verfahren robust gegenüber dem Verhältnis von Zeitschrittweite und Ortsauflösung. Wie bei geometrischen Mehrgittermethoden im elliptischen Fall sind die vorgestellten Löser nicht robust gegenüber Sprüngen in den Koeffizienten des elliptischen Anteils des parabolischen Operators, liefern aber für moderat große Sprünge immer noch gute Resultate, so dass auch solche Probleme effizient gelöst werden können. Für große Sprünge könnte der Einsatz von matrixabhängigen Prolongationsoperatoren und Vergrößerungen ähnlich wie im elliptischen Fall zu robusten Verfahren führen. Dies

sollte Gegenstand zukünftiger Forschung sein.



## Kapitel 6

# Adaptivität in Raum-Zeit

Die numerischen Experimente in Abschnitt 4.4 dokumentieren, dass das Raum-Zeit Dünngitter Crank-Nicolson sowie das Euler und das Discontinuous-Galerkin Verfahren mit stückweise linearen Ansatzfunktionen in der Zeit die gleichen Fehlerreduktionsraten aufweisen, wie der entsprechende Interpolant des Ansatzraumes der exakten Lösung. Wie in Abschnitt 2.4 beschrieben, benötigen wir allerdings eine gewisse Regularität der zu interpolierenden Funktion, um die entsprechenden Fehlerreduktionsraten zu erhalten. Die Lösungen der in dem Abschnitt 4.4 betrachteten parabolischen Probleme wiesen dabei allesamt die benötigte Glattheit auf.

Wie die Diskussion in Abschnitt 3.2 gezeigt hat, erfüllt die Lösung unter geeigneten Annahmen an das Gebiet, die rechte Seite, die Anfangs- und Randbedingungen des parabolischen Problems diese Glattheitsvoraussetzungen. Allerdings sind die Einschränkungen an das Gebiet und die Bedingungen an die rechte Seite, die Anfangs- und Randbedingungen in realen Anwendungen nicht immer erfüllt. Somit besitzt auch die Lösung der parabolischen Gleichung unter Umständen eine geringere Glattheit. In diesem Fall verschlechtern sich jedoch die Approximationsraten der entsprechenden Raum-Zeit Dünngitterräume. Somit benötigen wir zur Diskretisierung dieser Probleme wesentlich mehr Freiheitsgrade um die gleiche Größenordnung des Fehlers, wie bei Problemen mit ausreichend glatten Lösungen, zu erhalten. Doch selbst wenn das betrachtete Problem eine genügend glatte Lösung  $u$  besitzt, kann die entsprechende Norm von  $u$ , die in der Interpolationsfehlerabschätzung auftaucht, sehr groß sein. Wir erhalten dann zwar die von der Theorie vorhergesagten Reduktionsraten, allerdings ist der Fehler auf dem größten Gitter so groß, dass dennoch eine sehr feine Auflösung notwendig ist, um die gewünschte Genauigkeit zu erreichen.

Abhilfe schaffen hier adaptive Diskretisierungsverfahren. Hierbei wird der Fehler zwischen exakter und diskretisierter Lösung zunächst mit Hilfe so genannter a posteriori Fehlerschätzer bzw. Fehlerindikatoren geschätzt. Mit Hilfe dieser Information werden dann nur solche Freiheitsgrade zu der bestehenden Diskretisierung hinzugefügt, die den größten Einfluß auf den geschätzten Fehler haben. Für elliptische

Probleme, insbesondere in dem Kontext Finiter Elemente, existieren eine Vielzahl an Arbeiten zu dem Thema Fehlerschätzer und adaptive Verfeinerung, vgl. [115] und die Referenzen dort. Darüber hinaus wurden in den letzten Jahrzehnten auch zahlreiche Fortschritte für die klassische Diskretisierung parabolischer Probleme mittels der Rothe- oder der Linienmethode erzielt.

Im Fall der Linienmethode, bei der zunächst der Raum mittels Finiter Elemente diskretisiert wird, kann das anfallende System gewöhnlicher Differentialgleichungen mit Hilfe herkömmlicher adaptiver Zeitschrittverfahren gelöst werden. Um bei dieser Vorgehensweise den Ortsfehler zu kontrollieren, gelangen herkömmliche Fehlerschätzer für elliptische Probleme zur Anwendung, vgl. [2, 16, 17].

Die Interpretation der Rothe Methode, bei der zunächst die Zeit diskretisiert wird, kann abstrakt als Cauchy Problem in einem Funktionenraum aufgefasst werden [19, 20, 21]. Dabei wird die Diskretisierung der Ortsprobleme als Störung des diskreten Orbits des in der Zeit semidiskretisierten Problems interpretiert. Diese Störung muß soweit verkleinert werden, dass die Genauigkeit des semidiskreten Problems nicht verschlechtert wird. Um die Genauigkeit der Ortsdiskretisierung zu steuern, können wiederum herkömmliche adaptive Verfahren für elliptische Probleme verwendet werden.

Offensichtlich wird bei den obigen Vorgehensweisen weiterhin zwischen Raum und Zeit unterschieden, was sich durch die Trennung der Fehlerschätzung und Adaption für die Ortsprobleme auf der einen Seite und die Zeitschrittweite auf der anderen Seite bemerkbar macht. In vielen Fällen ist es allerdings sinnvoll, dass in manchen Regionen des Ortsgebietes größere Zeitschrittweiten verwendet werden als in anderen Regionen, was auch als *local time stepping* bezeichnet wird. Die Trennung zwischen Raum und Zeit in der Fehlerschätzung ist daher nicht von Vorteil, sondern lediglich damit zu erklären, dass ein volles Raum-Zeit Gitter aufgrund der größeren Komplexität durch die zusätzliche Zeitdimension im Vergleich zu stationären Problemen nicht mehr im Hauptspeicher des Computers zu behandeln ist. Hier stoßen auch die Discontinuous-Galerkin Methoden auf vollen Gittern an ihre Grenzen. Obwohl formal zunächst nicht zwischen Raum und Zeit unterschieden wird, und daher Fehlerschätzer in Raum und Zeit sehr einfach zu konstruieren sind [49, 50], sind auf Grund des limitierten Hauptspeichers die Verfahren derart angelegt, dass am Ende wiederum nur eine Sequenz von stationären Problemen gelöst werden muß. Aus diesem Grund ist eine wirkliche Raum-Zeit Adaptivität mit *local time stepping* auch in diesem Kontext in der algorithmischen Umsetzung ausgesprochen aufwändig.

Hier wird ein weiterer Vorteil der Diskretisierung mit Hilfe der Raum-Zeit Dünn-gitter erkennbar. Da dank der geringen Dimension der Raum-Zeit Dünngiterräume keine explizite Trennung von Raum und Zeit notwendig ist, sondern nun die volle Raum-Zeit Diskretisierung im Hauptspeicher verwaltet werden kann, ergibt sich durch das einfache Einfügen von Basisfunktionen von feineren Leveln das *local time stepping* auf natürliche Weise. Wir unterscheiden nicht zwischen einer Orts- oder

Zeitadaptivität, sondern erhalten bei richtiger Wahl eines Fehlerschätzers oder Fehlerindikators ein Raum-Zeit adaptives Verfahren mit local time stepping.

In diesem Kapitel diskutieren wir nun den Einsatz von Adaptivität im Kontext der Raum-Zeit Dünngitter. Hierzu übertragen wir in Abschnitt 6.1 einen Fehlerindikator, basierend auf den hierarchischen Überschüssen der diskreten Lösung, der für klassische Dünngitter sowohl bei elliptischen Problemen [29, 30, 92, 126] als auch bei parabolischen Problemen [5] gute Resultate geliefert hat. Dabei folgen wir im Wesentlichen der Darstellung in [62] für klassische Dünngitter. Danach zeigen wir in Abschnitt 6.2 anhand einiger numerischer Ergebnisse, dass die Verwendung dieses Fehlerindikators zusammen mit einer geeigneten Verfeinerungsstrategie auch die effiziente Behandlung parabolischer Probleme ermöglicht, deren Lösung nicht die notwendigen Glattheitsvoraussetzungen aus Abschnitt 2.4 erfüllen. In 6.3 untersuchen wir die Möglichkeit der Fehlerschätzung bzgl. linearer Funktionale im Kontext der Raum-Zeit Dünngitterdiskretisierungen und präsentieren in diesem Rahmen in Abschnitt 6.4 die Ergebnisse einiger numerischer Experimente.

## 6.1 A posteriori Fehlerschätzer und Fehlerindikatoren

Wie in Abschnitt 2.4 bezeichne  $\{\psi_{\mathbf{j},\mathbf{i}}\} = \{\psi_{j_1,i_1}^\Omega \cdot \psi_{j_2,i_2}^T\}$  eine Basis des Raum-Zeit Dünngitters. Wir beschreiben ein Raum-Zeit Dünngitter über eine Menge  $G$  von Multiindizes der Basisfunktionen, welche das Dünngitter aufspannen. Zum Beispiel besitzt das Raum-Zeit Dünngitter  $\tilde{V}_l^0$  die Indexmenge

$$G = \{(\mathbf{j}, \mathbf{i}) \mid \|\mathbf{j}\|_1 \leq l, 1 \leq i_1 \leq \dim W_{j_1}^\Omega, 1 \leq i_2 \leq \dim W_{j_2}^T\}.$$

Weiterhin definieren wir für eine Basisfunktion  $\psi_{\mathbf{j},\mathbf{i}}$  die Menge  $\mathcal{H}(\psi_{\mathbf{j},\mathbf{i}})$  der hierarchischen Söhne

$$\begin{aligned} \mathcal{H}(\psi_{\mathbf{j},\mathbf{i}}) := \{ & \psi_{\bar{\mathbf{j}},\bar{\mathbf{i}}} \mid (\bar{j}_1, \bar{j}_2) = (j_1 + 1, j_2) \text{ oder } (\bar{j}_1, \bar{j}_2) = (j_1, j_2 + 1), \\ & \psi_{j_1,i_1}^T > \psi_{\bar{j}_1,\bar{i}_1}^T \text{ oder } \psi_{j_2,i_2}^\Omega > \psi_{\bar{j}_2,\bar{i}_2}^\Omega \}, \end{aligned} \quad (6.1)$$

wobei  $>$  die in Abschnitt 4.3 definierte hierarchische Relation ist. Betrachten wir für eine Funktion  $u$  die unendliche Darstellung

$$u = \sum_{\mathbf{j}} \sum_{\mathbf{i}} u_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}} \quad (6.2)$$

bezüglich der Basis  $\psi_{\mathbf{j},\mathbf{i}}$  in Raum-Zeit, so können wir diese unendliche Reihe abschneiden, indem wir für fest vorgegebenes  $\varepsilon > 0$  und eine gewählte Norm  $\|\cdot\|$  nur die Summanden  $u_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}}$  beibehalten, die in der Norm gemessen größer sind als die Schranke  $\varepsilon$ , vgl. [62]. Damit erhalten wir die Approximation

$$u_\varepsilon := \sum_{\|u_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}}\| \geq \varepsilon} u_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}} \quad (6.3)$$

mit dem Fehler  $u - u_\varepsilon = \sum_{\|u_{\mathbf{j},\mathbf{i}}\psi_{\mathbf{j},\mathbf{i}}\| < \varepsilon} u_{\mathbf{j},\mathbf{i}}\psi_{\mathbf{j},\mathbf{i}}$ . Somit bietet sich für eine fest gewählte Norm als Fehlerindikator die Größe

$$\gamma_{\mathbf{j},\mathbf{i}} \cdot |u_{\mathbf{j},\mathbf{i}}| := \|\psi_{\mathbf{j},\mathbf{i}}\| \cdot |u_{\mathbf{j},\mathbf{i}}| \quad (6.4)$$

an, wobei  $\gamma_{\mathbf{j},\mathbf{i}} = \|\psi_{\mathbf{j},\mathbf{i}}\|$  die von der gewählten Norm  $\|\cdot\|$  abhängige Gewichtung des hierarchischen Koeffizienten darstellt. Für gegebenes  $\varepsilon > 0$  ist jedoch weder die exakte Lösung  $u$  des Problems noch die Darstellung (6.3) von  $u_\varepsilon$  bekannt, sondern lediglich die approximierte Lösung  $u_l$  gegeben. Wir verwenden daher die Koeffizienten dieser approximierten Lösung in der Berechnung des Fehlerindikator.

Haben wir nun eine feste Schranke  $\varepsilon > 0$  vorgegeben, so bestimmen wir zunächst die Indexmenge  $G^+$  aller hierarchischen Söhne  $\psi_{\mathbf{j},\mathbf{i}}$  von Basisfunktionen des aktuellen Dünngitters, für die der Fehlerindikator größer als die Schranke  $\varepsilon$  ist, d.h.

$$G^+ := \{(\bar{\mathbf{j}}, \bar{\mathbf{i}}) \mid \exists(\mathbf{j}, \mathbf{i}) \in G \text{ so dass } \psi_{\bar{\mathbf{j}},\bar{\mathbf{i}}} \in \mathcal{H}(\psi_{\mathbf{j},\mathbf{i}}) \text{ und } \gamma_{\mathbf{j},\mathbf{i}} \cdot |u_{\mathbf{j},\mathbf{i}}| \geq \varepsilon\}. \quad (6.5)$$

Wir erhalten dann die neue Indexmenge des verfeinerten Dünngitters durch

$$G = G \cup G^+.$$

Man beachte, dass bei diesem Vorgehen der Ort und die Zeit gleichzeitig verfeinert werden. Entscheidenden Einfluss auf die Vergrößerung hat die Wahl des Parameters  $\varepsilon$ . Im Wesentlichen gibt es zwei Möglichkeiten, diesen Parameter zu bestimmen.

Man kann einerseits  $\varepsilon$  unabhängig von der aktuell berechneten diskreten Lösung wählen, wobei in diesem Fall  $\varepsilon$  im Sinne von (6.3) als ein Maß für die gewünschte Genauigkeit der Approximation interpretiert werden kann [5, 62]. Hierbei ist allerdings a priori nicht zu entscheiden, ob das durch die Wahl von  $\varepsilon$  vorherbestimmte adaptive Gitter tatsächlich in den zur Verfügung stehenden Speicher passt. Der Adaptivitätszyklus wird unter Umständen vor Erreichen der vorgegebenen Schranke abgebrochen und damit auch vor Erreichen des optimalen Gitters beendet. Ein solcher Abbruch führt häufig zu Gittern, die die Singularität mit der gegebenen Anzahl Unbekannter nur schlecht auflösen. Ist  $\varepsilon$  beispielsweise sehr klein, so erhalten wir bei frühzeitigem Abbruch nur ein sehr feines reguläres Raum-Zeit Dünngitter, das keinerlei adaptive Struktur aufweist, was für viele Probleme in der Regel nicht optimal ist.

In [98] wird deshalb vorgeschlagen, für fest vorgegebenes  $\varepsilon$  zunächst mit einem größeren Startwert  $\varepsilon^0 > \varepsilon$  zu beginnen. Es wird dann mit diesem  $\varepsilon^0$  der adaptive Zyklus solange durchgeführt, bis keine neue Funktion mehr in den Raum aufgenommen wird. Der Wert  $\varepsilon^0$  wird dann um einen konstanten Faktor  $c < 1$  verkleinert, und mit dem so entstandenen neuen Wert  $\varepsilon^1 := c \cdot \varepsilon^0$  wird der Verfeinerungsprozess erneut gestartet. Man erhält damit rekursiv eine Folge  $\varepsilon^k$  und der Prozess wird abgebrochen, sobald  $\varepsilon^l \leq \varepsilon$  gilt.

In [100] wird ein etwas anderes Verfahren vorgeschlagen, das auch wir im Folgenden verwenden, bei dem der Parameter  $\varepsilon$  aus der aktuell berechneten diskreten Lösung bestimmt wird. Hierzu definiert man

$$\varepsilon_{max} := \max_{\substack{(j,i) \in G, \\ \mathcal{H}(\psi_{j,i}) \not\subset G}} \gamma_{j,i} \cdot |u_{j,i}|, \quad (6.6)$$

d.h. wir wählen das Maximum der gewichteten Koeffizienten, die zu den Basisfunktionen gehören, deren hierarchische Söhne nicht bereits komplett in dem aktuellen Raum-Zeit Dünngitter enthalten sind. Für fest gewähltes  $c < 1$  erhalten wir dann

$$\varepsilon := c \cdot \varepsilon_{max}. \quad (6.7)$$

Das adaptive Verfahren wird abgebrochen, wenn entweder eine vorher fest vorgegebene maximale Schranke an Verfeinerungsschritten durchlaufen wurde, oder wenn eine vorher festgelegte Zahl an Unbekannten überschritten ist. Dieses Vorgehen ist analog zu dem Algorithmus der Gitterverfeinerung bei Finiten Elementen, bei dem nur solche Elemente verfeinert werden, auf denen der Fehlerindikator oder Fehlerschätzer größer als das skalierte Maximum über alle Fehlerschätzer ist, siehe [3, 95, 114]. Insgesamt erhalten wir damit den Verfeinerungsalgorithmus 6.1.

ALGORITHMUS 6.1 (Verfeinerungsalgorithmus).

1. Gegeben: Schranke  $N$  der maximalen Anzahl an Unbekannten bzw. maximale Anzahl der Verfeinerungsschritte  $k_{max}$ , Faktor  $c < 1$ , Ausgangsgitter  $G_0$ .
2. Setze  $k = 0$ .
3. Solange  $|G| < N$  und  $k < k_{max}$ :
  - (a) Berechne diskrete Lösung  $u_k$  des parabolischen Problems bzgl. des durch  $G$  beschriebenen Raum-Zeit Dünngitters.
  - (b) Berechne  $\varepsilon_{max}$  aus  $u_k$  gemäß (6.6).
  - (c) Setze  $\varepsilon := c \cdot \varepsilon_{max}$ .
  - (d) Berechne  $G^+$  gemäß (6.5).
  - (e) Setze  $k = k + 1$  und  $G = G \cup G^+$ .

Bisher haben wir die frei wählbare Norm  $\|\cdot\|$  noch nicht näher spezifiziert. Bei den Untersuchungen für klassische Dünngitter im Rahmen elliptischer Probleme werden dabei üblicherweise die  $L^2$ -, die  $L^\infty$ -, die  $L^1$  oder die  $H^1$ -Norm verwendet. Für parabolische Probleme sind allerdings auch andere Normen denkbar, wobei die Orts- und Zeitableitungen jeweils unterschiedlich gewichtet werden, wie dies etwa bei den bereits betrachteten  $H_{mix}^{1,0}$ - und  $H_{mix}^{(1,2),(0,\infty)}$ -Normen der Fall war. Da bei unseren

**Tabelle 6.1.** Fehler für die Crank-Nicolson Diskretisierung auf regulären Dünngittern verschiedener Level  $l$  für das Problem (6.8) mit Lösung  $u(x, t) = (x^2 + t^2)^{1/3}$ .

l	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
4	3.562 <sub>-2</sub>	-	3.760 <sub>-3</sub>	-	7.068 <sub>-2</sub>	-
5	2.816 <sub>-2</sub>	0.79	2.626 <sub>-3</sub>	0.70	7.096 <sub>-2</sub>	1.00
6	2.244 <sub>-2</sub>	0.80	1.683 <sub>-3</sub>	0.64	6.994 <sub>-2</sub>	0.99
7	1.797 <sub>-2</sub>	0.80	1.024 <sub>-3</sub>	0.61	6.747 <sub>-2</sub>	0.96
8	1.432 <sub>-2</sub>	0.80	6.048 <sub>-4</sub>	0.59	6.391 <sub>-2</sub>	0.95
9	1.152 <sub>-2</sub>	0.80	3.505 <sub>-4</sub>	0.58	5.967 <sub>-2</sub>	0.93
10	9.089 <sub>-3</sub>	0.79	2.008 <sub>-4</sub>	0.57	5.510 <sub>-2</sub>	0.92
11	7.258 <sub>-3</sub>	0.80	1.142 <sub>-4</sub>	0.57	5.046 <sub>-2</sub>	0.92
12	5.727 <sub>-3</sub>	0.79	6.472 <sub>-5</sub>	0.57	4.591 <sub>-2</sub>	0.91

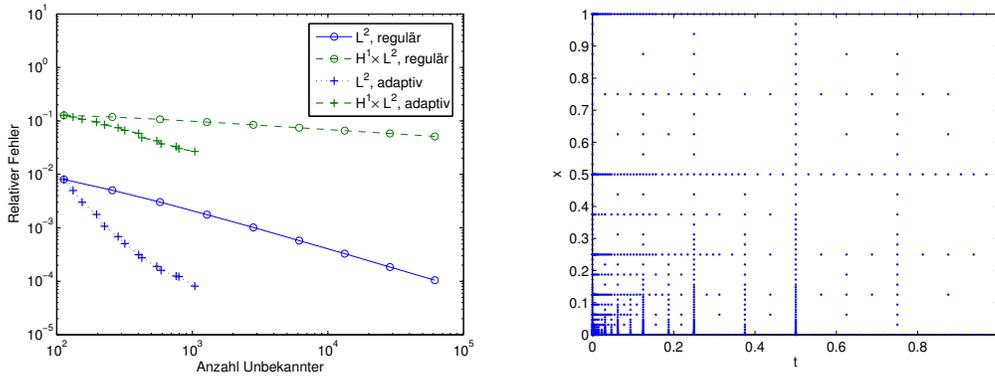
Untersuchungen die  $L^\infty$ -Norm zufrieden stellende Resultate geliefert hat, wollen wir uns in den folgenden Untersuchungen auf die Verwendung dieser Norm beschränken. Wir erhalten damit  $\gamma_{j,i} = 1.0$  als Gewichtung der hierarchischen Koeffizienten.

## 6.2 Numerische Ergebnisse

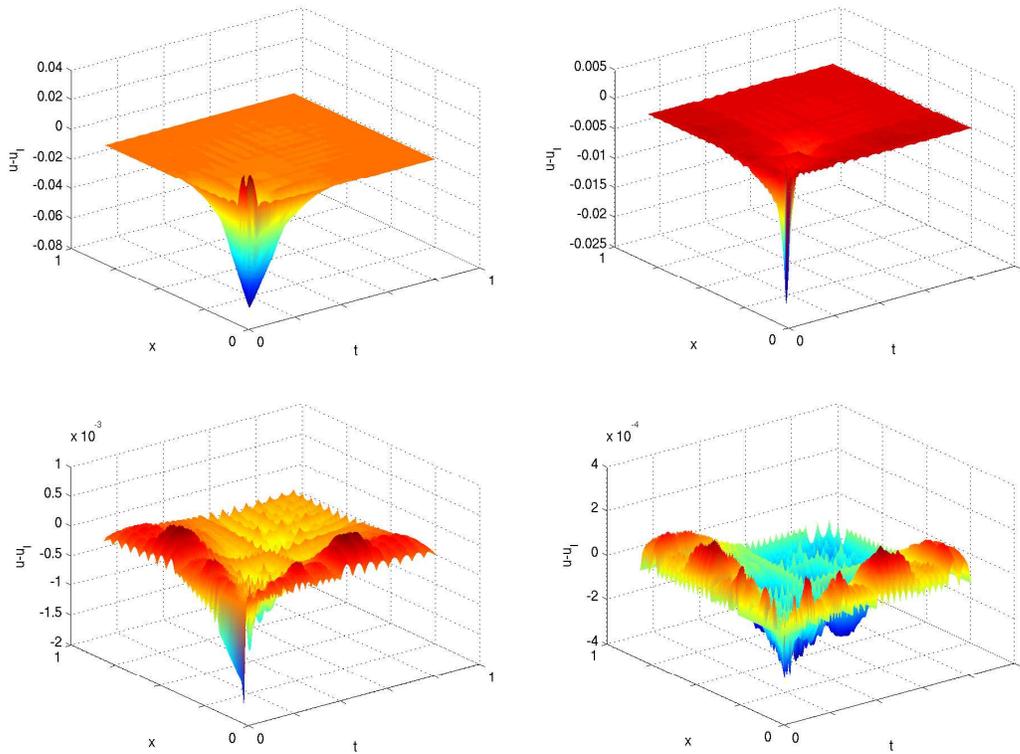
Anhand einiger Beispiele untersuchen wir nun den in dem vorherigen Abschnitt diskutierten Verfeinerungsalgorithmus. Dabei beschränken wir uns bei unseren Untersuchungen auf das Crank-Nicolson Verfahren. Numerische Experimente haben gezeigt, dass wir für das Discontinuous-Galerkin Verfahren ähnliche Ergebnisse erhalten, auf deren Präsentation wir aus Platzgründen deshalb verzichten. Zu beachten ist, dass wir zur Fehlermessung diesmal nicht die Approximation  $\hat{e}_l := u_l - u_h$  des Fehlers  $e_l := u_l - u$  wie in Abschnitt 4.4 heranziehen können, da die betrachteten Funktionen  $u$  in der Regel nicht mehr die notwendigen Glattheitsanforderungen erfüllen und somit der Interpolant  $u_h$  auf dem zu Level  $l$  gehörenden vollen Gitter die exakte Lösung  $u$  nicht mehr hinreichend genug approximiert. Stattdessen wählen wir ein sehr feines volles Gitter fest und verwenden den Interpolanten auf diesem vollen Gitter. Hierbei ist  $\varepsilon$  so gewählt, dass die in dem resultierenden adaptiven Raum-Zeit Dünngitter enthaltenen feinsten Basisfunktionen maximal zu einem Level gehören, das mindestens noch zwei Level über der Leveltiefe des zur Fehlermessung verwendeten vollen Gitters ist. Zur Bestimmung des Parameters  $\varepsilon := c \cdot \varepsilon_{max}$  verwenden wir in den folgenden Untersuchungen den Faktor  $c = 0.5$ .

*Beispiel 6.1.* Wir betrachten das Problem

$$\begin{aligned}
 \partial_t u - \partial_{xx} u &= f \text{ in } \Omega = (0, 1), \\
 u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega, \\
 u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega,
 \end{aligned}
 \tag{6.8}$$



**Abbildung 6.1.** Fehler der adaptiven und regulären Crank-Nicolson Diskretisierung des Problems (6.8) gegen Freiheitsgrade (links) und das nach 14 Verfeinerungsschritten resultierende Raum-Zeit Dünngitter (rechts).



**Abbildung 6.2.** Startfehler (oben links) und Fehler nach 5 (oben rechts), 10 (unten links) und 15 Verfeinerungsschritten der Crank-Nicolson Diskretisierung des Problems (6.8) ausgehend von Level  $l = 4$ .

**Tabelle 6.2.** Fehler für die Crank-Nicolson Diskretisierung auf adaptiven Dünngittern für das Problem (6.8) mit Lösung  $u(x, t) = (x^2 + t^2)^{1/3}$ .

dof	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
113	3.562 <sub>-2</sub>	-	3.760 <sub>-3</sub>	-	7.066 <sub>-2</sub>	-
166	2.244 <sub>-2</sub>	0.63	1.654 <sub>-3</sub>	0.44	7.011 <sub>-2</sub>	0.99
257	1.432 <sub>-2</sub>	0.64	6.019 <sub>-4</sub>	0.36	6.411 <sub>-2</sub>	0.91
367	9.023 <sub>-3</sub>	0.63	2.126 <sub>-4</sub>	0.35	5.534 <sub>-2</sub>	0.86
535	5.684 <sub>-3</sub>	0.63	8.291 <sub>-5</sub>	0.39	4.614 <sub>-2</sub>	0.83
749	3.581 <sub>-3</sub>	0.63	3.859 <sub>-5</sub>	0.47	3.784 <sub>-2</sub>	0.82
969	2.256 <sub>-3</sub>	0.63	2.360 <sub>-5</sub>	0.61	2.235 <sub>-2</sub>	0.59
1279	1.421 <sub>-3</sub>	0.63	1.531 <sub>-5</sub>	0.65	9.694 <sub>-3</sub>	0.43
1675	8.952 <sub>-4</sub>	0.63	8.883 <sub>-6</sub>	0.58	5.613 <sub>-3</sub>	0.58

wobei die rechte Seite, die Rand- und Anfangsbedingungen derart gewählt sind, dass wir als Lösung des Problems die Funktion

$$u(x, t) = (x^2 + t^2)^{1/3}$$

erhalten. Für diese Funktion gilt zwar  $u \in H^{5/3-\varepsilon}(\Omega_T)$ , siehe z.B. [80], für beliebig kleine  $\varepsilon > 0$ , die für die in Satz 2.12 angegebenen Interpolationsfehlerabschätzungen notwendigen Glattheitsannahmen sind jedoch nicht erfüllt. Da die numerischen Experimente des Abschnitts 4.4 gezeigt haben, dass sich die Diskretisierungsfehler wie die Interpolationsfehler verhalten, erwarten wir bei diesem Beispiel eine Verschlechterung der Konvergenzraten verglichen mit den Ergebnissen des Abschnitts 4.4. Tabelle 6.1 enthält die Ergebnisse für die Crank-Nicolson Diskretisierung auf verschiedenen Leveln  $l$ . In der Tat sind die Ergebnisse schlechter als in den Beispielen in Abschnitt 4.4. Für den Fehler gemessen in der  $H_{mix}^{(1,2),(0,2)}$ -Norm beobachten wir eine sehr geringe Reduktion. So liefert das Crank-Nicolson Verfahren von Level  $l = 4$  zu Level  $l = 5$  nahezu keine Reduktion des Fehler und von Level  $l = 11$  auf Level  $l = 12$  wird der Fehler lediglich um 0.91 reduziert. Wir erinnern uns daran, dass wir für genügend glatte Funktionen in den früheren numerischen Experimenten für diese Norm Raten um die 0.5 erhalten haben. Auch für den Fehler gemessen in der  $L^2$ - und der  $L^\infty$ -Norm führt das Crank-Nicolson Verfahren zu schlechteren Raten als in den vorherigen Beispielen. So erhalten wir beispielsweise eine Fehlerreduktion von Level  $l = 11$  auf Level  $l = 12$  von  $\rho_{L^2} = 0.57$ .

Betrachten wir nun die in Tabelle 6.2 dargestellten Ergebnisse des von Level  $l = 4$  ausgehenden adaptiven Verfahrens. Hierbei sind aus Platzgründen die Ergebnisse nur für jeden zweiten Verfeinerungsschritt angegeben. Es fällt zunächst auf, dass der Fehler ebenfalls recht langsam abzunehmen scheint. Allerdings handelt es sich hier ja nicht um reguläre Diskretisierungen bzgl. unterschiedlicher Level  $l$ , sondern um Diskretisierungen auf unterschiedlichen adaptiven Gitter, deren Freiheitsgrade bedeutend langsamer zunehmen als im regulären Fall. So besitzt der nach zwei Verfeinerungsschritten resultierende Raum-Zeit Dünngitterraum 166 Freiheitsgrade und

**Tabelle 6.3.** Fehler für die Crank-Nicolson Diskretisierung auf regulären Dünngittern verschiedener Level  $l$  für das Problem (6.8) mit Lösung  $u(x, t) = \exp(-(x - 0.5)^2/(t + 0.05))$ .

$l$	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^\infty}$	$\rho_{H^1 \otimes L^\infty}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
3	4.969 <sub>-1</sub>	-	1.438 <sub>-1</sub>	-	1.034 <sub>0</sub>	-	5.213 <sub>-1</sub>	-
4	2.721 <sub>-1</sub>	0.55	6.455 <sub>-2</sub>	0.45	6.357 <sub>-1</sub>	0.61	2.962 <sub>-1</sub>	0.57
5	9.796 <sub>-2</sub>	0.36	2.342 <sub>-2</sub>	0.36	3.137 <sub>-1</sub>	0.49	1.815 <sub>-1</sub>	0.61
6	3.820 <sub>-2</sub>	0.39	7.934 <sub>-3</sub>	0.34	1.991 <sub>-1</sub>	0.63	1.183 <sub>-1</sub>	0.65
7	1.402 <sub>-2</sub>	0.37	2.549 <sub>-3</sub>	0.32	1.233 <sub>-1</sub>	0.62	7.172 <sub>-2</sub>	0.61
8	4.649 <sub>-3</sub>	0.33	7.867 <sub>-4</sub>	0.31	7.238 <sub>-2</sub>	0.59	4.156 <sub>-2</sub>	0.58
9	5.826 <sub>-3</sub>	1.25	2.661 <sub>-4</sub>	0.34	5.258 <sub>-2</sub>	0.73	2.307 <sub>-2</sub>	0.56

nach zwei weiteren adaptiven Verfeinerungen ist die Zahl der Freiheitsgrade lediglich um den Faktor 1.6 auf 257 Unbekannte gestiegen. Erinnern wir uns daran, dass bei regulären dünnen Gittern  $\tilde{V}_l^0$  im eindimensionalen Fall der Anstieg einen (bis auf einen logarithmischen Faktor) Faktor zwei beträgt. Damit erhalten wir insgesamt ein wesentlich besseres Kosten-Nutzen Verhältnis als bei der regulären Verfeinerung, d.h. wir erreichen mit wesentlich weniger Freiheitsgraden einen kleineren oder vergleichbar großen Fehler wie bei den regulären Raum-Zeit Dünngitter. So liefert das adaptive Crank-Nicolson Verfahren mit 1675 Freiheitsgraden einen  $L^2$ -Fehler von  $8.88_{-6}$  und einen  $H_{mix}^{(1,2),(0,2)}$ -Fehler von  $5.613_{-3}$ , während wir im regulären Fall mit 61441 (Level  $l = 12$ ) einen  $L^2$ -Fehler von  $6.47_{-5}$  und einen  $H_{mix}^{(1,2),(0,2)}$ -Fehler von  $4.60_{-2}$  erhalten.

Besonders deutlich ist dieser Unterschied zwischen dem adaptiven und dem regulären Verfahren in Abbildung 6.1 zu sehen, in dem der  $L^2$ - und der  $H^{(1,2),(0,2)}$ -Fehler in abhängig von den Freiheitsgraden geplottet ist. Zusätzlich ist dort das nach acht adaptiven Verfeinerungsschritten resultierende Raum-Zeit Dünngitter zu sehen. Man sieht deutlich, wie besonders zu dem Punkt  $(0, 0)$ , in dem die Singularität liegt, Freiheitsgrade eingefügt wurden. Dadurch wird erreicht, dass der Fehler in dem Gebiet nahezu gleichmäßig verteilt ist. Dieser Effekt kann in Abbildung 6.1 beobachtet werden. Dort ist der Startfehler und der Fehler nach 5, 10 und 15 Verfeinerungsschritten zu sehen.

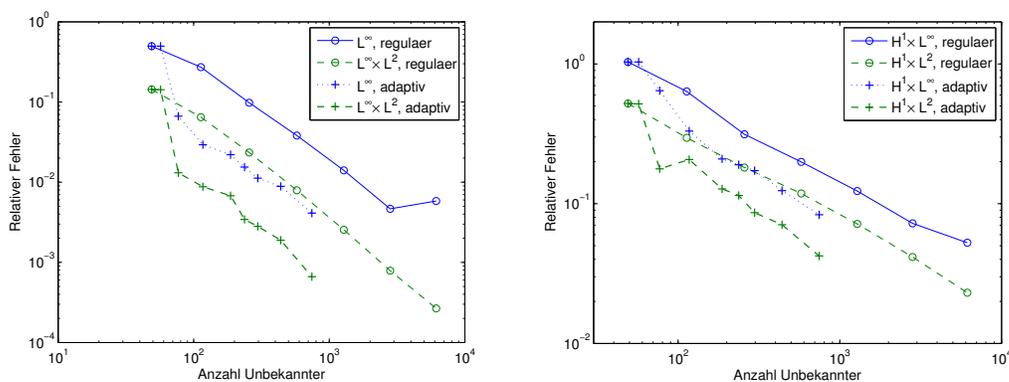
*Beispiel 6.2.* Wir betrachten erneut das Problem (6.8) des vorigen Beispiels, wobei wir jedoch die rechte Seite, die Anfangs- und die Randbedingung derart wählen, das wir als Lösung

$$u(x, t) = \exp(-(x - 0.5)^2/(t + 0.05))$$

erhalten. Diese Funktion ist zwar unendlich oft differenzierbar, besitzt aber zum Zeitpunkt  $t = 0$  einen sehr stark ausgeprägten Peak in  $x = 0.5$ , der bereits nach kurzer Zeit stark abflacht. Wir erwarten deshalb asymptotisch, d.h. für großes  $l$ , die gleichen Fehlerreduktionsraten wie in Abschnitt 4.4. Allerdings könnte, bedingt durch

**Tabelle 6.4.** Fehler für die Crank-Nicolson Diskretisierung auf adaptiven Dünngittern ausgehend von einem regulären Dünngitter auf Level  $l = 3$  für das Problem (6.8) mit Lösung  $u(x, t) = \exp(-(x - 0.5)^2/(t + 0.05))$ .

dof	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^\infty}$	$\rho_{H^1 \otimes L^\infty}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
49	4.969 <sub>-1</sub>	-	1.438 <sub>-1</sub>	-	1.034 <sub>0</sub>	-	5.213 <sub>-1</sub>	-
77	6.650 <sub>-2</sub>	0.13	1.307 <sub>-2</sub>	0.09	6.447 <sub>-1</sub>	0.62	1.776 <sub>-1</sub>	0.34
117	2.932 <sub>-2</sub>	0.44	8.761 <sub>-3</sub>	0.67	3.309 <sub>-1</sub>	0.51	2.070 <sub>-1</sub>	1.17
187	2.206 <sub>-2</sub>	0.75	6.779 <sub>-3</sub>	0.77	2.092 <sub>-1</sub>	0.63	1.276 <sub>-1</sub>	0.62
237	1.545 <sub>-2</sub>	0.70	3.431 <sub>-3</sub>	0.51	1.902 <sub>-1</sub>	0.91	1.146 <sub>-1</sub>	0.90
297	1.121 <sub>-2</sub>	0.73	2.809 <sub>-3</sub>	0.82	1.726 <sub>-1</sub>	0.91	8.633 <sub>-2</sub>	0.75
439	8.871 <sub>-3</sub>	0.79	1.887 <sub>-3</sub>	0.67	1.239 <sub>-1</sub>	0.72	7.080 <sub>-2</sub>	0.82
745	4.116 <sub>-3</sub>	0.46	6.611 <sub>-4</sub>	0.35	8.334 <sub>-2</sub>	0.67	4.228 <sub>-2</sub>	0.60

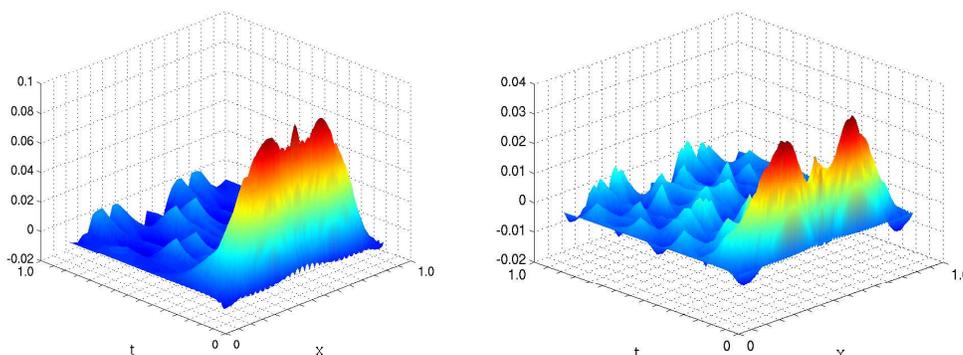


**Abbildung 6.3.** Fehler der adaptiven und regulären Crank-Nicolson Diskretisierung des Problems (6.8) bzgl. der  $L^2$ - und der  $L^\infty$ -Norm (links) und bzgl. der  $H_{mix}^{1,0}$ - und  $H_{mix}^{(1,2),(0,\infty)}$ -Norm (rechts) in Abhängigkeit von der Anzahl benötigter Freiheitsgrade.

den starken Peak, die Fehlerreduktion für kleine Level, d.h. in der Präasymptotik, durchaus schlechter sein.

In Tabelle 6.3 sind die Ergebnisse der Crank-Nicolson Diskretisierung auf verschiedenen Leveln aufgeführt. In der Tat ist die Fehlerreduktion für kleinere Levelzahlen geringer als bei den früheren Experimenten in Abschnitt 4.4. Im Fall der  $L^\infty$ -Norm beobachten wir sogar einen leichten Anstieg des Fehlers von Level  $l = 8$  zu Level  $l = 9$  von  $4.649_{-3}$  auf  $5.826_{-3}$ . In Abbildung 6.4 sind die Fehler der Diskretisierung für  $l = 5$  und  $l = 6$  dargestellt und man erkennt deutlich, dass der Fehler ungleich über den Raum-Zeit Zylinder verteilt ist und zu sehr frühen Zeitpunkten dominiert.

Wir betrachten nun die in Tabelle 6.4 dargestellten Ergebnisse des adaptiven Verfahrens, wobei wir als Ausgangslevel für die adaptive Verfeinerung das Level  $l = 4$



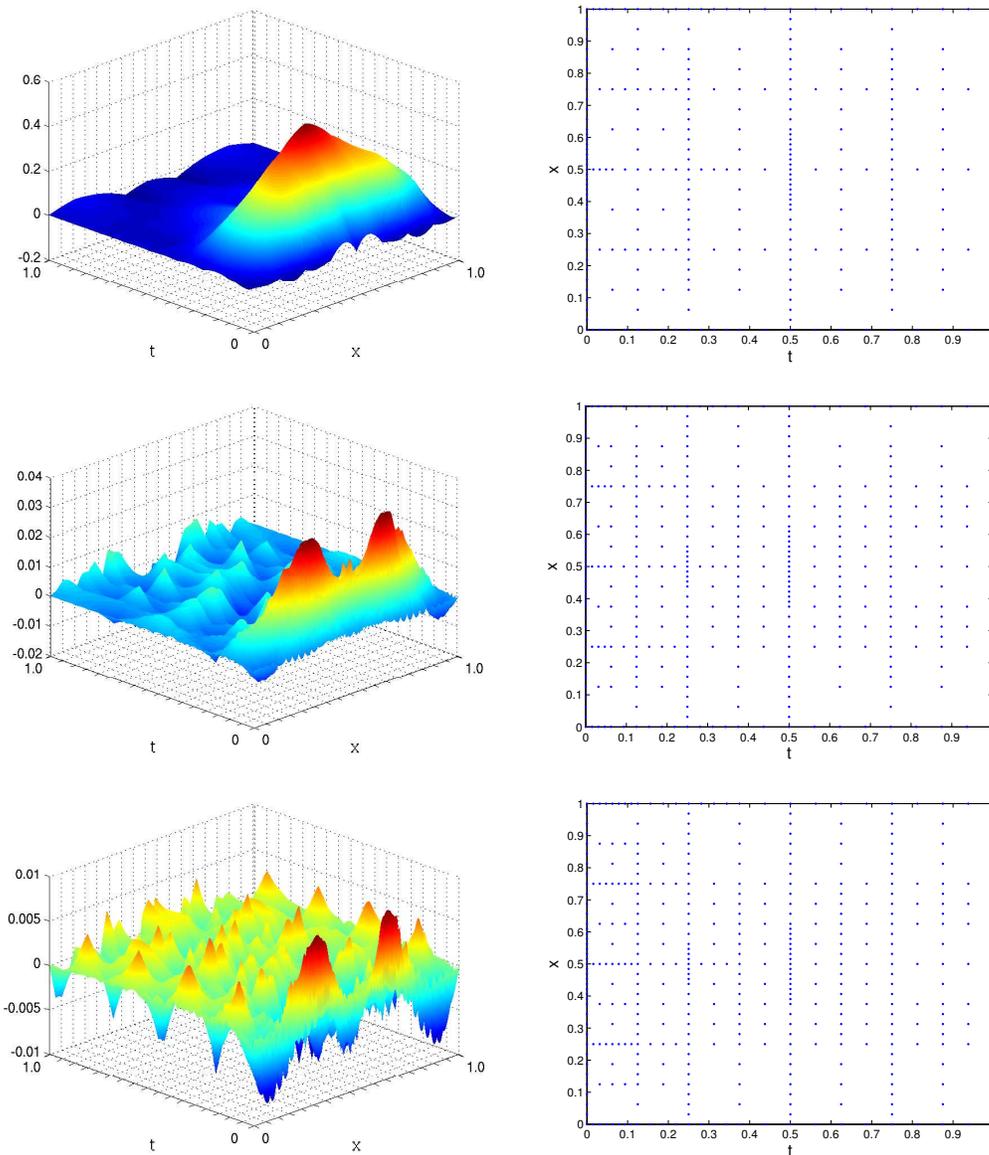
**Abbildung 6.4.** Fehler der Crank-Nicolson Diskretisierung des Problems (6.8) mit Lösung  $u(x, t) = \exp(-(x - 0.5)^2/(t + 0.05))$  auf Level  $l = 5$  (links) und  $l = 6$  (rechts).

**Tabelle 6.5.** Konvergenzresultate des Crank-Nicolson Verfahrens bei regulärer Verfeinerung für das Problem (6.9).

l	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^\infty}$	$\rho_{H^1 \otimes L^\infty}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
3	$3.207_{-2}$	-	$9.318_{-4}$	-	$5.323_{-2}$	-	$3.489_{-2}$	-
4	$2.019_{-2}$	0.63	$2.927_{-4}$	0.31	$3.378_{-2}$	0.63	$2.215_{-2}$	0.63
5	$1.272_{-2}$	0.63	$9.184_{-5}$	0.31	$2.124_{-2}$	0.63	$1.393_{-2}$	0.63
6	$7.949_{-3}$	0.63	$2.879_{-5}$	0.31	$1.314_{-2}$	0.62	$8.618_{-3}$	0.62
7	$5.007_{-3}$	0.63	$8.936_{-6}$	0.31	$7.822_{-3}$	0.60	$5.132_{-3}$	0.60

gewählt haben. Wir erhalten ein wesentlich besseres Kosten-Nutzen Verhältnis als bei der regulären Verfeinerung. So beträgt der  $L^\infty$ -Fehler im adaptiven Fall mit nur 745 Freiheitsgraden  $4.116_{-3}$ , während wir  $e_{L^\infty} = 5.826_{-3}$  auf dem regulären Dünngitter  $\tilde{V}_9^0$ , das 6145 Freiheitsgrade besitzt, erzielen. Dieser Unterschied wird besonders in Abbildung 6.3 deutlich, in der die Fehler der regulären und adaptiven Diskretisierungen in Abhängigkeit von der Zahl der Freiheitsgrade dargestellt sind. Abbildung 6.5 zeigt die bei einem von Level  $l = 3$  ausgehenden Verfeinerungszyklus entstandenen adaptiven Gitter und die dazugehörigen Fehler. Man beobachtet dort, dass zum einen feinere Zeitschritte zum Anfangszeitpunkt hin eingefügt werden, gleichzeitig aber auch der Ort an der Stelle des Peaks feiner aufgelöst wird. Dadurch verringert sich die Dominanz des Fehlers an frühen Zeitpunkten und nach wenigen Verfeinerungsschritten ist der Fehler nahezu gleichmäßig über das gesamte Gebiet verteilt.

*Beispiel 6.3.* Als ein Beispiel für Singularitäten, die durch die Geometrie erzeugt



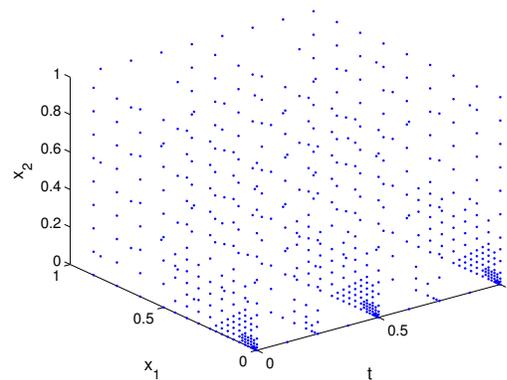
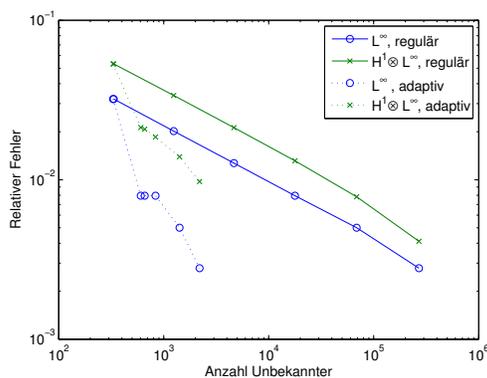
**Abbildung 6.5.** Fehler (links) und adaptives Raum-Zeit Dünngitter (rechts) für die Crank-Nicolson Diskretisierung des Problems (6.8) nach zwei (oben), vier (mitte) und sechs (unten) Verfeinerungsschritten ausgehend von Level  $l = 3$ .

werden, betrachten wir das Problem

$$\begin{aligned}
 \partial_t u - \Delta u &= f \text{ für alle } x \in \Omega = (0, 1)^2, \\
 u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega, \\
 u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega.
 \end{aligned}
 \tag{6.9}$$

**Tabelle 6.6.** Konvergenzergebnisse des Crank-Nicolson Verfahrens bei adaptiver Verfeinerung ausgehend von dem regulären Raum-Zeit Dünngitter auf Level  $l = 3$  für das Problem (6.9).

dof	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^\infty}$	$\rho_{H^1 \otimes L^\infty}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
329	$3.207_{-2}$	-	$9.318_{-4}$	-	$5.323_{-2}$	-	$3.489_{-2}$	-
605	$7.949_{-3}$	0.25	$1.770_{-4}$	0.19	$2.127_{-2}$	0.40	$1.397_{-2}$	0.40
659	$7.949_{-3}$	1.00	$1.659_{-4}$	0.94	$2.075_{-2}$	0.98	$1.363_{-2}$	0.98
835	$7.949_{-3}$	1.00	$1.388_{-4}$	0.84	$1.855_{-2}$	0.89	$1.217_{-2}$	0.89
1415	$5.007_{-3}$	0.63	$6.934_{-5}$	0.50	$1.393_{-2}$	0.75	$9.158_{-3}$	0.75
2194	$2.792_{-3}$	0.56	$4.226_{-5}$	0.61	$9.742_{-3}$	0.70	$6.409_{-3}$	0.70



**Abbildung 6.6.** Adaptives Raum-Zeit Dünngitter entstanden aus der adaptiven Verfeinerung von einem regulären Raum-Zeit Dünngitter auf Level  $l = 3$  für das Problem (6.9) mit Lösung  $u(r, \phi, t) = \exp(-t) \cdot r^{2/3} \sin(\frac{2\phi - \pi}{3})$ .

Wir wählen dabei die rechte Seite und die Randbedingungen so, dass wir die exakte Lösung  $u$  in Polarkoordinaten  $(r, \phi)$

$$u(r, \phi, t) = \exp(-t) \cdot r^{2/3} \sin\left(\frac{2\phi - \pi}{3}\right)$$

erhalten. Derartige Funktionen entstehen zum Beispiel bei elliptischen Problemen durch eine einspringende Ecke. Für jedes feste  $t$  ist die Funktion lediglich in  $H^s$  mit  $s < 5/3$  und erfüllt daher nicht die Glattheitsannahmen, die wir zur Abschätzung der Dünngitter Interpolationsfehler benötigen. Im Gegensatz zu dem vorherigen Beispiel erwarten wir daher im Allgemeinen nicht die gleichen asymptotischen Fehlerreduktionsraten, die bereits in Abschnitt 4.4 gemessen wurden. Tabelle 6.5 zeigt die Ergebnisse der Crank-Nicolson Diskretisierung für verschiedene Level. Die Fehlerreduktionsrate des  $L^\infty$ -Fehlers ist dabei mit 0.63 unabhängig vom Level besonders schlecht. Es sei daran erinnert, dass wir bei genügender Glattheit für die Crank-

Nicolson Diskretisierung Reduktionsraten gemessen haben, die sich mit wachsendem Level dem Wert 0.25 nähern.

Ausgehend von dem regulären Raum-Zeit Dünngitter  $\tilde{V}_3^0$  sind in Tabelle 6.6 die Ergebnisse des adaptiven Verfahrens ausgehend. Besonders interessant ist der Vergleich der Kosten-Nutzen Relationen in Abbildung 6.6, in der der Fehler in Abhängigkeit von der Anzahl der Freiheitsgrade für verschiedene reguläre und adaptive Raum-Zeit Dünngitter bzgl. der  $L^\infty$ - und der  $H_{mix}^{(1,2),(0,\infty)}$ -Norm dargestellt ist. Es ist deutlich erkennbar, dass das adaptive Verfahren mit wesentlich weniger Freiheitsgraden zu kleineren Fehlern als die reguläre Verfeinerung führt. In Abbildung 6.6 ist zusätzlich das nach drei Verfeinerungsschritten resultierende adaptive Raum-Zeit Dünngitter dargestellt. Wir beobachten hier, dass die Singularität im Punkte  $x = (0, 0)$  zu allen Zeitpunkten erkannt und aufgelöst wird.

### 6.3 Adaptivität für lineare Funktionale

In vielen Anwendungen ist man primär nicht an der Lösung  $u$ , sondern vielmehr an dem Wert eines bestimmten Funktionals  $l(\cdot)$  ausgewertet an der Stelle  $u$  interessiert. Ein Beispiel sind hier etwa Optimierungsprobleme, bei denen die parabolische Differentialgleichung als Nebenbedingung auftritt. Eine Diskretisierung für diese Probleme sollte demnach die Größe  $|l(u) - l(u_l)|$  berücksichtigen. Diese Aufgabenstellung sollte selbstverständlich auch bei der adaptiven Behandlung solcher Probleme nicht außer Acht gelassen werden. Insbesondere der Fehlerschätzer bzw. der Fehlerindikator muss dabei angepasst werden, so dass Fehler nicht mehr bzgl. einer Norm, sondern vielmehr bzgl. des Funktionals betrachtet werden.

Die Fehlerschätzung bzgl. linearer Funktionale hat daher in den vergangenen Jahren zunehmend an Bedeutung gewonnen und ist Gegenstand aktueller Forschung. Hierbei hat sich im Fall der Finiten Elemente vor allem der Ansatz der dualitätsbasierten Fehlerschätzer als besonders viel versprechend erwiesen [11, 13, 59], bei dem die Lösung eines dualen Problems zur Fehlerschätzung mit einbezogen wird.

In Anlehnung an diese dualitätsbasierten Fehlerschätzer für Finite Elemente ist in [32, 100] ein Fehlerindikator für die klassische Dünngitterdiskretisierung elliptischer Probleme zur Abschätzung des Fehlers bzgl. linearer Funktionale vorgeschlagen worden. Hierbei werden die hierarchischen Koeffizienten mit den hierarchischen Koeffizienten der Lösung eines dualen Problems gewichtet.

Wir möchten nun in diesem Abschnitt die in [32, 100] vorgeschlagenen Heuristiken auf die Raum-Zeit Dünngitterdiskretisierung übertragen. Hierbei beschränken wir uns auf den Fall der Crank-Nicolson Diskretisierung, deren Behandlung etwas schwieriger ist als die Untersuchung für das Discontinuous-Galerkin Verfahren, da es sich um ein Petrov-Galerkin Verfahren handelt. Insbesondere lassen sich die folgenden Überlegungen leicht auf den Discontinuous-Galerkin Fall übertragen. Des Weiteren behandeln wir zur Vereinfachung der Darstellung nur den Fall homoge-

ner Anfangsbedingungen  $u_0 = 0$  und homogener Dirichlet Randbedingungen. Für  $l_0 \in L^2(\Omega_T)$  betrachten wir das durch

$$l(\phi) := (l_0, \phi)_{L^2(\Omega_T)}$$

definierte Funktional  $l(\cdot)$  auf  $L^2(\Omega_T)$ . Sei  $u$  die Lösung des parabolischen Problems in schwacher Form (3.8). Wir bezeichnen mit  $u_l \in V_l^{0,Ansatz}$  die Lösung des mit dem Crank-Nicolson Verfahren aus Abschnitt (4.1) diskretisierten parabolischen Problems, d.h. nach (4.4) gilt

$$B(u_l, \varphi) = \int_0^T (f, \varphi)_0 \text{ für alle } \varphi \in V_l^{0,Test},$$

wobei  $V_l^{0,Ansatz}$  der aus der linearen hierarchischen Basis in der Zeit und der ( $d$ -) linearen hierarchischen Basis im Ort und  $V_l^{0,Test}$  der aus der stückweise konstanten hierarchischen Basis in der Zeit und der ( $d$ -) linearen hierarchischen Basis im Ort konstruierte Dünngitterraum ist. Bei dem zu (3.8) adjungierten Problem suchen wir nun ein  $z$ , das

$$\begin{aligned} -(\partial_t z, \varphi)_{L^2(\Omega)} + a(t; \varphi, z) &= l(\varphi) \text{ für alle } \varphi \in H_0^1(\Omega), t \in (0, T] \\ z(T, \cdot) &= 0, \text{ f.ü. in } \Omega, \\ z(x, t) &= 0, \text{ für alle } t \in (0, T] \text{ und } x \in \partial\Omega. \end{aligned} \quad (6.10)$$

erfüllt. Nach weiterer partieller Integration des Terms mit der Zeitableitung (hierbei nutzen wir aus, dass  $\varphi(x, 0) = 0$  für alle  $\varphi \in V_l^{0,Ansatz}$  gilt) erhalten wir das diskretisierte Problem: Finde  $z_l \in V_l^{0,Test}$  so dass

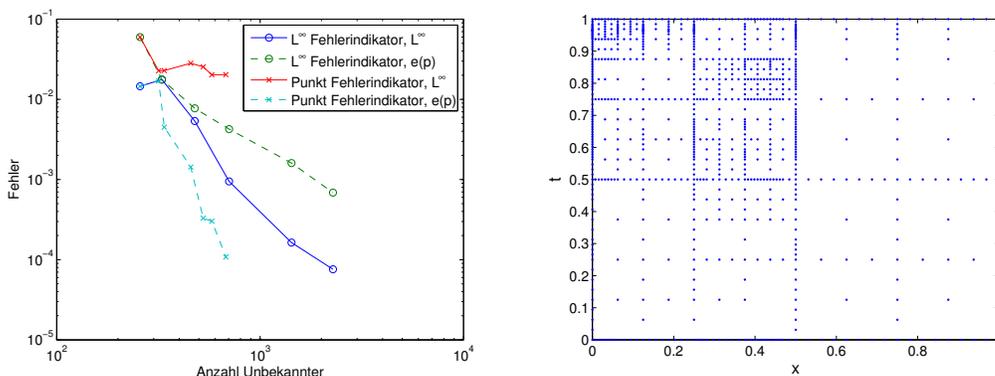
$$B(\varphi, z_l) = \int_0^T (l_0, \varphi)_0 \text{ für alle } \varphi \in V_l^{0,Ansatz} \quad (6.11)$$

gilt und  $z_l$  die Endzeitnullrandbedingung sowie die homogene Dirichletrandbedingung erfüllt. Man beachte, dass wir zur Diskretisierung des adjungierten Problems die Wahl des Ansatz- und Testraumes vertauscht haben, d.h. der Ansatzraum für das Ursprungsproblem ist nun der Testraum für das adjungierte Problem und umgekehrt ist der Testraum für das ursprüngliche Problem nun der Ansatzraum des adjungierten Problems. Mit Hilfe dieser Rollenvertauschung erhalten wir  $B(u - u_l, z_l) = 0$ , was wir nun ausnutzen wollen. Offensichtlich gilt für die Lösung  $z$  des adjungierten Problems (6.10)  $B(\varphi, z) = \int_0^T (l_0, \varphi)_0$  für alle  $\varphi \in H_0^1(\Omega) \otimes L^2((0, T))$ ,  $\varphi(x, 0) = 0$ , und damit insbesondere für alle  $\varphi \in V_l^{0,Ansatz}$ . Wir erhalten für den Fehler  $e_l := u - u_l$

$$l(e_l) = l(u - u_l) = B(u - u_l, z)$$

und mit  $B(u - u_l, z_l) = 0$  folgt

$$l(e_l) = B(u - u_l, z - z_l). \quad (6.12)$$



**Abbildung 6.7.** Links: Ergebnisse der adaptiven Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung des Beispiels 6.4 bei Verwendung des Absolutwertes der hierarchischen Koeffizienten als Fehlerindikator und des dualen Fehlerindikators (6.14). Rechts: Das ausgehend von dem regulären Raum-Zeit Dünngitter  $\tilde{V}_5^0$  nach sechs Verfeinerungsschritten mit Hilfe des dualen Fehlerindikators resultierende Raum-Zeit Dünngitter.

In [32] wird nun für klassische Dünngitter im Fall von elliptischen Problemen vorgeschlagen, die Bilinearform  $B(\cdot, \cdot)$  durch die Diagonale der diskretisierten Bilinearform und den Fehler  $e_l^z := z - z_l$  des adjungierten Problems sowie den Fehler  $e_l$  durch den Absolutwert der hierarchischen Koeffizienten der diskreten Lösungen  $u_l = \sum_{(j,i) \in G_l} u_{j,i} \varphi_{j,i}$  und  $z_l = \sum_{(j,i) \in G_l} z_{j,i} \psi_{j,i}$  zu ersetzen. Wir erhalten mit diesem Vorgehen

$$|l(e_l)| \approx \sum_{(j,i) \in G} |u_{j,i}| |b_{j,i}| |z_{j,i}|, \quad (6.13)$$

wobei  $b_{j,i} := B(\psi_{j,i}, \psi_{j,i})$ . Motiviert durch (6.13) verwenden wir dann die Größen

$$\eta_{j,i} := |u_{j,i}| |b_{j,i}| |z_{j,i}| \quad (6.14)$$

als Fehlerindikator. Man beachte, dass die obigen Überlegungen nur als Motivation für die Wahl des Indikators  $\eta_{j,i}$  dienen sollen und einer strengen mathematischen Prüfung nicht standhalten. Die folgenden numerischen Ergebnisse zeigen jedoch, dass dieser Fehlerindikator in der Praxis gute Resultate liefert.

## 6.4 Numerische Ergebnisse

*Beispiel 6.4.* Wir betrachten das eindimensionale Standardproblem

$$\begin{aligned} \partial_t u - \partial_{xx} u &= f \text{ in } \Omega = (0, 1), \\ u(x, 0) &= u_0(x) \text{ für alle } x \in \Omega, \\ u(x, t) &= g(x, t) \text{ für alle } t \in (0, 1] \text{ und } x \in \partial\Omega, \end{aligned} \quad (6.15)$$

wobei die rechte Seite und die Randbedingung derart gesetzt sind, dass wir als exakte Lösung

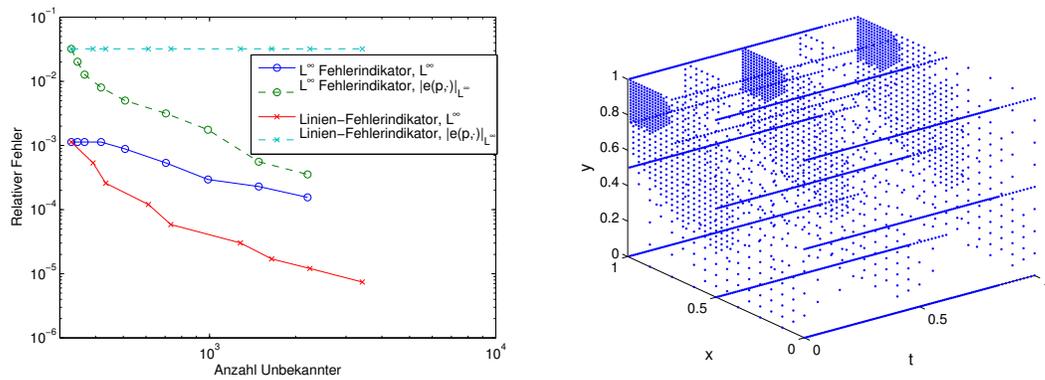
$$u(x, t) = \sin\left(\pi \cdot \frac{x}{t + 0.1}\right)$$

erhalten. Allerdings sind wir nun nicht mehr an dem Fehler innerhalb des gesamten Gebietes (bzgl. einer fest vorgegeben Norm), sondern lediglich an der Kontrolle des Fehlers in dem Punkt  $p = (x_0, t_0) = (0.9, 0.5)$  interessiert. Hierzu wählen wir das Funktional  $l(u) = u(p)$ , indem wir als  $l_0$  die Dirac Distribution an diesem Punkt einsetzen. Man beachte, dass in unserem Fall die exakte Lösung punktauswertbar ist, weshalb wir hier so vorgehen können. Ist die Lösung nicht bekannt und ist demnach auch unklar, ob die Lösung punktauswertbar ist, ersetzt man in der Regel die Dirac Distribution durch eine so genannte *mollification*, vgl. [89].

Wir vergleichen die Ergebnisse des adaptiven Verfahrens bei Verwendung des Fehlerindikators aus Abschnitt 6.1 (Absolutwert des hierarchischen Koeffizienten der diskreten Lösung) und des dualen Fehlerindikators (6.14). In Abbildung 6.7 sind der Punktfehler in dem Punkt  $p$  sowie der Fehler bzgl. der  $L^\infty$ -Norm in Abhängigkeit von der Zahl der Freiheitsgrade dargestellt. Während der adaptive Zyklus bei Verwendung des dualen Fehlerindikators den Fehler gemessen in der  $L^\infty$ -Norm nur in dem ersten Verfeinerungsschritt verkleinert und danach nicht weiter verringert, ist der Fehler im Punkt  $p$  bei nahezu gleicher Anzahl an Freiheitsgraden wesentlich kleiner als im Fall der adaptiven Verfeinerung mittels des hierarchischen Koeffizienten. Das adaptive Dünngitter nach sechs Verfeinerungsschritten ist in Abbildung 6.7 zu sehen. Wie zu erwarten war, ist das Gitter nur für Zeiten  $t \leq 0.5$  verfeinert worden, da spätere Zeitpunkte den Fehler im Punkte  $p$  nicht mehr beeinflussen. Ferner ist erkennbar, dass der größte Teil der Verfeinerung nur für  $x \geq 0.5$  stattgefunden hat, da nur solche Werte einen großen Einfluss auf den Fehler im Punkte  $p$  haben.

*Beispiel 6.5.* Wir untersuchen nun die Verfeinerung mittels des dualen Fehlerschätzers anhand des zweidimensionalen Problems aus Beispiel 6.3 mit Lösung  $u(r, \phi, t) = \exp(-t) \cdot r^{2/3} \sin(\frac{2\phi - \pi}{3})$ . Wir wählen für  $x_0 = (0.9, 0.9)$  das Funktional  $l(u) = \int_0^T u(x_0, t) dt$ , d.h. wir sind daran interessiert, den Fehler an der Stelle  $x_0$  im integralen Mittel über die Zeit möglichst klein zu halten.

Die Ergebnisse für den dualen Fehlerindikator (6.14) und den Fehlerindikator aus Abschnitt 6.1 bzgl. der  $L^\infty$ -Norm und bzgl. des Funktionals sind aus Abbildung 6.5 ersichtlich. Wie schon im vorherigen Beispiel wird der  $L^\infty$ -Fehler bei Verfeinerung mittels des dualen Fehlerindikators kaum reduziert. Betrachten wir hingegen den Fehler im Kostenfunktional, so beobachten wir eine stärkere Reduktion des Fehlers pro Freiheitsgrad bei Verwendung des dualen Indikators als bei Verwendung des Absolutwertes des hierarchischen Koeffizienten. Somit erhalten wir beispielsweise einen Fehler von  $1.2 \cdot 10^{-5}$  bzgl. des Kostenfunktional für 2242 Unbekannte bei Verwendung des dualen Fehlerschätzers, während die Verwendung des Absolutwertes der hierarchischen Koeffizienten zu einem Fehler  $1.6 \cdot 10^{-4}$  mit 2200 Unbekannten



**Abbildung 6.8.** Links: Ergebnisse für die adaptive Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung des Beispiels 6.5 bei Verwendung der hierarchischen Koeffizienten als Fehlerindikator und des dualen Fehlerindikators. Rechts: Das ausgehend von  $\tilde{V}_4^0$  nach sechs Verfeinerungsschritten mit Hilfe des dualen Fehlerindikators resultierende Raum-Zeit Dünngitter für Beispiel 6.5.

führt.

Abbildung 6.5 zeigt zusätzlich das ausgehend von  $\tilde{V}_4^0$  nach acht Verfeinerungsschritten resultierende Raum-Zeit Dünngitter bei Verwendung des dualen Fehlerindikators. Es ist offenkundig, dass nun, obwohl die Singularität der Lösung im Nullpunkt liegt, nicht zu der Singularität, sondern zu dem Punkt  $(0.9, 0.9)$  verfeinert wird, an dem wir mit dem gewählten Funktional den Fehler messen.

In diesem Kapitel haben wir Adaptivität für die Raum-Zeit Dünngitterdiskretisierungen untersucht. Hierbei haben wir sowohl einen Fehlerindikator basierend auf den hierarchischen Koeffizienten der diskreten Lösung als auch einen mit Hilfe der diskreten Lösung des dualen Systems konstruierten Fehlerindikator zur Fehlerschätzung bzgl. linearer Funktionale entwickelt. Insbesondere führt die Verfeinerung des Raum-Zeit Dünngitters mit den vorgestellten Fehlerindikatoren zu Gittern, die das so genannte local time stepping beinhalten, d.h. wir erhalten an unterschiedlichen Orten unterschiedlich feine Zeitaufösungen. Dieses local time stepping bereitet bei klassischen Diskretisierungsverfahren sowohl unter dem Blickwinkel der Fehlerschätzung, als auch unter dem Gesichtspunkt der Algorithmik erhebliche Schwierigkeiten. Die durch die Adaptivität mit Hilfe der beiden Fehlerindikatoren erzielten Resultate der numerischen Ergebnisse sind dabei sehr viel versprechend und zeigen, dass wir mit dem adaptiven Raum-Zeit Dünngitterverfahren ein effizientes Verfahren zur Lösung von Problemen mit nichtglatten Lösungen erhalten.

## Kapitel 7

# Instationäre verteilte Kontrollprobleme

In der Praxis werden häufig Probleme betrachtet, bei denen nicht nur eine Größe  $u$ , sondern mehrere Größen  $u_1, \dots, u_n$  auftreten. Wir erhalten dann ein System parabolischer Gleichungen. Dabei sucht man eine vektorwertige Funktion  $\mathbf{u} = (u_1, \dots, u_n)$ , so dass

$$\partial_t \mathbf{u} - \sum_{i=1}^n \partial_{x_i} \left( \sum_{j=1}^n a_{ij}(x, t) \partial_{x_j} \mathbf{u} + A_i(x, t) \mathbf{u} \right) + \sum_{i=1}^n B_i(x, t) \partial_{x_i} \mathbf{u} + A(x, t) \mathbf{u} = \mathbf{f} \quad (7.1)$$

für gegebene  $a_{ij}(x, t) \in \mathbb{R}$ ,  $1 \leq i \leq j$ ,  $A_i(x, t), B_i(x, t), A(x, t) \in \mathbb{R}^n$ ,  $1 \leq i \leq n$  und  $\mathbf{f}(x, t) \in \mathbb{R}^n$  gilt, und vorgegebene Anfangs- und Randbedingungen erfüllt sind. Unter ähnlichen Voraussetzungen wie für den in Kapitel 3 vorgestellten skalaren Fall besitzt das Problem eine eindeutige Lösung, vgl. [82]. Wir können Probleme der Art (7.1) mit den in Kapitel 4 beschriebenen Raum-Zeit Dünngitterdiskretisierungen lösen, indem wir nun für jede Unbekannte  $u_i$  einen eigenen Ansatz- und Testraum definieren. Im Folgenden illustrieren wir diese Vorgehensweise an dem Beispiel von instationären Kontrollproblemen.

Der Abschnitt 7.1 gibt zunächst eine kurze Einführung in instationäre verteilte Kontrollprobleme. Danach beschreiben wir in Abschnitt 7.2 wie die bisher betrachteten Raum-Zeit Dünngitterdiskretisierungen von skalaren Gleichungen auf den Fall von Systemen parabolischer Gleichungen und insbesondere von instationären Kontrollproblemen verallgemeinert werden können. Dabei beschränken wir uns in unseren Untersuchungen auf die Crank-Nicolson Diskretisierung, da sich die Discontinuous-Galerkin Diskretisierung analog auf den Systemfall übertragen läßt. Wir zeigen schließlich exemplarisch an einigen numerischen Ergebnissen, dass wir bei der Diskretisierung der instationären Kontrollprobleme die gleichen Konvergenzraten, sowohl für den Zustand als auch für die Kontrolle, wie bei den skalaren Problemen erhalten.

In Abschnitt 7.3 behandeln wir die Lösung der aus der Diskretisierung entstandenen diskreten Optimalitätssysteme. Dabei sind die vorgestellten Löser Erweiterungen

rungen des Gauß-Seidel Verfahrens bzw. des Waveform-Gauß-Seidel Verfahrens. Die numerischen Ergebnisse des Abschnitts zeigen, dass wir mit diesen Erweiterungen ähnlich gute Konvergenzresultate für die diskreten Kontrollprobleme wie im Fall der skalaren Problemen erhalten.

Abschließend geben wir in Abschnitt 7.4 einen kurzen Ausblick auf die adaptive Diskretisierung der instationären verteilten Kontrollprobleme und beenden den Abschnitt sowie dieses Kapitel mit einigen Beispielrechnungen.

## 7.1 Instationäre Kontrollprobleme

Oftmals werden parabolische Gleichungen gelöst, um die Auswirkungen eines von außen beeinflussbaren Parameters auf das Gesamtverhalten des Systems zu studieren. Ziel ist es dabei, den Parameter in dem Sinne optimal zu wählen, dass die Lösung der Gleichung bzw. das Verhalten des Systems einem vorher festgelegten Verhalten möglichst ähnlich ist. Diese Aufgabenstellung kann als ein instationäres Kontrollproblem mathematisiert werden, d.h. als Optimierungsproblem, bei dem die parabolische Differentialgleichung als Nebenbedingung auftritt. Als Beispiele sind hier etwa inverse Probleme zur Bestimmung von in den Gleichungen vorkommenden Materialparametern mit Hilfe einiger Messwerte zu nennen.

Zur Lösung solcher Optimierungsprobleme ist dabei als Teilproblem sehr häufig ein System parabolischer Gleichungen zu lösen, das eine interessante Struktur aufweist: Während einige (skalare) Gleichungen dieses Systems in der Zeit vorwärts laufen, sind andere Gleichungen in der Zeit rückwärts gerichtet. Diese Struktur des Systems stellt für die numerische Lösung dieser Probleme eine besondere Herausforderung dar. Um solche Systeme, bei denen miteinander gekoppelte skalare Vorwärts- und Rückwärtsgleichungen auftreten, zu lösen, müssen Lösungen einzelner skalarer Gleichungen zu beliebigen Zeitpunkten auswertbar sein. Damit dies auf dem Computer schnell durchführbar ist, müssen alle Werte zu den berechneten Zeitpunkten im Hauptspeicher gehalten werden. Aufgrund der hohen Zahl ( $O(N^{d+1})$  bzw.  $O(N^{d+2})$  bei uniformer Verfeinerung in Raum und Zeit) an Freiheitsgraden des vollen Raum-Zeit Gitters werden dabei schnell die Grenzen des technisch Machbaren erreicht.

Aus diesen Gründen ist die Fragestellung nach effizienten Verfahren zur Lösung parabolischer Gleichungen insbesondere innerhalb der instationären Kontrollprobleme Gegenstand aktueller Forschung. Dabei ist es nicht nur von Bedeutung, dass die parabolischen Gleichungen effizient gelöst, sondern dass die berechneten Lösungen (zum Beispiel die der Zustandsvariablen) auch zu beliebigen Zeitpunkten schnell ausgewertet werden können (um beispielsweise innerhalb des Lösungsprozesses für die zu der adjungierten Variablen gehörenden skalaren Gleichung verwendet zu werden).

Für Vollgitterdiskretisierungen sind dabei insbesondere zwei Methoden entwickelt worden. Eine Verfahrensklasse zur Reduktion des Speicheraufwands sind die so ge-

nannten *reduced order models* [74, 81], die aus der Lösung des zeitabhängigen Problems an wenigen Zeitpunkten, so genannten *snapshots*, die für die Lösung wichtigsten Orts-Moden berechnen. Dies führt zu wenigen Funktionen im Ort, die die Lösung (hoffentlich) gut über die Zeit approximieren können. Mit Hilfe dieser wenigen Funktionen werden nun die anfallenden parabolischen Probleme gelöst. Wie erwähnt muß dazu allerdings zunächst ein instationäres Problem zu Beginn gelöst werden und es kann sogar notwendig sein, während der Optimierung die errechneten Moden um neue zu ergänzen, so dass auch dann wieder ein parabolisches Problem mit den oben beschriebenen Verfahren gelöst werden muß.

Ein weiteres Verfahren ist das so genannte *checkpointing* [69], bei dem die Lösung der parabolischen Gleichung an verschiedenen, wenigen Zeitpunkten abgespeichert wird. Zur Berechnung der Lösung an einem Zeitschritt, der nicht abgespeichert wurde, wird das parabolische Problem von dem Zeitpunkt aus gelöst, der dem gesuchten Zeitpunkt am nächsten ist, wobei die dort abgespeicherte Lösung als Anfangswert verwendet wird.

Die Diskussion zeigt, wie schwierig die Behandlung von instationären Kontrollproblemen mit klassischen Diskretisierungsverfahren auf vollen Gittern aufgrund der hohen Komplexität ist. Für die Raum-Zeit Dünngitterdiskretisierungen, bei denen alle Freiheitsgrade und somit die Funktion in Raum-Zeit komplett im Hauptspeicher des Rechners verwaltet werden können, stoßen wir nicht auf die beschriebenen Probleme.

### 7.1.1 Verteilte Kontrollprobleme

Bei verteilten Kontrollproblemen versucht man, über eine Veränderung der rechten Seite das System in einen gewünschten Zustand einzustellen. Bezeichnen wir mit  $z$  den gewünschten Zustand und sei  $\gamma > 0$  vorgegeben, so suchen wir eine Kontrolle  $u \in L^2(\Omega_T)$ , so dass das Kostenfunktional  $J(\cdot, \cdot)$

$$J(y, u) := \frac{1}{2} \|y - z\|_{L^2(\Omega_T)}^2 + \frac{\gamma}{2} \|u\|_{L^2(\Omega_T)}^2, \quad (7.2)$$

unter der Nebenbedingung

$$\partial_t y + L(t)y = f + u \text{ für alle } x \in \Omega, t \in (0, T], \quad (7.3)$$

$$y(x, t) = 0 \text{ auf } \partial\Omega \times (0, T], \quad (7.4)$$

$$y(x, 0) = y_0(x) \text{ für alle } x \in \Omega \quad (7.5)$$

minimiert wird. Für eine eingehendere Diskussion und allgemeinere Kostenfunktionale verweisen wir an dieser Stelle auf [84]. Bezeichnen wir mit  $L^*(t)$  die Adjungierte von  $L(t)$  und nehmen wir an, dass die zu  $L(t)$  gehörige Sesquilinearform  $a(t; \cdot, \cdot)$  die Bedingungen A1 und A2 des Abschnitts 3.2 erfüllt, so ist eine Lösung  $(y, u)$  des

Kontrollproblems durch das System

$$\begin{aligned} \partial_t y + L(t)y &= f + u \text{ für alle } x \in \Omega, t \in (0, T], \\ -\partial_t p + L^*(t)p + (y - z) &= 0 \text{ für alle } x \in \Omega, t \in [0, T), \\ \gamma u - p &= 0 \text{ in } \Omega_T \end{aligned}$$

mit den zugehörigen Randbedingungen

$$\begin{aligned} y(x, t) &= 0 \text{ auf } \partial\Omega \times (0, T], \quad y(x, 0) = y_0(x) \text{ für alle } x \in \Omega, \\ p(x, t) &= 0 \text{ auf } \partial\Omega \times [0, T), \quad p(x, T) = 0 \text{ für alle } x \in \Omega, \end{aligned}$$

charakterisiert [84]. Nutzen wir die Relation  $\gamma u = p$  zwischen der Kontrolle  $u$  und dem Lagrangeparameter  $p$ , der üblicherweise auch als Adjungierte bezeichnet wird, aus, so erhalten wir schließlich das System

$$\partial_t y + L(t)y - \frac{1}{\gamma}p = f \text{ für alle } x \in \Omega, t \in (0, T], \quad (7.6)$$

$$-\partial_t p + L^*(t)p + (y - z) = 0 \text{ für alle } x \in \Omega, t \in [0, T), \quad (7.7)$$

mit den Randbedingungen

$$y(x, t) = 0 \text{ auf } \partial\Omega \times (0, T], \quad y(x, 0) = y_0(x) \text{ für alle } x \in \Omega, \quad (7.8)$$

$$p(x, t) = 0 \text{ auf } \partial\Omega \times [0, T), \quad p(x, T) = 0 \text{ für alle } x \in \Omega. \quad (7.9)$$

Somit kann eine Lösung des ursprünglichen verteilten Kontrollproblems (7.2) mit den Nebenbedingungen (7.3), (7.4) und (7.5) durch die Lösung des Systems (7.6), (7.7) mit den entsprechenden Randbedingungen bestimmt werden.

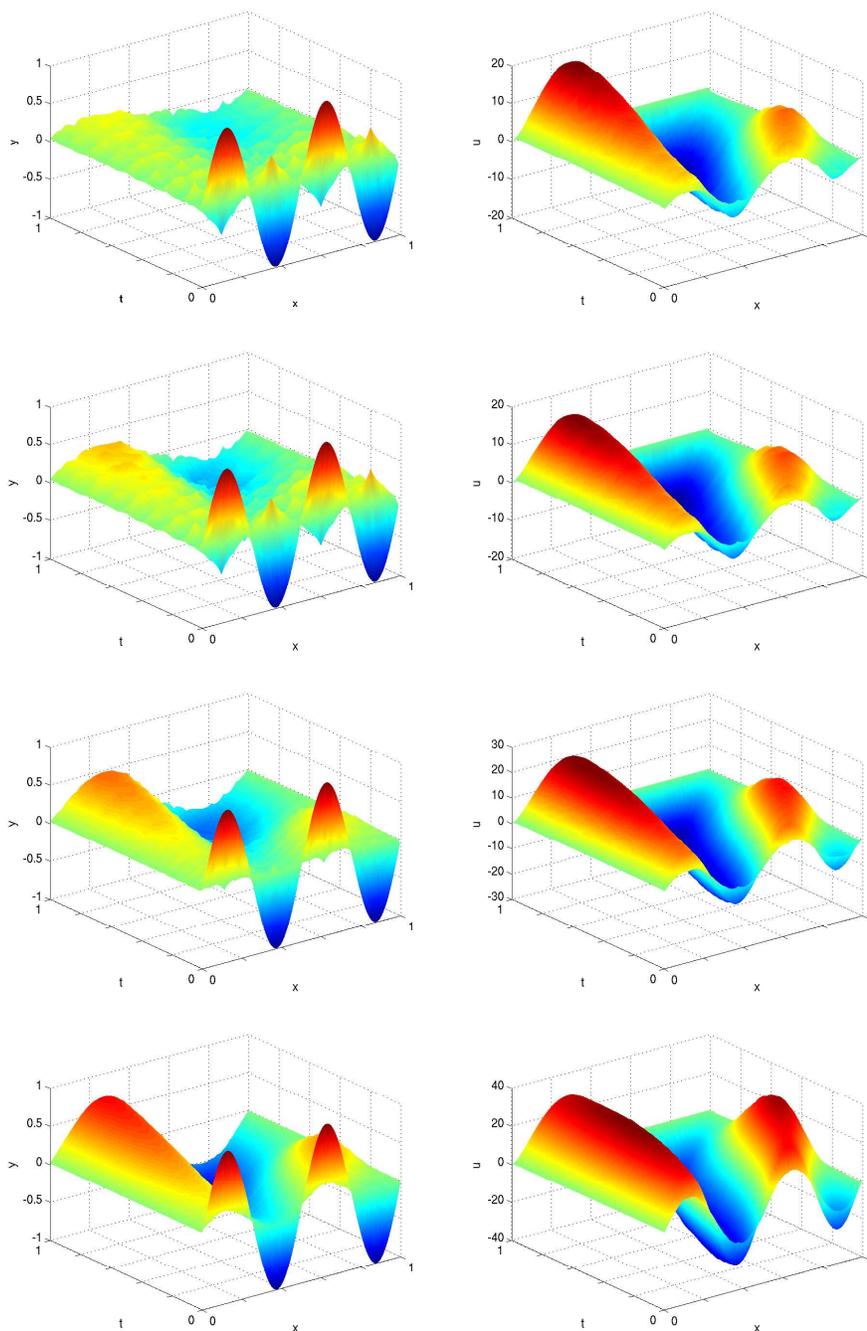
Bevor wir im nächsten Abschnitt die Diskretisierung dieses Systems behandeln, möchten wir an dieser Stelle kurz die Rolle des Parameters  $\gamma$  diskutieren. Verschiedene Werte für den Parameter  $\gamma$  führen zu unterschiedlichen Optimierungsproblemen und damit zu unterschiedlichen Lösungen. Als erstes Beispiel betrachten wir auf  $\Omega_T = (0, 1)^2$  das Kontrollproblem

$$\min_{(y,u)} \frac{1}{2} \|y - \sin(4\pi x/(t+1))\|_{L^2(\Omega_T)}^2 + \frac{\gamma}{2} \|u\|_{L^2(\Omega_T)}^2 \text{ mit} \quad (7.10)$$

$$y(x, t) = 0 \text{ auf } \partial\Omega \times (0, 1], \quad (7.11)$$

$$y(x, 0) = \sin(4\pi x) \text{ für alle } x \in (0, 1). \quad (7.12)$$

Dabei verwenden wir für  $\gamma$  die Werte  $2 \cdot 10^{-3}, 10^{-3}, 5 \cdot 10^{-4}, 2 \cdot 10^{-4}$ . Abbildung 7.1 zeigt die entsprechenden numerisch berechneten Resultate für den Zustand  $y$  und die Kontrolle  $u$ . Man sieht deutlich, dass die betragsmäßig größten Funktionswerte der Kontrolle  $u$  mit kleiner werdendem  $\gamma$  wachsen. Dieser Effekt ist auch zu erwarten, da für kleiner werdendes  $\gamma$  der Einfluss der Kontrolle auf das Kostenfunktional



**Abbildung 7.1.** Zustand  $y$  (linke Spalte) und Kontrolle  $u$  (rechte Spalte) für das Kontrollproblem (7.11) mit  $\gamma = 2 \cdot 10^{-3}, 10^{-3}, 5 \cdot 10^{-4}, 2 \cdot 10^{-4}$  (von oben nach unten).

abnimmt. Gleichzeitig nähert sich zu dem Endzeitpunkt  $T = 1$  der Zustand  $y$  immer mehr dem gewünschten Zustand  $z(x, t) = \sin(4\pi x/(t + 1))$ .

Wie wir sehen, beeinflusst die Wahl des Parameters  $\gamma$  sowohl den Zustand als auch die Kontrolle. Gleichzeitig bestimmt die Größe des Parameters die Stärke der Kopplung zwischen dem Zustand und der Kontrolle über die Gleichung (7.6). Die Stärke der Kopplung kann allerdings Auswirkungen auf die Resultate der zugrunde liegenden numerischen Verfahren haben, wie beispielsweise auf das Konvergenzverhalten des linearen Löser.

Aus diesen Gründen ist die Frage nach einem typischen Wertebereich für den Parameter  $\gamma$  interessant. Wird der Parameter  $\gamma$  als Regularisierungsparameter interpretiert, der lediglich eingeführt wird, um die Lösbarkeit des Optimierungsproblems zu garantieren, wird man bestrebt sein,  $\gamma$  so klein wie möglich zu wählen. Allerdings läßt sich der Term  $\frac{\gamma}{2} \|u\|_{L^2(\Omega_T)}^2$  auch als die Kosten für die Kontrolle  $u$  interpretieren, die sich beispielsweise direkt auf die Kosten eines Produktionsprozesses (Energieaufwand) auswirken. Aus diesem Blickwinkel betrachtet hat der Anwender von Fall zu Fall  $\gamma$  selbst zu wählen, wobei nun je nach Anwendung  $\gamma$  keinesfalls sehr klein sein muß. Wir werden in dem Abschnitt zur Lösung der diskreten Optimalitätsprobleme auf diese Diskussion zurückkommen.

## 7.2 Raum-Zeit Dünngitterdiskretisierung instationärer verteilter Kontrollprobleme

Wir beschäftigen uns nun mit der Raum-Zeit Dünngitterdiskretisierung des durch (7.6) und (7.7) beschriebenen Systems.

Die Crank-Nicolson Diskretisierung dieses Vorwärts-Rückwärts Systems erhalten wir ohne große Modifikationen. Sei wie in Abschnitt 4.1  $\tilde{V}_l^{0, Ansatz}$  der aus der hierarchischen Basis im Ort und der linearen hierarchischen Basis in der Zeit konstruierte Raum-Zeit Dünngitterraum  $\tilde{V}_l^0$  sowie  $\tilde{V}_l^{0, Test}$  der Dünngitterraum, bei dem wir in der Zeit die stückweise konstante Basis verwendet haben. Dann erhalten wir als Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung das diskrete Problem: Finde  $y_l, p_l \in \tilde{V}_l^{0, Ansatz}$  die die homogenen Rand- und Anfangs- bzw. Endbedingungen sowie

$$B(y_l, \varphi) - \frac{1}{\gamma} (p_l, \varphi)_{L^2(\Omega_T)} = \int_0^T (f, \varphi)_0 dt \quad \forall \varphi \in \tilde{V}_l^{0, Test}, \quad (7.13)$$

$$B^*(p_l, \varphi) + (y_l, \varphi)_{L^2(\Omega_T)} = \int_0^T (z, \varphi)_0 dt \quad \forall \varphi \in \tilde{V}_l^{0, Test} \quad (7.14)$$

mit

$$B(v, w) := \int_0^T (\partial_t v, w)_0 + a(t; v, w) dt$$

$$B^*(v, w) := \int_0^T -(\partial_t v, w)_0 + a(t; w, v) dt$$

erfüllen.

Bei der Discontinuous-Galerkin Diskretisierung müssen wir allerdings den speziellen Rückwärtscharakter der Gleichung (7.7) berücksichtigen. Zudem sind wir an den Werten der adjungierten Variablen zum Zeitpunkt  $t = 0$  interessiert, an dem die stückweise konstanten oder linearen Funktionen formal nicht definiert sind. Für festes  $l \in \mathbb{N}$  unterteilen wir das Intervall  $[0, 1)$  nun in  $2^l$  Teilintervalle  $I_k^l$ ,  $1 \leq k \leq 2^l$ ,

$$\tilde{I}_k^l = [(k-1)2^{-l}, k2^{-l})$$

und betrachten analog zu Abschnitt 2.3 Multilevel-Zerlegungen in stückweise konstante oder lineare Funktionen, die nun im Gegensatz zu Abschnitt 2.3 rechts- und nicht mehr linksseitig stetig sind. Sei  $V_l^{0,y}$  der aus der stückweise konstanten oder linearen hierarchischen Basis in der Zeit und der hierarchischen Basis im Ort konstruierte Raum-Zeit Dünngitterraum  $\tilde{V}_l^0$  und  $V_l^{0,p}$  der Dünngitterraum, bei dem wir stattdessen die Zerlegung in rechtsseitig stetige, stückweise konstante bzw. lineare Funktionen in der Zeit verwendet haben. Dann lautet die Discontinuous-Galerkin Diskretisierung des Optimierungsproblems: Gesucht sind  $y_l \in V_l^{0,y}$ ,  $p_l \in V_l^{0,p}$ , die die homogenen Anfangs- bzw. Endbedingungen, die homogenen Randbedingungen sowie

$$B(y_l, \varphi) - \frac{1}{\gamma} (p_l, \varphi)_{L^2(\Omega_T)} = (f, \varphi)_{L^2(\Omega_T)} \text{ für alle } \varphi \in V_l^{0,y},$$

$$B^*(p_l, \varphi) + (y_l, \varphi)_{L^2(\Omega_T)} = (z, \varphi)_{L^2(\Omega_T)} \text{ für alle } \varphi \in V_l^{0,p}$$

mit

$$B(v, w) := \sum_{k=1}^N \int_{I_k} (\partial_t v, w)_{L^2(\Omega)} + a(t; v, w) dt + \sum_{k=0}^{N-1} ([v]_k, w_k^+)_{L^2(\Omega)},$$

$$B^*(v, w) := \sum_{k=1}^N \int_{I_k} -(\partial_t v, w)_{L^2(\Omega)} + a^*(t; v, w) dt - \sum_{k=1}^N ([v]_k, w_k^-)_{L^2(\Omega)}$$

und  $(p_l)_N^+ := 0$  erfüllen.

Für die effiziente Anwendung des Matrix-Vektorproduktes (und den von dem Multilevellöser benötigten Operationen) läßt sich das in Abschnitt 4.3 diskutierte Unidirektionale Prinzip ohne besondere Schwierigkeiten auf den Systemfall übertragen.

**Tabelle 7.1.** Konvergenz der Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung für die Kontrolle  $u = 10^2 \cdot p$  und den Zustand  $y$  für das Problem (7.15) mit  $d = 1$ .

1	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^\infty}$	$\rho_{H^1 \otimes L^\infty}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
Zustand $y$								
4	1.342 <sub>-3</sub>	-	1.237 <sub>-3</sub>	-	1.064 <sub>-2</sub>	-	8.390 <sub>-3</sub>	-
5	3.124 <sub>-4</sub>	0.23	2.653 <sub>-4</sub>	0.21	5.126 <sub>-3</sub>	0.48	3.670 <sub>-3</sub>	0.44
6	7.622 <sub>-5</sub>	0.24	5.583 <sub>-5</sub>	0.21	2.546 <sub>-3</sub>	0.50	1.780 <sub>-3</sub>	0.48
7	1.811 <sub>-5</sub>	0.24	1.143 <sub>-5</sub>	0.20	1.272 <sub>-3</sub>	0.50	8.856 <sub>-4</sub>	0.50
8	4.245 <sub>-6</sub>	0.23	2.262 <sub>-6</sub>	0.20	6.359 <sub>-4</sub>	0.50	4.427 <sub>-4</sub>	0.50
9	9.912 <sub>-7</sub>	0.23	4.341 <sub>-7</sub>	0.19	3.180 <sub>-4</sub>	0.50	2.215 <sub>-4</sub>	0.50
10	2.467 <sub>-7</sub>	0.25	8.506 <sub>-8</sub>	0.20	1.590 <sub>-4</sub>	0.50	1.108 <sub>-4</sub>	0.50
Kontrolle $u$								
4	1.387 <sub>-2</sub>	-	1.617 <sub>-2</sub>	-	4.385 <sub>-2</sub>	-	5.321 <sub>-2</sub>	-
5	3.701 <sub>-3</sub>	0.27	4.217 <sub>-3</sub>	0.26	1.522 <sub>-2</sub>	0.35	1.556 <sub>-2</sub>	0.29
6	9.742 <sub>-4</sub>	0.26	1.097 <sub>-3</sub>	0.26	6.764 <sub>-3</sub>	0.44	5.365 <sub>-3</sub>	0.34
7	2.559 <sub>-4</sub>	0.26	2.850 <sub>-4</sub>	0.26	3.260 <sub>-3</sub>	0.48	2.246 <sub>-3</sub>	0.42
8	6.758 <sub>-5</sub>	0.26	7.394 <sub>-5</sub>	0.26	1.613 <sub>-3</sub>	0.49	1.057 <sub>-3</sub>	0.47
9	1.788 <sub>-5</sub>	0.26	1.915 <sub>-5</sub>	0.26	8.042 <sub>-4</sub>	0.50	5.193 <sub>-4</sub>	0.49
10	4.716 <sub>-6</sub>	0.26	4.957 <sub>-6</sub>	0.26	4.018 <sub>-4</sub>	0.50	2.584 <sub>-4</sub>	0.50

**Tabelle 7.2.** Konvergenz der Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung für die Kontrolle  $u = 10^2 \cdot p$  und den Zustand  $y$  für das Kontrollproblem (7.15) mit  $d = 2$ .

1	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^\infty}$	$\rho_{H^1 \otimes L^\infty}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
Zustand $y$								
4	2.581 <sub>-3</sub>	-	2.121 <sub>-3</sub>	-	1.382 <sub>-2</sub>	-	1.180 <sub>-2</sub>	-
5	4.973 <sub>-4</sub>	0.19	4.171 <sub>-4</sub>	0.20	6.411 <sub>-3</sub>	0.46	4.708 <sub>-3</sub>	0.40
6	1.225 <sub>-4</sub>	0.25	7.749 <sub>-5</sub>	0.19	3.160 <sub>-3</sub>	0.49	2.238 <sub>-3</sub>	0.48
7	2.835 <sub>-5</sub>	0.23	1.328 <sub>-5</sub>	0.17	1.578 <sub>-3</sub>	0.50	1.111 <sub>-3</sub>	0.50
8	6.797 <sub>-6</sub>	0.24	2.233 <sub>-6</sub>	0.17	7.887 <sub>-4</sub>	0.50	5.552 <sub>-4</sub>	0.50
9	1.792 <sub>-6</sub>	0.26	5.782 <sub>-7</sub>	0.26	3.944 <sub>-4</sub>	0.50	2.777 <sub>-4</sub>	0.50
Kontrolle $u$								
4	1.272 <sub>-2</sub>	-	1.760 <sub>-2</sub>	-	4.834 <sub>-2</sub>	-	5.955 <sub>-2</sub>	-
5	3.501 <sub>-3</sub>	0.28	4.797 <sub>-3</sub>	0.27	1.789 <sub>-2</sub>	0.37	1.840 <sub>-2</sub>	0.31
6	9.543 <sub>-4</sub>	0.27	1.296 <sub>-3</sub>	0.27	8.273 <sub>-3</sub>	0.46	6.608 <sub>-3</sub>	0.36
7	2.649 <sub>-4</sub>	0.28	3.482 <sub>-4</sub>	0.27	4.031 <sub>-3</sub>	0.49	2.813 <sub>-3</sub>	0.43
8	7.280 <sub>-5</sub>	0.27	9.307 <sub>-5</sub>	0.27	2.000 <sub>-3</sub>	0.50	1.327 <sub>-3</sub>	0.47
9	1.976 <sub>-5</sub>	0.27	2.477 <sub>-5</sub>	0.27	9.975 <sub>-4</sub>	0.50	6.517 <sub>-4</sub>	0.49

## 7.2.1 Numerische Ergebnisse

Wir untersuchen nun anhand einiger numerischer Experimente die Konvergenzeigenschaften der Crank-Nicolson Dünngitterdiskretisierung für das durch (7.6) und (7.7) beschriebene Optimalitätssystem.

**Tabelle 7.3.** Konvergenz der Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung für die Kontrolle  $u = 10^2 \cdot p$  und den Zustand  $y$  für das Kontrollproblem (7.15) mit  $d = 3$ .

1	$e_{L^\infty}$	$\rho_{L^\infty}$	$e_{L^2}$	$\rho_{L^2}$	$e_{H^1 \otimes L^\infty}$	$\rho_{H^1 \otimes L^\infty}$	$e_{H^1 \otimes L^2}$	$\rho_{H^1 \otimes L^2}$
Zustand $y$								
4	$3.805_{-3}$	-	$1.443_{-3}$	-	$1.195_{-2}$	-	$1.016_{-2}$	-
5	$7.673_{-4}$	0.20	$2.446_{-4}$	0.17	$5.531_{-3}$	0.46	$4.023_{-3}$	0.40
6	$1.600_{-4}$	0.21	$3.735_{-5}$	0.15	$2.732_{-3}$	0.49	$1.930_{-3}$	0.48
Kontrolle $u$								
4	$1.167_{-2}$	-	$1.083_{-2}$	-	$3.597_{-2}$	-	$4.552_{-2}$	-
5	$3.217_{-3}$	0.28	$3.117_{-3}$	0.29	$1.474_{-2}$	0.41	$1.502_{-2}$	0.33
6	$9.459_{-4}$	0.29	$8.799_{-4}$	0.28	$7.072_{-3}$	0.48	$5.628_{-3}$	0.37

Ebenso wie in Abschnitt 4.4 messen wir den Fehler zwischen der Lösung des diskreten Problems und der Interpolierenden auf dem entsprechenden vollen Raum-Zeit Gitter. Hierzu betrachten wir für  $d = 1, 2, 3$  das Problem

$$\begin{aligned} \partial_t y - \Delta y - \frac{1}{\gamma} p &= f \text{ für alle } x \in \Omega := (0, 1)^d, t \in (0, 1], \\ -\partial_t p - \Delta p + y &= z \text{ für alle } x \in \Omega, t \in [0, 1) \end{aligned} \quad (7.15)$$

mit den homogenen Anfangs- und Randbedingungen

$$\begin{aligned} y(x, t) &= 0 \text{ auf } \partial\Omega \times (0, 1], \quad y(x, 0) = 0 \text{ für alle } x \in \Omega, \\ p(x, t) &= 0 \text{ auf } \partial\Omega \times [0, 1), \quad p(x, 1) = 0 \text{ für alle } x \in \Omega. \end{aligned}$$

Dabei wählen wir  $\gamma = 10^{-2}$  und die rechte Seite  $f$  sowie den bevorzugten Zustand  $z$  derart, dass wir als Lösung

$$y(x, t) = \exp(-t) \prod_{i=1}^d \sin\left(\pi \cdot \frac{x_i}{t + 0.5}\right), \quad p(x, t) = 10^{-2}(1-t)^2 \cdot \prod_{i=1}^d x_i \cdot (1-x_i)$$

erhalten. Tabelle 7.1 enthält die Ergebnisse für den Zustand  $y$  und die Kontrolle  $u = 10^2 p$  im Fall  $d = 1$ . Wir beobachten im Wesentlichen das gleiche Verhalten des Fehlers wie bei den numerischen Experimenten des Abschnitts 4.4 für die skalaren Probleme. Sowohl für die Kontrolle als auch für den Zustand liefert das Verfahren bzgl. der  $H_{mix}^{1,0}$ - und der  $H_{mix}^{(1,2),(0,\infty)}$ -Norm eine Fehlerreduktion  $\rho \approx 0.5$ . Für die  $L^2$ - und die  $L^\infty$ -Normen erreichen wir bei diesem Beispiel sogar etwas bessere Fehlerreduktionsraten als in den früheren numerischen Experimenten. So beträgt die Fehlerreduktionsrate bzgl. der  $L^2$ -Norm für den Zustand  $y$  bei den betrachteten Levels etwa 0.2, während wir für die Kontrolle eine Reduktion von 0.26 erreichen. Wir erwarten dennoch, dass sich die Reduktion bzgl. der  $L^2$ -Norm asymptotisch dem Wert  $\rho = 0.25$  nähert.

Für den Fall  $d = 2$  sind die entsprechenden Ergebnisse in Tabelle 7.2 abgebildet. Auffällig ist, dass sich die Reduktionsraten bzgl. der  $H_{mix}^{1,0}$ -Norm für den Zustand dem Wert  $\rho = 0.5$  mit steigender Levelzahl von unten nähern. Insgesamt verhält sich auch in diesem Fall das Crank-Nicolson Verfahren wie bei den früher betrachteten skalaren Problemen. Auch für das dreidimensionale Problem, für das die Ergebnisse in Tabelle 7.3 zu finden sind, liefert das Crank-Nicolson Verfahren das gleiche Konvergenzverhalten wie in Abschnitt 4.4.

Die Ergebnisse deuten darauf hin, dass wir das Raum-Zeit Dünngitterprinzip mit wenig Aufwand zur Diskretisierung von in der Zeit gekoppelten Vorwärts-Rückwärts Gleichungen der Art (7.6) und (7.7) verwenden können. Hierbei ändert sich das Konvergenzverhalten im Vergleich zu dem skalaren Fall nicht. Dabei haben wir uns in unseren Untersuchungen auf die Crank-Nicolson Diskretisierung beschränkt. Allerdings ist das gleiche Verhalten ebenfalls für die Discontinuous-Galerkin Diskretisierung in Experimenten zu beobachten. Daher verzichten wir an dieser Stelle aus Platzgründen auf die Darstellung der Ergebnisse.

Wir untersuchen nun im nächsten Abschnitt, wie die bei der Diskretisierung der Optimalitätssysteme anfallenden linearen Gleichungssysteme gelöst werden können.

### 7.3 Die Lösung des diskretisierten Optimalitätssystems

Um die bei der Raum-Zeit Dünngitterdiskretisierung von (7.6) und (7.7) anfallenden linearen Gleichungssysteme zu lösen, können wir erneut Multilevel-Iterationsverfahren, wie in Kapitel 5 beschrieben, aus klassischen Iterationsverfahren über dem Erzeugendensystem konstruieren.

Hierbei ist allerdings darauf zu achten, dass die anfallenden Gleichungssysteme aus der Diskretisierung eines System von parabolischen Differentialgleichungen stammen und nicht mehr nur aus einer einzigen skalaren Gleichung hervorgehen. Im Fall geometrischer Mehrgittermethoden hat sich für solche Probleme oftmals herausgestellt, dass Glätter, die für die Lösung der Diskretisierung skalarer Probleme im Mehrgitterzyklus funktionieren, für den Systemfall angepasst werden müssen [108], um ähnlich gute Konvergenzresultate zu liefern. Dies ist insbesondere dann der Fall, wenn die Kopplung zwischen den einzelnen Gleichungen besonders stark ist. An dieser Stelle ist bei den instationären verteilten Kontrollproblemen dem Regularisierungsparameter  $\gamma$  besondere Bedeutung beizumessen, da er die Stärke der Kopplung beeinflusst.

Die Arbeiten [22, 23, 24] zeigen für die Lösung mittels Raum-Zeit Mehrgitterverfahren zur Lösung klassischer Diskretisierungen auf vollen Gittern der obigen Optimalitätssysteme, dass man nur unter dem Einsatz eines speziellen Glätters bzgl. des Parameters  $\gamma$  robuste Verfahren erhält. Diese Glätter beruhen auf der zweimaligen Anwendung des Block-Gauß-Seidel Verfahrens. Hierbei liegen die zu dem gleichen Zeit- und Ortsgitterpunkt gehörigen Unbekannten der Adjungierten und des

Zustands in dem gleichen Block. Die zu dem Zustand gehörigen Freiheitsgrade werden dabei durch ein Block-Gauß-Seidel Verfahren, das die Unbekannten in der Zeit vorwärts durchläuft, relaxiert, während die zur Adjungierten gehörigen Freiheitsgrade durch ein Block-Gauß-Seidel Verfahren, das in der Zeit das System rückwärts durchläuft, korrigiert werden.

Dieser Ansatz läßt sich ohne Schwierigkeiten auf die Multilevelverfahren des Kapitels 5 übertragen. Ein Nachteil ist dabei jedoch, dass durch die zwei separaten Block-Relaxationen bzgl. der Adjungierten und des Zustands die resultierenden Löser doppelt so viele Rechenoperationen pro Freiheitsgrad benötigen, wie die Löser für die skalaren Probleme. Aus diesem Grund möchten wir auf den Verfahren des Kapitels 5 beruhende alternative Lösungsverfahren vorstellen, die den gleichen Rechenaufwand wie die skalaren Löser besitzen, dabei allerdings nicht robust bzgl. des Regularisierungsparameters  $\gamma$  sind. In den numerischen Experimenten hat sich jedoch herausgestellt, dass wir für Werte  $\gamma \approx 10^{-3}$  noch gute Konvergenzresultate erhalten. Wie die Diskussion bzgl. des Parameters  $\gamma$  in Abschnitt 7.1.1 gezeigt hat, sind wir damit in der Lage, eine Vielzahl von Problemen zu lösen, bei denen die Kontrolle tatsächlich Kosten verursacht und dementsprechend  $\gamma$  nicht beliebig klein gewünscht wird. Um vollständig robuste Verfahren zu erhalten, sollten allerdings in Zukunft gemäß [22, 23, 24] entsprechend modifizierte, auf das Kapitel 5 basierende, Multilevellöser entwickelt und untersucht werden.

Bei der Weiterentwicklung des Gauß-Seidel und des Waveform-Gauß-Seidel Verfahren zur Lösung der diskreten Optimalitätssysteme sollte beachtet werden, dass der Lagrangeparameter  $p$  rückwärts in der Zeit durch die Gleichung (7.7) bestimmt wird, was auch in der Durchlaufreihenfolge des zugrunde liegenden Iterationsverfahrens berücksichtigt werden muß.

Zunächst unterteilen wir die Unbekannten in zwei Mengen. Eine Menge enthält alle zur adjungierten Variablen gehörenden Unbekannten, während die andere Menge aus allen zum Zustand gehörenden Freiheitsgraden besteht. Damit erhalten wir das  $2 \times 2$ -Blocksystem

$$\begin{pmatrix} L_{yy}^E & L_{yp}^E \\ L_{py}^E & L_{pp}^E \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{z} \end{pmatrix} \quad (7.16)$$

wobei  $L_{yy}^E$  die Diskretisierung von  $B(\cdot, \cdot)$  bzgl. des für den Zustand  $y$  verwendeten Erzeugendensystems und entsprechend  $L_{pp}^E$  die Diskretisierung von  $B^*(\cdot, \cdot)$  bzgl. des für die Adjungierte gewählten Erzeugendensystems sind. Die Blöcke auf den Nebendiagonalen sind dementsprechend die Diskretisierungen der Kopplungsterme  $-\frac{1}{\gamma}(\cdot, \cdot)_{L^2(\Omega_T)}$  und  $(\cdot, \cdot)_{L^2(\Omega_T)}$ . Wir könnten versuchen, das Block-System mit Hilfe eines Block-Gauß-Seidel Verfahrens mit inexaktem Block-Löser zu lösen. Hierbei würden sich die Multilevelverfahren des Kapitels 5 als Teilproblemlöser anbieten. Wir erwarten jedoch, dass die Konvergenzeigenschaften dieser Methode sehr stark von der Struktur der Nebendiagonalblöcke und damit von der Größe des Parameters  $\gamma$  abhängen. Aus diesem Grund relaxieren wir stattdessen mit einem Teilschritt des

Gauß-Seidel oder des Waveform-Gauß-Seidel Verfahrens abwechselnd jeweils immer nur eine Unbekannte in jedem Block. Dabei durchlaufen wir die zur adjungierten Variablen gehörigen Unbekannten rückwärts in der Zeit.

Je nachdem, ob wir das Gauß-Seidel oder das Waveform-Gauß-Seidel Verfahren dabei verwenden, sprechen wir für das resultierende Verfahren zur Vereinfachung wiederum von dem Gauß-Seidel bzw. Waveform-Gauß-Seidel Verfahren.

*Beispiel 7.1.* Wir betrachten das Problem

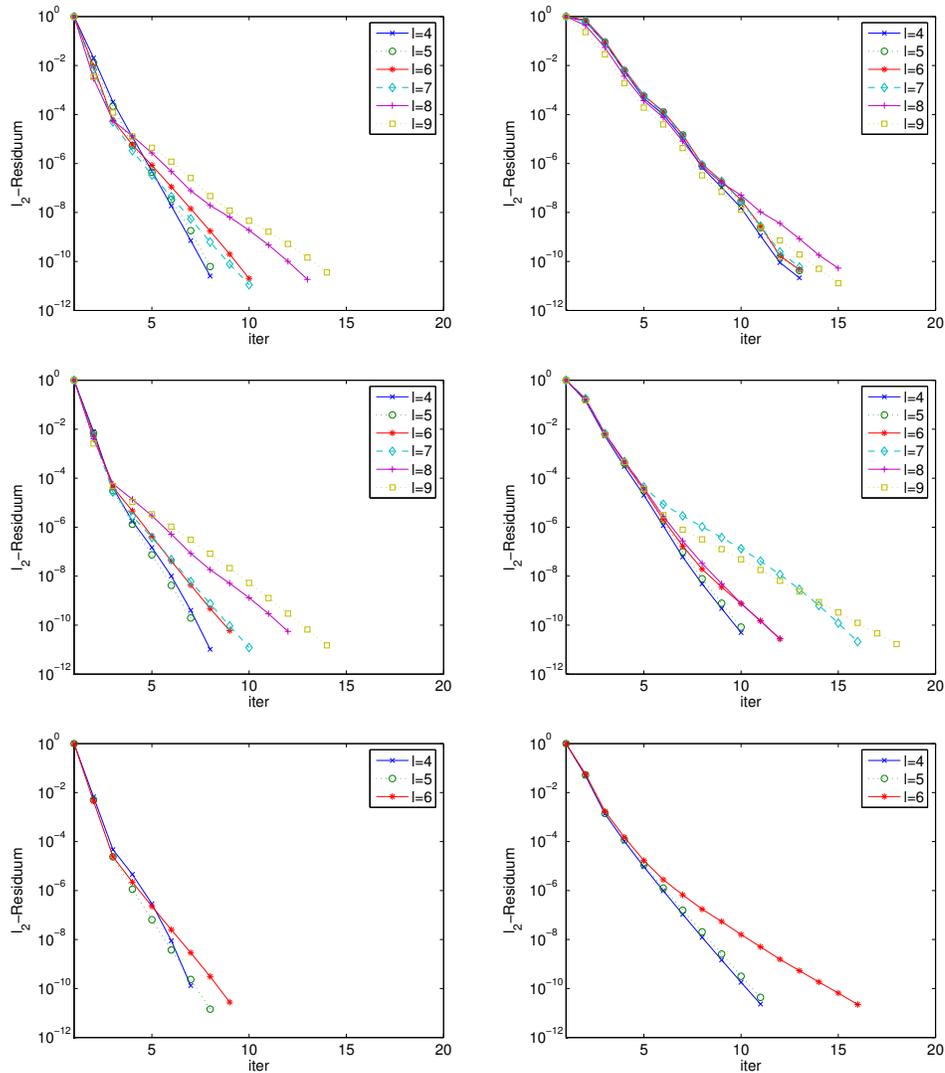
$$\begin{aligned} \partial_t y - \Delta y - \frac{1}{\gamma} p &= f \text{ für alle } x \in \Omega := (0, 1)^d, t \in (0, 1], \\ -\partial_t p - \Delta p + y &= z \text{ für alle } x \in \Omega, t \in [0, 1) \end{aligned}$$

mit  $\gamma = 10^{-3}$ ,  $f = 0$  sowie homogenen Rand- und Anfangs- bzw. Endzeitbedingungen. Dementsprechend ist die exakte Lösung des mit dem Crank-Nicolson Verfahren diskretisierten Problems durch den Nullvektor gegeben. Als Startwert verwenden wir einen Zufallsvektor, der derart normiert ist, dass die euklidische Norm des Residuums dem Wert eins entspricht.

Die Abbildung 7.2 zeigt die Ergebnisse für das Gauß-Seidel und das Waveform-Gauß-Seidel Verfahren mit  $d = 1, 2, 3$ . Wir beobachten im wesentlichen das gleiche Konvergenzverhalten wie für die skalaren Probleme des Abschnitts 5.4. Während sich die Konvergenz des Waveform-Gauß-Seidel Verfahrens mit wachsender Ortsdimension verlangsamt, konvergiert das Gauß-Seidel Verfahren für festes Level  $l$  mit von  $d$  unabhängigen Konvergenzraten. Allerdings beobachten wir für beide Verfahren keine Unabhängigkeit der Konvergenzrate von der Leveltiefe  $l$  der Diskretisierung. Dieser Effekt ist bereits in Abschnitt 5.4 für skalare Probleme diskutiert worden, wobei sich herausstellte, dass eine Kombination aus beiden Verfahren, wobei wir zunächst eine Vorwärtsiteration des Gauß-Seidel Verfahrens und dann eine Rückwärtsiteration des Waveform-Gauß-Seidel Verfahrens ausführen, zu levelunabhängigen Konvergenzraten führt. Diesen Effekt beobachten wir auch in Abbildung 7.3, in der die entsprechenden Resultate für die Kombination aus beiden Verfahren dargestellt sind.

## 7.4 Adaptivität

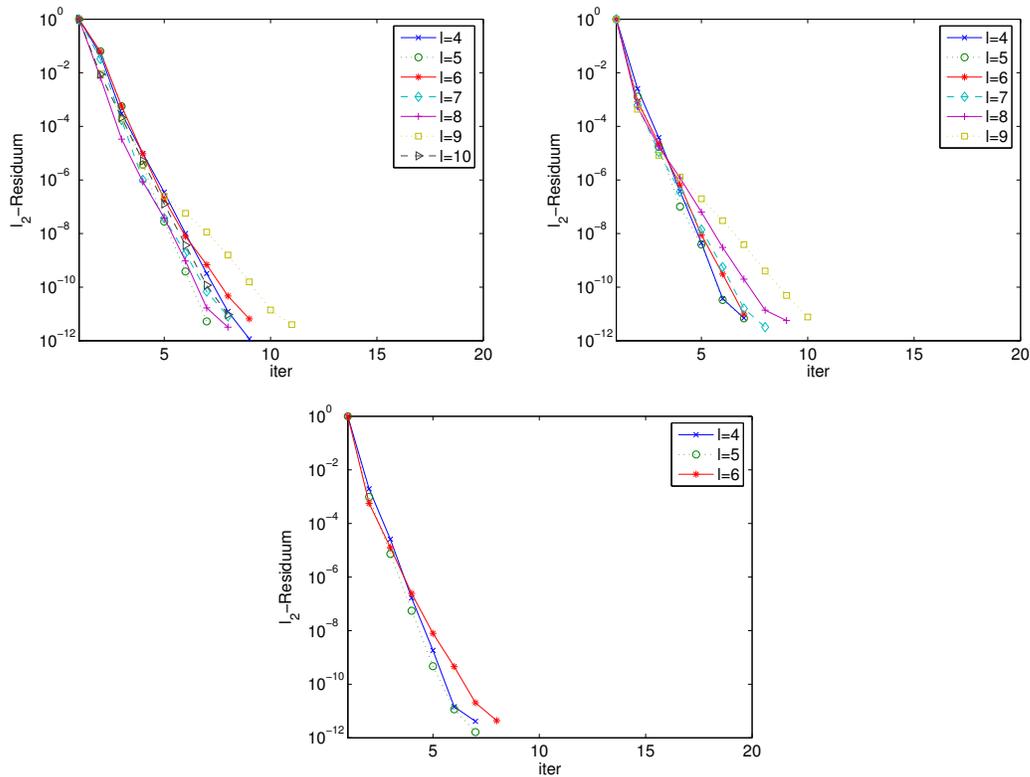
Zum Abschluss dieses Kapitels geben wir in diesem Abschnitt einen Ausblick auf den Einsatz von Adaptivität bei der Diskretisierung instationärer verteilter Kontrollprobleme. Adaptivität für die Diskretisierung von Kontrollproblemen mittels Finiter Elemente ist Gegenstand reger Forschung [11, 12, 14]. Im Bereich von klassischen Dünngitterdiskretisierungen sind zu diesem Thema bisher keinerlei Untersuchungen angestellt worden. Aus diesen Gründen würde eine tiefer gehende Behandlung von Adaptivität im Kontext der Raum-Zeit Dünngitterdiskretisierung von Kontrollproblemen sicherlich den Rahmen der vorliegenden Arbeit sprengen.



**Abbildung 7.2.** Konvergenzhistorie des Gauß-Seidel (links) und des Waveform-Gauß-Seidel Verfahrens (rechts) für die Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung auf verschiedenen Leveln  $l$  des Problems (7.15) und  $d = 1$  (oben),  $d = 2$  (mitte),  $d = 3$  (unten).

Stattdessen möchten wir einen kleinen Einblick geben, wie ohne einen Mehraufwand mit den bestehenden Datenstrukturen und Fehlerindikatoren instationäre Kontrollprobleme adaptiv behandelt werden können.

Hierzu benötigen wir zunächst einen Fehlerindikator bzw. Fehlerschätzer. Die oben zitierten Arbeiten geben dazu im Kontext Finiter Elemente Fehlerschätzer an, die die Größe  $|J(y, u) - J(y_l, u_l)|$  zwischen der exakten Lösung  $(y, u)$  und der



**Abbildung 7.3.** Konvergenzhistorie des aus dem Gauß-Seidel und dem Waveform-Gauß-Seidel Verfahren kombinierte Iterationsverfahrens für die Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung auf verschiedenen Leveln  $l$  des Problems (7.15) für  $d = 1$  (links oben),  $d = 2$  (rechts oben),  $d = 3$  (unten).

approximierten Lösung  $(y_l, u_l)$  abschätzen. Die dort beschriebenen Ansätze lassen sich im Wesentlichen ähnlich wie im Fall der Fehlerschätzung bzgl. linearer Funktionale des Abschnitts 6.3 zur Konstruktion von Fehlerindikatoren für die Raum-Zeit Dünngitter Diskretisierung instationärer Kontrollprobleme heranziehen.

Zur Demonstration der grundsätzlichen Funktionsweise der adaptiven Raum-Zeit Dünngitterdiskretisierung verwenden wir jedoch den in Abschnitt 6.1 vorgestellten, auf den hierarchischen Koeffizienten der diskreten Lösung basierenden Fehlerindikator. Die Konstruktion anderer Fehlerindikatoren sollte Gegenstand zukünftiger intensiver Untersuchungen sein. Wir beschränken unsere Darstellung auf den Fall der Crank-Nicolson Diskretisierung, bei dem sowohl der Ansatz- als auch der Testraum für die adjungierte Variable und den Zustand jeweils gleich sind. Die Überlegungen können allerdings direkt auf die Discontinuous-Galerkin Diskretisierung übertragen werden.

Wie in Abschnitt 6.1 sei  $G$  die Menge von Multiindizes, die beschreibt, welche

Basisfunktionen  $\{\psi_{\mathbf{j},\mathbf{i}}\}$  zur Diskretisierung des Zustands  $y$  und der Adjungierten  $p$  herangezogen werden sollen. Wir verwenden also eine gemeinsame Indexmenge, d.h. das gleiche Gitter zur Diskretisierung der adjungierten Variablen und des Zustands. Es wäre durchaus auch denkbar, die beiden Größen auf zwei verschiedenen Gittern zu diskretisieren. Wie ein späteres Beispiel zeigen wird, kann es tatsächlich vorkommen, dass für die Adjungierte und den Zustand in jeweils unterschiedlichen Regionen des Gebietes das Gitter verfeinert werden muß. Allerdings ist die Anzahl der Rechenoperationen pro Freiheitsgrad des in Abschnitt 4.3 behandelten Unidirektionalen Prinzips zur Berechnung des Matrix-Vektor Produktes im Fall getrennt verwalteter Gitter gleich der Zahl an Rechenoperationen für ein Gitter, das aus der Vereinigung der beiden Gitter besteht. Eine separate Verwaltung von zwei Gittern bringt daher keine Reduktion bezüglich des Rechenaufwands. Man würde also bei zwei Gittern lediglich weniger Unbekannte benötigen und den damit verbundenen Speicher sparen. Gleichzeitig steigt für zwei Gitter verglichen mit nur einem Gitter der Verwaltungsaufwand und der Speicherplatzbedarf, so dass bei der Verwendung zweier Gitter unter Umständen zwar weniger Unbekannte, jedoch ein ähnlicher Speicherplatz benötigt wird und zusätzlich die Rechenzeit ansteigt. Aus diesen Gründen verwenden wir im Folgenden für beide Größen ein gemeinsames adaptives Raum-Zeit Dünngitter.

Wir verfolgen nun im Wesentlichen die gleiche Strategie wie in Abschnitt 6.1: Sind ein Gitter  $G$  und dazugehörige Lösungen der Diskretisierung  $y_G = \sum_{(\mathbf{j},\mathbf{i})} y_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}}$  und  $p_G = \sum_{(\mathbf{j},\mathbf{i})} p_{\mathbf{j},\mathbf{i}} \psi_{\mathbf{j},\mathbf{i}}$  gegeben, so berechnen wir für gegebene  $\gamma_{\mathbf{j},\mathbf{i}}, \varepsilon_p, \varepsilon_y > 0$  die Menge  $G^+$  durch

$$G^+ := \{(\bar{\mathbf{j}}, \bar{\mathbf{i}}) \mid \exists (\mathbf{j}, \mathbf{i}) \in G \text{ so dass } \psi_{\bar{\mathbf{j}},\bar{\mathbf{i}}} \in \mathcal{H}(\psi_{\mathbf{j},\mathbf{i}}) \text{ und } \gamma_{\mathbf{j},\mathbf{i}} \cdot |p_{\mathbf{j},\mathbf{i}}| \geq \varepsilon_u \text{ oder } \gamma_{\mathbf{j},\mathbf{i}} \cdot |y_{\mathbf{j},\mathbf{i}}| \geq \varepsilon_y\}.$$

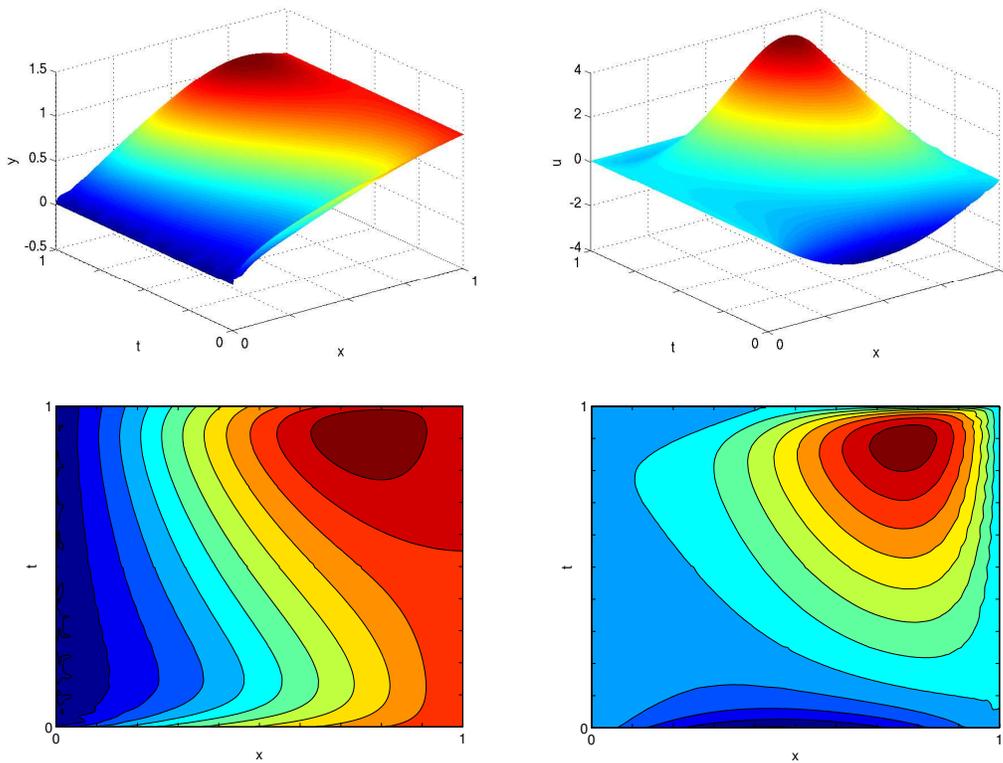
Hierbei kann  $\gamma_{\mathbf{j},\mathbf{i}}$  je nach der Norm, bzgl. der wir den Fehler verringern möchten, variieren. Wie in Abschnitt 6.1 wählen wir für unsere Betrachtungen  $\gamma_{\mathbf{j},\mathbf{i}} = 1$ . Im Gegensatz zum skalaren Fall benötigen wir nun anstelle von einem Parameter zwei Parameter  $\varepsilon_p, \varepsilon_y$ , die wir analog zum skalaren Fall für  $0 \leq c < 1$  fest gewählt durch

$$\varepsilon_p := c \cdot \max_{\substack{(\mathbf{j},\mathbf{i}) \in G, \\ \mathcal{H}(\psi_{\mathbf{j},\mathbf{i}}) \not\subset G}} \gamma_{\mathbf{j},\mathbf{i}} \cdot |p_{\mathbf{j},\mathbf{i}}|, \varepsilon_y := c \cdot \max_{\substack{(\mathbf{j},\mathbf{i}) \in G, \\ \mathcal{H}(\psi_{\mathbf{j},\mathbf{i}}) \not\subset G}} \gamma_{\mathbf{j},\mathbf{i}} \cdot |y_{\mathbf{j},\mathbf{i}}|, \quad (7.17)$$

berechnen. Wir erhalten damit das neue Gitter als die Vereinigung des alten Gitters mit dem Gitter  $G^+$ ,  $G = G \cup G^+$ .

*Beispiel 7.2.* Wir betrachten das Problem

$$\begin{aligned} \partial_t y - \partial_{xx} y - 5 \cdot 10^3 p &= 0 \text{ für alle } x \in \Omega := (0, 1), t \in (0, 1], \\ -\partial_t p - \partial_{xx} p + y &= x^2(1+t)^2 \text{ für alle } x \in \Omega, t \in [0, 1), \end{aligned}$$

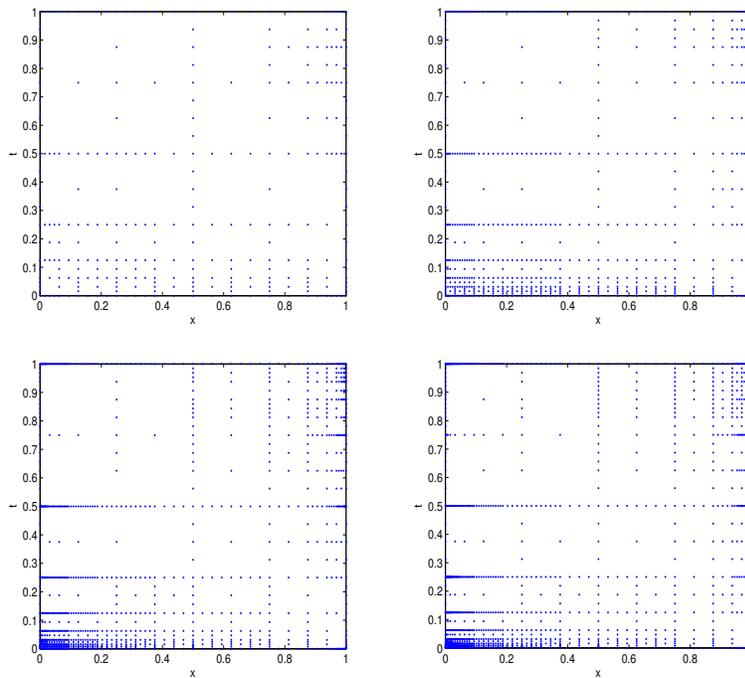


**Abbildung 7.4.** *Mit der Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung berechnete Lösung zu Beispiel 7.2. Es sind dargestellt: Links oben der Surface-Plot des Zustand, rechts oben der Surface-Plot der Kontrolle, links unten Kontourflächen des Zustands und rechts unten Kontourflächen der Kontrolle.*

mit den Randbedingungen

$$\begin{aligned}
 y(x, 0) &= \sqrt{x} \text{ für alle } x \in \Omega, \\
 y(0, t) &= 0 \text{ für alle } t \in (0, T], \quad y(1, t) = 1 \text{ für alle } t \in (0, T], \\
 p(x, T) &= 0 \text{ für alle } x \in \Omega, \quad p(x, t) = 0 \text{ auf } \partial\Omega \times [0, 1].
 \end{aligned}$$

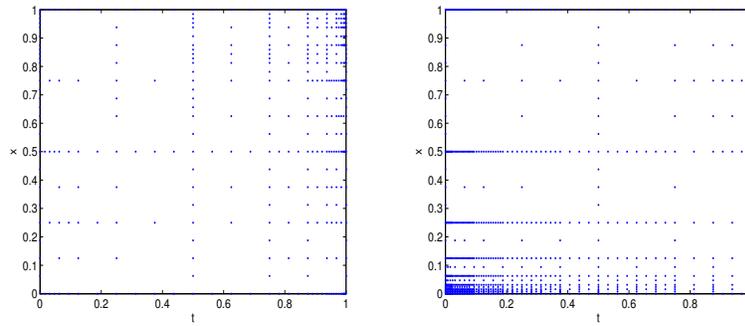
Bedingt durch die Anfangsbedingung  $y(x, 0) = \sqrt{x}$  besitzt  $y$  zum Zeitpunkt  $t = 0$  an der Stelle  $x = 0$  eine Singularität. Wir erwarten daher eine Verfeinerung des Gitters zu diesem Punkt. Die Zielfunktion  $z = x^2(1+t)^2$  nimmt auf dem Gebiet  $[0, 1] \times [0, 1]$  ihr Maximum im Punkte  $(x, t) = (1, 1)$  an und es gilt  $z(1, 1) = 4$ . Zusätzlich fixiert die Randbedingung den Zustand  $y$  in diesem Punkt auf den Wert 1. Um dies zu kompensieren, wird die Kontrolle vermutlich die größten Werte in einem Gebiet in der Nähe dieses Punktes annehmen. Da durch die Randbedingung in diesem Punkt die Kontrolle auf 0 gesetzt ist, liegt die Vermutung nahe, dass in einer Region um den Punkt  $(1, 1)$  die Kontrolle sehr große Ableitungen besitzt. Daher sollte das adaptive



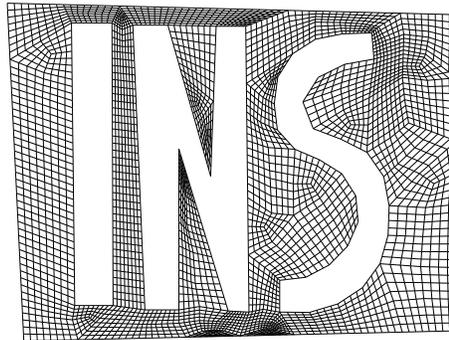
**Abbildung 7.5.** Von einem regulären Dünngitter auf Level 4 nach 4 (links oben), 7 (rechts oben), 10 (links unten) und 13 (rechts unten) Verfeinerungsschritten resultierendes Raum-Zeit Dünngitter für Beispiel 7.2.

Verfahren das Gitter zusätzlich in der Nähe des Punktes  $(1, 1)$  verfeinern.

Wir wählen für den adaptiven Zyklus zur Bestimmung der Schranken  $\varepsilon_y$  und  $\varepsilon_p$  in (7.17)  $c = 0.7$  und starten mit einem regulären Dünngitter auf Level  $l = 4$ . Abbildung 7.4 zeigt die Ergebnisse der nach 13 Verfeinerungsschritten resultierenden adaptiven Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung. Hierbei ist deutlich zu sehen, dass die Kontrolle in der Nähe des Punktes  $(1, 1)$  große Ableitungen besitzt. Die in Abbildung 7.5 dargestellten adaptiven Dünngitter sind, wie vermutet, besonders in der Nähe der Eckpunkte  $(0, 0)$  und  $(1, 1)$  fein aufgelöst. Ziehen wir hingegen nur die hierarchischen Koeffizienten des Zustands  $y$  zur Verfeinerung heran, wird das Dünngitter vornehmlich in Richtung des Eckpunktes  $(0, 0)$  verfeinert, während wir bei Verfeinerung nach den Koeffizienten der Adjungierten  $p$  ein zum Punkt  $(1, 1)$  feiner aufgelöstes Gitter erhalten. Dies ist in Abbildung 7.6 gut zu erkennen, in der die nach 10 adaptiven Verfeinerungsschritten resultierenden Dünngitter, bei denen jeweils nur die hierarchischen Koeffizienten von  $y_l$  bzw.  $p_l$  zur Verfeinerung verwendet wurden, dargestellt sind.



**Abbildung 7.6.** Von einem regulären Dünngitter auf Level 4 nach 10 Verfeinerungsschritten resultierendes adaptives Dünngitter bei Verwendung der hierarchischen Koeffizienten der Kontrolle (links) und des Zustands (rechts) als Fehlerindikator für Problem 7.2.



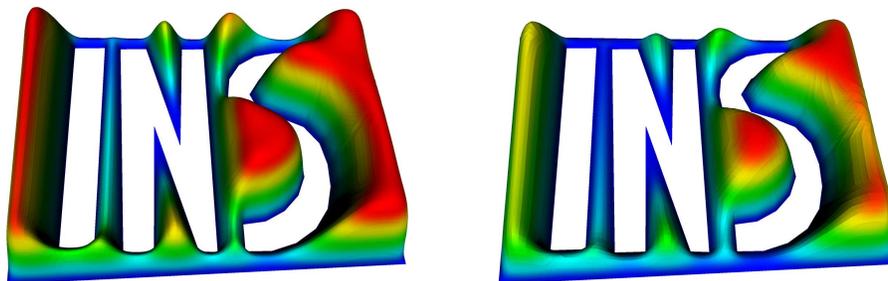
**Abbildung 7.7.** Das Gebiet  $\Omega$  aus Beispiel 7.3 mit einem Beispielgitter.

*Beispiel 7.3.* Wir betrachten das Problem

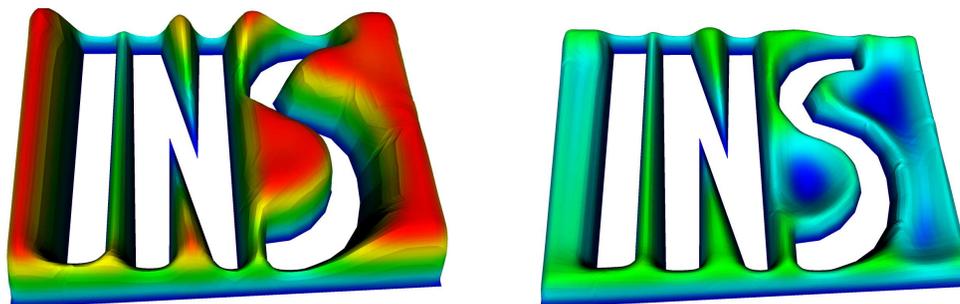
$$\begin{aligned} \partial_t y - \Delta y - 10^3 p &= 0 \text{ für alle } x \in \Omega, t \in (0, 1], \\ -\partial_t p - \Delta p + y &= 1 \text{ für alle } x \in \Omega, t \in [0, 1), \end{aligned}$$

mit homogenen Anfangs- bzw. Endzeit- und Randbedingungen auf dem Gebiet  $\Omega$ , das zusammen mit einem Beispielgitter in Abbildung 7.7 dargestellt ist. Wir erwarten durch die homogenen Randbedingungen an den einspringenden Ecken des Gebietes, das sowohl der Zustand als auch die Kontrolle an diesen Punkten besonders große Ableitungen besitzen. Das adaptive Verfahren sollte dementsprechend Gitter erzeugen, die zu den einspringenden Ecken feiner aufgelöst sind.

Abbildung 7.8 zeigt den durch das adaptive Crank-Nicolson Verfahren mit  $c = 0.5$  nach fünf Verfeinerungsschritten berechneten Zustand  $y$  zum Zeitpunkt  $t = 0.2$  und  $t = 1.0$ . Die dazugehörige Kontrolle ist in Abbildung 7.9 zu den Zeitpunkten  $t = 0$  und  $t = 0.8$  zu sehen. Das entsprechende adaptive Raum-Zeit Dünngitter ist



**Abbildung 7.8.** Der mit dem adaptiven Raum-Zeit Dünngitter Crank-Nicolson Verfahren nach fünf Verfeinerungsschritten berechnete Zustand  $y$  zu den Zeitpunkten  $t = 0.2$  (links) und  $t = 1.0$  (rechts).



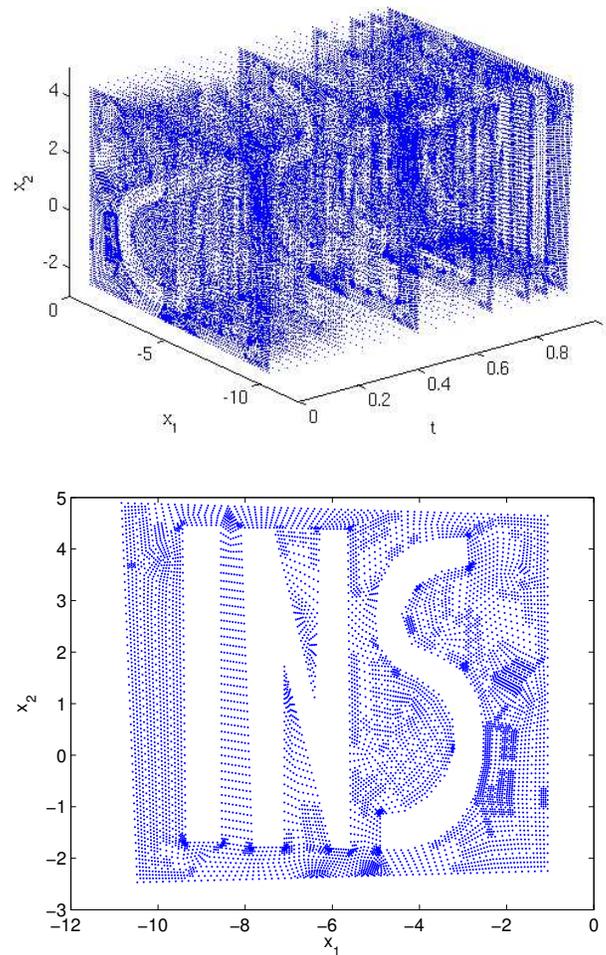
**Abbildung 7.9.** Die mit dem adaptiven Raum-Zeit Dünngitter Crank-Nicolson Verfahren nach fünf Verfeinerungsschritten berechnete Kontrolle  $u$  zu den Zeitpunkten  $t = 0$  (links) und  $t = 0.8$  (rechts).

in Abbildung 7.10 dargestellt. Wie oben diskutiert, ist das Gitter besonders an den einspringenden Ecken verfeinert

*Beispiel 7.4.* Zum Abschluß dieses Kapitels betrachten wir auf dem Gebiet  $\Omega = (-1, 1)^3 \setminus (-1, 0)^3$  die adaptive Diskretisierung des dreidimensionalen Sattelpunktproblems

$$\begin{aligned} \partial_t y - \Delta y - 10^3 p &= 0 \text{ für alle } x \in \Omega, t \in (0, 1], \\ -\partial_t p - \Delta p + y &= 1 \text{ für alle } x \in \Omega, t \in [0, 1), \end{aligned}$$

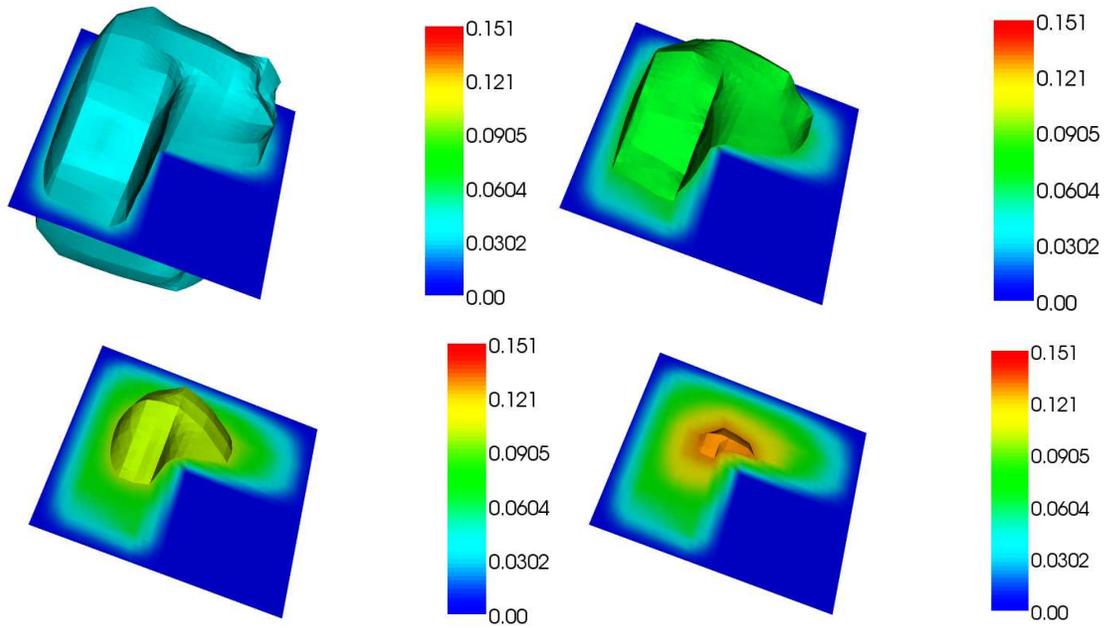
mit homogenen Rand- und Anfangs- bzw. Endzeitbedingungen. Durch die einspringende Ecke und die Nullrandbedingungen erwarten wir, dass sowohl die Kontrolle als auch der Zustand im Nullpunkt die benötigte Regularität nicht erfüllen. Man beachte, dass durch die hohe Komplexität klassischer Diskretisierungsverfahren die direkte Diskretisierung eines solchen Sattelpunktproblems in drei Ortsdimensionen in bisherigen Arbeiten noch nicht betrachtet wurde. Wir verwenden zur Lösung das adaptive Raum-Zeit Dünngitter Crank-Nicolson Verfahren ausgehend von Level  $l = 3$  mit vier Verfeinerungsschritten, wobei wir  $c = 0.5$  gewählt haben.



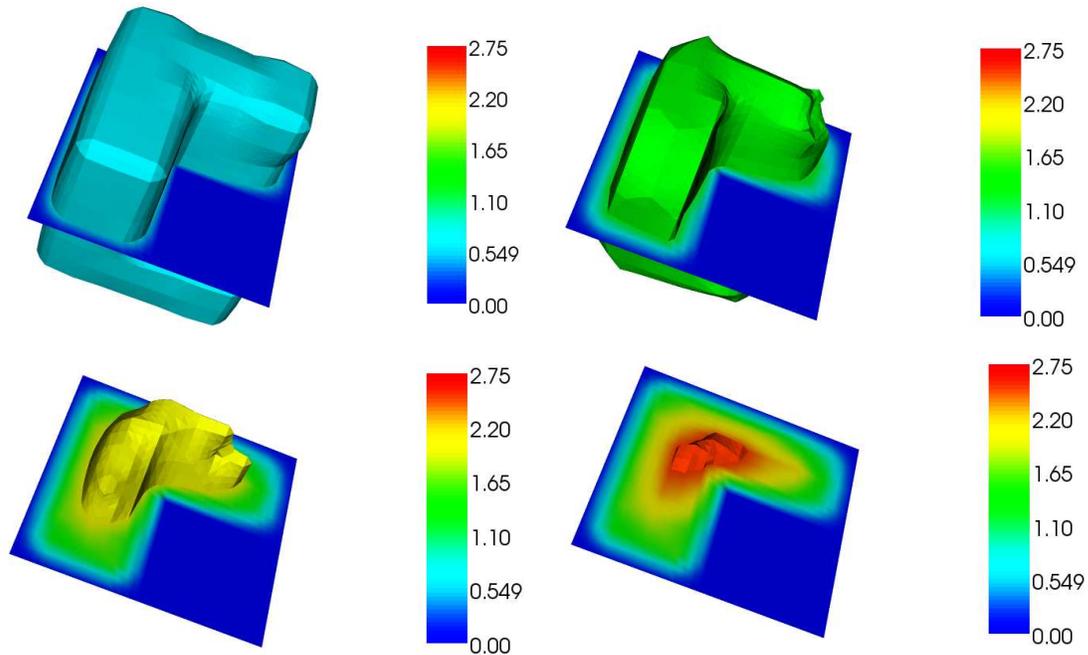
**Abbildung 7.10.** Nach fünf Verfeinerungsschritten resultierendes Raum-Zeit Dünngitter (oben) und Ausschnitt des Gitters zum Zeitpunkt  $t=0$  (unten) für Beispiel 7.3.

Abbildung 7.11 zeigt einige Isoflächen für den Zustand  $y$  zum Zeitpunkt  $t = 1.0$  und in Abbildung 7.12 sind Isoflächen der Kontrolle zum Zeitpunkt  $t = 0$  dargestellt.

In diesem Kapitel haben wir die Verallgemeinerung der Raum-Zeit Dünngitterdiskretisierungen für lineare skalare parabolische Probleme auf den Systemfall diskutiert. Hierbei haben wir uns auf die Diskretisierung der Sattelpunktsformulierung von instationären verteilten Kontrollproblemen als Beispiel für ein System parabolischer Gleichungen konzentriert. Durch die spezielle Struktur der Sattelpunktsprobleme, bei denen miteinander gekoppelte Vorwärts- und Rückwärtsgleichungen auftreten, zerfallen klassische Diskretisierungsverfahren nicht mehr in Zeitschritt-



**Abbildung 7.11.** *Isoflächen für den mittels des adaptiven Raum-Zeit Dünngitter Crank-Nicolson Verfahrens berechneten Zustand  $y$  des Beispiels 7.4.*



**Abbildung 7.12.** *Isoflächen für die mittels des adaptiven Raum-Zeit Dünngitter Crank-Nicolson Verfahrens berechnete Kontrolle  $u$  des Beispiels 7.4.*

verfahren, so dass alle Unbekannten zu jedem Zeitschritt im Hauptspeicher gehalten werden müssen. Aufgrund der hohen Komplexität voller Gitter, sind deshalb bisher Ergebnisse klassischer Diskretisierungsverfahren für solche Sattelpunktsprobleme nur für Probleme in zwei Ortsdimensionen vorhanden. Im Gegensatz dazu waren wir in der Lage, mit Raum-Zeit Dünngittern auch den dreidimensionalen Fall zu behandeln. Die numerischen Ergebnisse zeigen dabei, dass sowohl die Konvergenzraten der Diskretisierung als auch das Konvergenzverhalten der entsprechenden Multilevellöser für die linearen Gleichungssysteme vergleichbar mit den Ergebnissen im skalaren Fall sind. Schließlich haben wir auf Grundlage des in Abschnitt 6.1 entwickelten Fehlerindikators Adaptivität für die instationären verteilten Kontrollprobleme an dem Beispiel der Raum-Zeit Dünngitter Crank-Nicolson Diskretisierung untersucht. Die präsentierten adaptiven Ergebnisse sind viel versprechend und motivieren eine eingehendere, ausführlichere Forschung in diesem Bereich.

## Kapitel 8

# Zusammenfassung und Ausblick

In dieser Arbeit haben wir die Diskretisierung linearer parabolischer Probleme mit so genannten Raum-Zeit Dünngittern untersucht. Die Raum-Zeit Dünngitter sind eine Verallgemeinerung klassischer Dünngitter und basieren auf der natürlichen Produktstruktur zwischen Raum und Zeit.

Die theoretischen und numerischen Ergebnisse dieser Arbeit zeigen, dass diese Räume unter etwas stärkeren Glattheitsannahmen als bei vollen Gittern zu den gleichen Approximationsraten führen. Gleichzeitig ist die Dimension der Raum-Zeit Dünngitterräume (bis auf einen logarithmischen Faktor) um eine Ordnung geringer als die Dimension der Vollgitterräume. Aufgrund dessen ist ein wesentlich besseres Kosten-Nutzen Verhältnis erzielbar. Die Betrachtung der klassischen Regularitätstheorie für lineare parabolische Gleichungen hat in dieser Arbeit gezeigt, dass die benötigten Glattheitsvoraussetzungen keinesfalls rein akademisch gewählt sind, um die Theorie zu ermöglichen, sondern tatsächlich auch bei parabolischen Problemen unter sinnvollen Annahmen an das Gebiet, die Glattheit der Koeffizienten des Operators, die Randbedingungen und die rechte Seite erfüllt sind.

Ferner haben wir in dieser Arbeit die Crank-Nicolson und die Discontinuous-Galerkin Diskretisierung auf vollen Gittern auf die Diskretisierung mit den Raum-Zeit Dünngittern übertragen. Eine Analyse des Discontinuous-Galerkin Verfahrens mit Raum-Zeit Dünngittern führte zu einer a priori Fehlerabschätzung, die den Diskretisierungsfehler gegen entsprechende Interpolationsfehler abschätzt. Numerische Experimente haben schließlich die Effizienz der Raum-Zeit Dünngitterdiskretisierungen in der Praxis bestätigt.

Ein wesentlicher Aspekt für die Effizienz des gesamten Diskretisierungsprozesses ist die Effizienz des iterativen Löser für die anfallenden linearen Gleichungssysteme. In diesem Zusammenhang sind Multilevel- bzw. Mehrgittermethoden von besonderer Bedeutsamkeit, da deren Konvergenzgeschwindigkeit in vielen Fällen unabhängig von der Feinheit der Diskretisierung ist. Aus diesem Grund gelten Multilevellöser auch als optimal. Nach [61] können eine Vielzahl von Multilevellösern als klassische

Iterationsverfahren über dem Erzeugendensystem aufgefasst werden. Beruhend auf dem (Block-) Gauss-Seidel Verfahren über dem Erzeugendensystem wurden in dieser Arbeit dementsprechend zwei verschiedene Multilevellöser in Raum-Zeit vorgestellt und numerisch untersucht. Dabei haben wir durch die Kombination der beiden Löser ein Verfahren erhalten, dessen Konvergenzrate in den Experimenten unabhängig von der gewählten Feinheit der Diskretisierung und in diesem Sinne optimal ist.

Um Probleme, deren Lösung nicht den für die Approximationsraten nötigen Glattheitsvoraussetzungen genügt, effizient behandeln zu können, haben wir adaptive Raum-Zeit Dünngitter betrachtet. Hierbei haben wir, analog zu klassischen Dünngittern, einen auf den hierarchischen Koeffizienten der diskreten Lösung basierenden Fehlerindikator entwickelt. Numerische Experimente zeigen, dass das resultierende adaptive Verfahren auch im Fall von singulären Lösungen zu einem guten Kosten-Nutzen Verhältnis führt. Besonders bemerkenswert ist dabei, dass sich bei dieser Vorgehensweise das so genannte local time stepping, d.h. verschiedene Ortsgebiete werden mit unterschiedlichen Zeitschrittweiten diskretisiert, automatisch ergibt. Im Gegensatz dazu ist im Fall von Diskretisierungen auf adaptiven vollen Gittern die Behandlung des local time stepping mit einem erheblichen Mehraufwand in der Implementierung verbunden.

Bei einer Vielzahl von Problemstellungen steht nicht die Lösung selbst im Vordergrund, sondern lediglich eine gewisse gemittelte Größe der Lösung, wie zum Beispiel der Wert eines linearen Funktionals ausgewertet an der Lösung. Aufgrund dessen haben wir zusätzlich einen Fehlerindikator bzgl. linearer Funktionale entwickelt, der auf der diskreten Lösung des ursprünglichen und der diskreten Lösung des dazu dualen Problems basiert. Dabei sind die mit diesem Fehlerindikator in dieser Arbeit erzielten numerischen Ergebnisse recht viel versprechend.

Schließlich haben wir die Diskretisierung von Systemen parabolischer Gleichungen mit den Raum-Zeit Dünngittern diskutiert. Hierbei haben wir als besonders interessantes Beispiel instationäre verteilte Kontrollprobleme untersucht, bei deren Sattelpunktsformulierung die Gleichung, die den Zustand bestimmt, vorwärts und die Gleichung, die die Kontrolle bestimmt, rückwärts in der Zeit läuft. Es konnte gezeigt werden, dass sich die Diskretisierungen für skalare Probleme ohne große Schwierigkeit auf diesen Fall übertragen lassen. Während bisherige Arbeiten zu klassischen Diskretisierungsverfahren für solche Sattelpunktsprobleme auf Grund der großen Zahl an benötigten Freiheitsgraden bei vollen Gittern lediglich den zweidimensionalen Fall behandelt haben, konnten wir erstmalig diese Probleme auch in drei Ortsdimensionen lösen.

Insgesamt ist es damit in dieser Arbeit gelungen, neue, hocheffiziente Diskretisierungsverfahren zu entwickeln, die es sogar ermöglichen, lineare parabolische Probleme mit der gleichen Kosten-Nutzen Relation, d.h. Relation von Freiheitsgraden zu erzieltm Fehler, wie die Diskretisierung rein stationärer, elliptischer Probleme mit Finiten Elementen zu lösen. Hierbei werden mit den Raum-Zeit Dünngittern viele

Nachteile dünner Gitter durch die explizite Nutzung der natürlichen Produktstruktur zwischen Raum und Zeit innerhalb der Konstruktion vermieden. So können lineare parabolische Probleme mit im Ort beliebig variierenden Koeffizientenfunktionen behandelt werden. Auch die benötigten Glattheitsvoraussetzungen der Raum-Zeit Dünngitter sind bedeutend schwächer als bei klassischen Dünngittern und werden auch bei herkömmlichen Zeitschrittverfahren benötigt. Somit stellen die Glattheitsanforderungen im Vergleich zu herkömmlichen Verfahren keine Einschränkung dar. Des Weiteren benötigt man zur Konstruktion lediglich eine Sequenz von ineinander geschachtelten Finite Element Räumen im Ort, die von vielen Finite Element Paketen, die Mehrgitterverfahren zur Lösung beinhalten, bereitgestellt wird. Somit lassen sich wesentlich einfacher auch Probleme auf komplizierteren Geometrien als im klassischen Dünngitterfall behandeln.

Für die Anwendung auf praxisrelevante Probleme ergeben sich einige weitere interessante Fragestellungen, die in dieser Arbeit nicht behandelt wurden. So haben wir in unserer Raum-Zeit Dünngitterkonstruktion die aus einer Sequenz ineinander geschachtelter Finite Element Räume konstruierte hierarchische Basis als Multilevelzerlegung des Ortes betrachtet. Einige Finite Element Pakete bieten geometrische Mehrgitterverfahren zur Lösung der anfallenden linearen Gleichungssysteme an, so dass diese Pakete auf die eine oder andere Weise einen Zugriff auf die entsprechenden Gitterhierarchien ermöglichen und die in dieser Arbeit untersuchte Konstruktion angewendet werden kann. Allerdings besteht im Fall sehr komplizierter Geometrien das Problem, dass selbst auf dem größten Level noch sehr viele Freiheitsgrade benötigt werden, um die Geometrie aufzulösen. In diesen Fällen besitzt der Raum-Zeit Dünngitterraum eine dementsprechend große Dimension. Um die Dimension weiter zu verringern, sollte die Zahl der Freiheitsgrade auf dem größten Ortsgitter weiter reduziert werden. Hierzu wurde in [64] vorgeschlagen, die groben Probleme weiter mit Hilfe von Vegröberungsalgorithmen, die aus algebraischen Mehrgittermethoden stammen, zu vergrößern, wobei die benötigten Interpolationsoperatoren über einen least-square Ansatz bestimmt werden, so dass die ( $d$ -) linearen Funktionen möglichst gut approximiert werden. Eine weitere Möglichkeit zur Bestimmung der Gittervergrößerung und der Interpolationsoperatoren, bei denen fest vorgegebene Vektoren exakt interpoliert werden, findet man in [113] und sollte Gegenstand zukünftiger Forschung sein.

In dieser Arbeit haben wir uns auf die Behandlung linearer Probleme beschränkt. In vielen Anwendungen auftretende Probleme sind jedoch häufig nichtlinearer Natur. Durch die von den Raum-Zeit Dünngittern induzierte gleichzeitige Diskretisierung von Raum und Zeit ist es jedoch nicht möglich, die Nichtlinearität wie bei Zeitschrittverfahren üblich, entweder auf den Ort zu übertragen oder gar in dem Lösungsprozess durch eine explizite Zeitdiskretisierung ganz zu eliminieren. Stattdessen muß die Nichtlinearität in dem gesamten Raum-Zeit Zylinder behandelt werden, wozu möglicherweise Fixpunktiterationen verwendet werden könnten.

Obwohl die Raum-Zeit Dünngitter gegenüber klassischen Diskretisierungen den benötigten Rechen- und Speicheraufwand drastisch reduzieren, sind der Behandlung komplexer Probleme durch die Rechenleistung aktuell verfügbarer Computer technische Grenzen gesetzt. Aus diesem Grund ist es auch im Fall der in dieser Arbeit vorgestellten Diskretisierungsverfahren von Bedeutung, möglichst effiziente Parallelisierungsstrategien zu entwickeln. Hierbei ist insbesondere die Lastbalancierung der Daten interessant, da wir nicht, wie bei herkömmlichen Zeitschrittverfahren, nur das Ortsproblem durch Gebietszerlegungsmethoden auf die Prozessoren aufzuteilen haben, sondern zusätzlich noch die Zeitkomponente berücksichtigen müssen. Während für die bei Zeitschrittverfahren anfallenden Ortsprobleme Lösungen für das Lastbalancierungsproblem mittels raumfüllender Kurven [67, 68] oder paralleler Graphpartitionierungsalgorithmen [77, 78] existieren, und für die parallele Diskretisierung von gewöhnlichen Differentialgleichungen parareal-Algorithmen [4, 85, 86] verwendet werden können, benötigen wir für die Raum-Zeit Dünngitter völlig neuartige Ansätze. Vorstellbar ist in diesem Zusammenhang eine Kombination von Gebietszerlegungsmethoden im Ort und parareal-Algorithmen in der Zeit. Zukünftige Forschungsansätze, die diesen Kontext berücksichtigen, erscheinen aus unserer Sicht viel versprechend.

# Kapitel 9

## Anhang

### 9.1 Beweis von Satz 5.1

Wir beweisen nun Satz 5.1. Nach den Betrachtungen in 5.1 reicht es dafür, für die Iterationsmatrix des SOR Verfahrens

$$\tilde{\rho}(I - \omega(D^E + \omega U^E)^{-1}L^E) < 1$$

zu zeigen. Hierbei ist  $\tilde{\rho}(A)$  der verallgemeinerte Spektralradius (5.7) einer Matrix  $A$  ist. Wir schreiben  $\omega = 1 + d$  und schätzen im folgenden  $d > -1$  ab. Sei  $\lambda$  ein Eigenwert von  $I - \omega(D^E + \omega U^E)^{-1}L^E$  und  $x$  ein dazugehöriger normierter Eigenvektor,  $x^*x = 1$  ( $x^* = (\bar{x})^T$ ). Da wir in dem verallgemeinerten Spektralradius nur Eigenwerte ungleich eins betrachten, nehmen wir ohne Einschränkung  $\lambda \neq 1$  an. Da  $\lambda \neq 1$  ist insbesondere  $x \notin \ker(L^E)$  und damit  $P^\perp x \neq 0$ . Nach der Voraussetzung des Satzes gilt  $D^E = I$  woraus

$$\lambda(I + (1 + d)U^E)x = -((1 + d)R^E + dI)x$$

folgt. Damit erhalten wir

$$\begin{aligned} x^* ((1 - d)I - (1 + d)(L^E - U^E + R^E))x &= 2x^*(-dI - (1 + d)R^E)x \\ &= 2\lambda x^*(I + (1 + d)U^E)x \\ &= \lambda x^* ((1 - d)I + (1 + d)(L^E + \\ &\quad U^E - R^E))x. \end{aligned}$$

Da wir  $d > -1$  voraussetzen, können wir obige Gleichung durch  $\omega = (1 + d)$  teilen und erhalten mit den Abkürzungen  $c := (1 - d)/(1 + d)$ ,  $b := x^*Bx$  und  $\tilde{b} := x^*\tilde{B}x$

$$\lambda = \frac{c - (1 + b) + \tilde{b}}{c + (1 + b) + \tilde{b}}. \quad (9.1)$$

Da  $Px \in \ker(L^E)$  gilt

$$0 = (P^\perp x)^* L^E (Px) = (P^\perp x)^* (I + B)(Px) = (P^\perp x)^* B(Px). \quad (9.2)$$

Da nach Voraussetzung  $\ker(L^E) = \ker((L^E)^T)$  folgt genauso  $(Px)^* B(P^\perp x) = 0$ .  
Damit erhalten wir

$$\begin{aligned} b &= x^* Bx \\ &= (Px + P^\perp x)^* B(Px + P^\perp x) \\ &= (Px)^* BPx + (P^\perp x)^* BP^\perp x. \end{aligned}$$

Zusätzlich ist  $0 = (Px)^* L^E Px = (Px)^* Px + (Px)^* BPx$  und deshalb  $(Px)^* BPx = -(Px)^* Px$  und es folgt

$$\begin{aligned} 1 + b &= (Px)^*(Px) + (P^\perp x)^* P^\perp x + (Px)^* BPx + (P^\perp x)^* BP^\perp x \\ &= (P^\perp x)^* P^\perp x + (P^\perp x)^* BP^\perp x = (1 + \hat{b}) \|P^\perp x\|^2, \end{aligned}$$

wobei  $\hat{b} := (P^\perp x)^* BP^\perp x / \|P^\perp x\|^2$ . Wir setzen

$$\hat{b} := \hat{g} + i\hat{s}, \quad \tilde{b} := \tilde{g} + i\tilde{s},$$

mit reellen Zahlen  $g, \tilde{g}, s, \tilde{s}$ . Bezeichnen wir mit  $z$  den Zähler von (9.1) und mit  $x$  den Nenner, so erhalten wir  $|\lambda| < 1$  wenn

$$\bar{x}x - \bar{z}z > 0.$$

Es gilt

$$\bar{x}x - \bar{z}z = 4\gamma((c + \tilde{g})(1 + \hat{g}) + \hat{s}\tilde{s}), \quad \text{mit } \gamma := \|P^\perp x\|^2.$$

Da nach Annahme  $x \notin \ker(L^E)$  und deshalb insbesondere  $\gamma > 0$  ist, ist

$$((c + \tilde{g})(1 + \hat{g}) + \hat{s}\tilde{s}) > 0 \quad (9.3)$$

eine hinreichende Bedingung für  $|\lambda| < 1$ . Es gilt

$$|\tilde{s}| \leq \tilde{\sigma}, \quad \tilde{g}_1 \leq \tilde{g}, \quad |\hat{s}| \leq \hat{\sigma}, \quad \text{und } \hat{g}_1 \leq \hat{g}.$$

Da außerdem  $I + \hat{G}$  positiv definit auf dem orthogonalen Komplement des Kerns von  $L^E$  ist, muß  $\hat{g}_1 > -1$  sein und es ist  $1 + \hat{g} > 1 + \hat{g}_1 > 0$ . Wir formen nun (9.3) weiter um und erhalten

$$c(1 + \hat{g}) > -\hat{s}\tilde{s} - \tilde{g}(1 + \hat{g}).$$

Division durch  $(1 + \hat{g})$  liefert

$$c > -\frac{\hat{s}\tilde{s}}{1 + \hat{g}} - \tilde{g}.$$

Wir schätzen die rechte Seite der Ungleichung weiter ab durch

$$\frac{\hat{s}\tilde{s}}{1+\hat{g}} - \tilde{g} \leq \frac{\hat{\sigma}\tilde{\sigma}}{1+\hat{g}_1} - \tilde{g}_1.$$

Damit ist (9.3) auf jeden Fall erfüllt, wenn

$$c > \frac{\hat{\sigma}\tilde{\sigma}}{1+\hat{g}_1} - \tilde{g}_1.$$

Verwenden wir nun  $c = (1-d)/(1+d) = -1 + 2/(1+d)$ , so folgt die Behauptung.

□



# Literaturverzeichnis

- [1] S. ACHATZ, *Adaptive finite Dünngitter-Elemente höherer Ordnung für elliptische partielle Differentialgleichungen mit variablen Koeffizienten*, Dissertation, TU München, 2003.
- [2] S. ADJERID, J. FLAHERTY, AND Y. WANG, *Adaptive method-of-lines techniques for parabolic systems*, tech. rep., Rensselaer Polytechnic Institute, 1993.
- [3] I. BABUSKA AND W. RHEINBOLDT, *Error estimates for adaptive finite elements computations*, SIAM J. Numer. Anal., 15 (1978), pp. 736–754.
- [4] G. BAL, *On the convergence and the stability of the parareal algorithm to solve partial differential equations*, in Proceedings of DD15, LNCSE, Springer, 2004, pp. 425–432.
- [5] R. BALDER, *Adaptive Verfahren für elliptische und parabolische Differentialgleichungen*, Dissertation, Technische Universität München, 1994.
- [6] R. BALDER AND C. ZENGER, *The solution of the multidimensional real helmholtz equations on sparse grids*, SIAM J. Sci. Comp., 17 (1996), pp. 631–646.
- [7] W. BANGERTH, R. HARTMANN, AND G. KANSCHAT, *deal.II Differential Equations Analysis Library, Technical Reference*. <http://www.dealii.org>.
- [8] W. BANGERTH AND R. RANNACHER, *Adaptive Finite Element Methods for Differential Equations*, Birkhäuser, 2003.
- [9] R. BARRET, M. BERRY, T. CHAN, J. DEMMEL, J. DONATO, J. DONGARRA, V. EIJKHOUT, R. POZO, C. ROMINE, AND H. VAN DER VORST, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, 1994.
- [10] G. BASZENSKI AND F. DELVOS, *Multivariate boolean midpoint rules*, in Numerical Integration IV, H. Brass and G. Hämmerlin, eds., vol. 112 of ISNM, Birkhäuser, 1993, pp. 1–11.

- 
- [11] R. BECKER, *Adaptive finite elements for optimal control problems*, Habilitation, Fakultät für Mathematik, Ruprecht-Karls-Universität Heidelberg, 2001.
- [12] R. BECKER, H. KAPP, AND R. RANNACHER, *Adaptive finite element methods for optimal control of partial differential equations: Basic concepts*, SIAM J. Control Optim., 39 (2000), pp. 113–132.
- [13] R. BECKER AND R. RANNACHER, *A feed-back approach to error control in finite element methods: Basic analysis and examples*, East-West J. Numer. Anal., 37 (1996), pp. 209–225.
- [14] R. BECKER AND B. VEXLER, *A posteriori error estimation for finite element discretization of parameter identification problems*, Numer. Math., 96 (2004), pp. 435–459.
- [15] R. BELLMANN, *Adaptive Control Processes: A Guided Tour*, Princeton University Press, 1961.
- [16] M. BIETERMAN AND I. BABUSKA, *The finite element method for parabolic equations. I. A posteriori error estimation*, Numer. Math., 40 (1982), pp. 339–371.
- [17] —, *An adaptive method of lines with error control for parabolic problems*, J. Comp. Phys., 63 (1986), pp. 33–66.
- [18] T. BONK, *Ein rekursiver Algorithmus zur adaptiven numerischen Quadratur mehrdimensionaler Funktionen*, Dissertation, Institut für Informatik, Technische Universität München, 1994.
- [19] F. BORNEMANN, *An adaptive multilevel approach to parabolic problems. I. General theory and 1d implementation*, Impact Comp. Sci. Engrg., 2 (1990), pp. 279–317.
- [20] —, *An Adaptive Multilevel Approach for Parabolic Equations in Two Space Dimensions*, Dissertation, FU Berlin, 1991.
- [21] —, *An adaptive multilevel approach to parabolic problems. II. Variable-order time discretization based on a multiplicative error correction*, Impact Comp. Sci. Engrg., 3 (1991), pp. 93–122.
- [22] A. BORZI, *Fast multigrid methods for parabolic distributed optimal control problems*, in Proceedings of the 18<sup>th</sup> GAMM-Seminar Leipzig, 2002, pp. 1–10.
- [23] —, *Multigrid Methods for Optimality Systems*, Habilitation, Universität Graz, 2003.

- 
- [24] —, *Multigrid methods for parabolic distributed optimal control problems*, J. Comp. Appl. Math., 157 (2003), pp. 365–382.
- [25] D. BRAESS, *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*, Cambridge University Press, 2001.
- [26] A. BRANDT, *Algebraic Multigrid Theory: The Symmetric Case*, in Preliminary Proceedings for the International Multigrid Conference, Copper Mountain, Colorado, April 1983.
- [27] A. BRANDT, *Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics*, Tech. Rep. GMD-Studien Nr. 85, GMD - Forschungszentrum Informationstechnik GmbH, 1984.
- [28] A. BRANDT AND J. GREENWALD, *Parabolic multigrid revisited*, in Multigrid Methods III, W. Hackbusch and U. Trottenberg, eds., 1991, pp. 143–154.
- [29] H.-J. BUNGARTZ, *An adaptive poisson solver using hierarchical bases and sparse grids*, in Iterative Methods in Linear Algebra, Elsevier, 1992, pp. 293–310.
- [30] H.-J. BUNGARTZ, *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*, Dissertation, Technische Universität München, 1992.
- [31] —, *A multigrid algorithm for higher order finite elements on sparse grids*, ETNA, Electron. Trans. Numer. Anal., 6 (1997), pp. 63–77.
- [32] H.-J. BUNGARTZ, *Finite Elements of Higher Order on Sparse Grids*, Shaker Verlag, 1998.
- [33] H.-J. BUNGARTZ AND T. DORNSEIFER, *Sparse grids: Recent developments for elliptic partial differential equations*, in Multigrid Methods V, W. e. a. Hackbusch, ed., vol. 3 of Lecture Notes in Computational Science and Engineering, Springer, 1998, pp. 45–70.
- [34] H.-J. BUNGARTZ, T. DORNSEIFER, AND C. ZENGER, *Tensor product approximation spaces for the efficient numerical solution of partial differential equations*, in Proc. Int. Workshop on Scientific Computations, Nova Science Publishers, Konya, 2996.
- [35] H.-J. BUNGARTZ AND M. GRIEBEL, *Sparse grids*, Acta Numerica, 13 (2004), pp. 1–121.
- [36] J. BURIE AND M. MARION, *Multilevel methods in space and time for the navier stokes equation*, SIAM J. Numer. Anal., 34 (1997), pp. 1574–1599.

- [37] —, *Adaptive multilevel methods in space and time for parabolic problems - the periodic case*, Math. Comp., 69 (2000), pp. 547–581.
- [38] C. CHUI AND Y. WANG, *A general framework for compactly supported splines and wavelets*, J. Approx. Theory, 71 (1992), pp. 263–304.
- [39] P. CIARLET, *The Finite Element Method for Elliptic Problems*, Classics in Applied Mathematics, SIAM, 2002.
- [40] A. COHEN, *Numerical Analysis of Wavelet Methods*, Studies in Mathematics and its Applications, Vol. 32, North Holland, 2003.
- [41] W. DAHMEN, *Wavelet and multiscale methods for operator equations*, Acta Numerica, (1997), pp. 55–228.
- [42] W. DAHMEN AND R. SCHNEIDER, *Wavelets on manifolds I: Construction and domain decomposition*, SIAM J. Math. Anal., 31 (1999), pp. 184–230.
- [43] W. DAHMEN AND R. STEVENSON, *Element-by-element construction of wavelets satisfying stability and moment conditions*, SIAM J. Numer. Anal., 37 (1999), pp. 319–352.
- [44] I. DAUBECHIES, *Orthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 41 (1988), pp. 909–996.
- [45] J. DENDY, *Blackbox multigrid for nonsymmetric problems*, Appl. Math Comp., (1983), pp. 261–283.
- [46] T. DORNSEIFER, *Diskretisierung allgemeiner elliptischer Differentialgleichungen in krummlinigen Koordinatensystemen auf dünnen Gittern*, Dissertation, Technische Universität München, 1997.
- [47] T. DORNSEIFER AND C. PFLAUM, *Discretization of elliptic differential equations on curvilinear bounded domains with sparse grids*, Computing, 56 (1996), pp. 197–213.
- [48] K. ERIKSSON, D. ESTEP, P. HANSBO, AND C. JOHNSON, *Computational Differential Equations*, Cambridge University Press, 1996.
- [49] K. ERIKSSON AND C. JOHNSON, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.
- [50] —, *Adaptive finite element methods for parabolic problems II: Optimal error estimates in  $l_\infty l_2$  and  $l_\infty l_\infty$* , SIAM J. Numer. Anal., 32 (1995), pp. 706–740.

- [51] —, *Adaptive finite element methods for parabolic problems IV: Nonlinear problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1729–1749.
- [52] —, *Adaptive finite element methods for parabolic problems V: Long-time integration*, SIAM J. Numer. Anal., 32 (1995), pp. 1750–1763.
- [53] K. ERIKSSON, C. JOHNSON, AND S. LARSSON, *Adaptive finite element methods for parabolic problems VI: Analytic semigroups*, SIAM J. Numer. Anal., 35 (1998), pp. 1315–1325.
- [54] D. ESTEP AND D. FRENCH, *Global error control for the continuous Galerkin finite element method for ordinary differential equations*, Math. Anal. Numer., 28 (1994), pp. 815–852.
- [55] G. FABER, *Über stetige Funktionen*, Mathematische Annalen, (1909), pp. 81–94.
- [56] C. FEUERSÄNGER, *Dünngitterverfahren für hochdimensionale elliptische partielle Differentialgleichungen*, Diplomarbeit, Institut für Numerische Simulation, Universität Bonn, 2005.
- [57] J. GARCKE AND M. GRIEBEL, *On the computation of the eigenproblems of hydrogen and helium in strong magnetic and electric fields with the sparse grid combination technique*, Journal of Computational Physics, 165 (2000), pp. 694–716.
- [58] T. GERSTNER AND M. GRIEBEL, *Numerical integration using sparse grids*, Numerical Algorithms, 18 (1998), pp. 209–232.
- [59] M. GILES, M. LARSON, M. LEVENSTAM, AND E. SÜLI, *Adaptive error control for finite element approximations of the lift and drag in a viscous flow*, tech. rep., Oxford University Computing Laboratory, 1997.
- [60] M. GRIEBEL, *Parallel multigrid methods on sparse grids*, in Multigrid Methods III, International Series of Numerical Mathematics, vol. 98, Birkhäuser Verlag, Basel, 1991, pp. 211–221. also as SFB Bericht, 342/30/90 A, Institut für Informatik, TU München, 1990.
- [61] M. GRIEBEL, *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen*, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.
- [62] —, *Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences*, Computing, 61 (1998), pp. 151–179.
- [63] M. GRIEBEL AND S. KNAPEK, *Optimized tensor-product approximation spaces*, Constructive Approximation, 16 (2000), pp. 525–540.

- [64] M. GRIEBEL, D. OELTZ, AND P. VASSILEVSKI, *Space-time approximation with sparse grids*. eingereicht bei SIAM J. Sci. Comp.
- [65] M. GRIEBEL AND P. OSWALD, *Tensor product type subspace splitting and multilevel iterative methods for anisotropic problems*, Adv. Comput. Math., 4 (1995), pp. 171–206. also as SFB-Bericht 342/15/94A, Institut für Informatik, TU München, 1994.
- [66] M. GRIEBEL, P. OSWALD, AND T. SCHIEKOFER, *Sparse grids for boundary integral equations*, Numer. Mathematik, 83 (1999), pp. 279–312. also as SFB 256 report 554, Universität Bonn.
- [67] M. GRIEBEL AND G. ZUMBUSCH, *Hash based adaptive parallel multilevel methods with space-filling curves*, in NIC Symposium 2001, H. Rollnik and D. Wolf, eds., vol. 9 of NIC Series, ISBN 3-00-009055-X, Germany, 2002, Forschungszentrum Jülich, pp. 479–492.
- [68] M. GRIEBEL AND G. W. ZUMBUSCH, *Parallel multigrid in an adaptive PDE solver based on hashing and space-filling curves*, Parallel Computing, 25 (1999), pp. 827–843.
- [69] A. GRIEWANK, *Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation*, Optim. Methods Softw., 1 (1992), pp. 35–54.
- [70] C. GROSSMANN AND H.-G. ROOS, *Numerik partieller Differentialgleichungen*, Teubner Verlag, 1992.
- [71] W. HACKBUSCH, *Parabolic multigrid methods*, in Computing Methods in Applied Sciences and Engineering VI, R. Glowinski and J.-L. Lions, eds., Elsevier Science Publishers B.V., 1984.
- [72] ———, *Multi-Grid Methods and Applications*, vol. 4 of Springer Series in Computational Mathematics, Springer, 1985.
- [73] ———, *The efficient computation of certain determinants arising in the treatment of schrödinger's equation*, Computing, 67 (2001), pp. 35–56.
- [74] M. HINZE AND K. KUNISCH, *Three control methods for time-dependent fluid flow*, Journal of Flow, Control and Combustion, 65 (2000), pp. 273–298.
- [75] R. HIPTMAIR AND C. SCHWAB, *Numerical Treatment of Partial Differential Equations*. Lecture notes for a course held by Ralf Hiptmair.

- [76] M. HOLST AND F. SAIED, *Parallel performance of some multigrid solvers for three-dimensional parabolic equations*, tech. rep., Numerical Computing Group, University of Illinois at Urbana-Champaign, 1991.
- [77] G. KARYPIS AND V. KUMAR, *A fast and high quality multilevel scheme for partitioning irregular graphs*, SIAM J. Sci. Comp., 20 (1989), pp. 359–392.
- [78] K. KARYPIS AND V. KUMAR, *Multilevel  $k$ -way partitioning scheme for irregular graphs*, Journal of Parallel and Distributed Computing, 48 (1998), pp. 96–129. to appear.
- [79] S. KNAPEK, *Approximation und Kompression mit Tensorprodukt-Multiskalen-Approximationsäumen*, Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, 2000.
- [80] F. KOSTER, *Multiskalen-basierte Finite-Differenzen-Verfahren auf adaptiven dünnen Gittern*, Dissertation, Rheinische Friedrich-Hilhelms-Universität Bonn, 2001.
- [81] K. KUNISCH AND S. VOLKWEIN, *Galerkin proper orthogonal decomposition methods for parabolic systems*, Numer. Math., 90 (2001), pp. 117–148.
- [82] O. LADYZENSKAJA, V. SOLONNIKOV, AND N. URAL'CEVA, *Linear and Quasilinear Equations of Parabolic Type*, vol. 23 of Translations of Mathematical Monographs, American Mathematical Society, 1988.
- [83] J. LANG, *Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems*, Springer, 2001.
- [84] J. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, 1971.
- [85] J. LIONS, Y. MADAY, AND G. TURINICI, *A "parareal" in time discretization of pde's*, Comptes Rendus de l'Academie des Sciences Series I Mathematics, 332 (2001), pp. 661–668.
- [86] Y. MADAY AND G. TURINICI, *A parareal in time procedure for the control of partial differential equations*, Comptes Rendus Mathematique, 335 (2002), pp. 387–392.
- [87] H. NGUYEN AND R. STEVENSON, *Finite-element wavelets on manifolds*, IMA J. of Numer. Anal., 23 (2003), pp. 149–173.
- [88] W. NIETHAMMER, *Relaxation bei nichtsymmetrischen Matrizen*, Math. Zeitschr., 85 (1964), pp. 319–327.

- [89] T. ODEN AND S. PRUDHOMME, *Goal-oriented error estimation and adaptivity for the finite element method*, Computers and Mathematics with Applications, 41 (2001), pp. 735–756.
- [90] S. PASKOV, *Average case complexity of multivariate integration for smooth functions*, J. Complexity, 9 (1993).
- [91] T. PETERSDORFF AND C. SCHWAB, *Numerical solution of parabolic equations in high dimensions*, Mathematical Modelling and Numerical Analysis, 38 (2004), pp. 93–128.
- [92] C. PFLAUM, *Anwendung von Mehrgitterverfahren auf dünnen Gittern*, Diplomarbeit, Technische Universität München, 1992.
- [93] —, *Diskretisierung elliptischer Differentialgleichungen mit dünnen Gittern*, Dissertation, Technische Universität München, 1996.
- [94] C. REISINGER, *Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben*, Dissertation, Ruprecht-Karls-Universität Heidelberg, 2004.
- [95] W. RHEINBOLDT, *On a theory of mesh-refinement processes*, SIAM J. Numer. Anal., 17 (1980), pp. 766–778.
- [96] J. W. RUGE AND K. STÜBEN, *Efficient Solution of Finite Difference and Finite Element Equations by Algebraic Multigrid*, in Multigrid Methods for Integral and Differential Equations, D. J. Paddon and H. Holstein, eds., The Institute of Mathematics and its Applications Conference Series, Clarendon Press, 1985.
- [97] R. RYAN, *Introduction to Tensor Products of Banach Spaces*, Springer Verlag, 2002.
- [98] T. SCHIEKOFER, *Die Methode der Finiten Differenzen auf Dünnen Gittern*, Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn, 1998.
- [99] S. SCHNEIDER, *Extrapolationsmethoden zur Lösung parabolischer Gleichungen*, Diplomarbeit, Technische Universität München, 1994.
- [100] S. SCHNEIDER AND C. ZENGER, *Multigrid Methods for Hierarchical Adaptive Finite Elements*, tech. rep., Technische Universität München, 1999.
- [101] C. SCHWAB AND R. TODOR, *Sparse finite elements for stochastic elliptic problems-higher order moments*, Computing, 71 (2003), pp. 43–63.

- 
- [102] R. STEVENSON, *Locally supported, piecewise polynomial biorthogonal wavelets on non-uniform meshes*, Constr. Approx., 19 (2003), pp. 477–508.
- [103] W. SWELDENS, *The lifting scheme: A construction of second generation wavelets*, SIAM J. Math. Anal., 29 (1997), pp. 511–546.
- [104] V. TEMLYAKOV, *Approximation of functions with bounded mixed derivative*, Proc. Steklov Inst. Math., 1 (1989).
- [105] ———, *On approximate recovery of functions with bounded mixed derivatives*, J. Complexity, 9 (1993), pp. 41–59.
- [106] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, Springer, 1997.
- [107] J. TRAUB AND H. WOŹNIAKOWSKI, *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.
- [108] U. TROTTEBERG, C. W. OSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, 2001.
- [109] S. VANDEWALLE, *Parallel Multigrid Waveform Relaxation for Parabolic Problems*, B. G. Teubner, 1993.
- [110] S. VANDEWALLE AND G. HORTIN, *A space-time multigrid method for parabolic p.d.e.s*, SIAM J. Sci. Comput., 16 (1995).
- [111] S. VANDEWALLE AND R. PIESSENS, *Numerical experiments with nonlinear multigrid waveform relaxation on a parallel processor*, Appl. Numer. Math., 8 (1991), pp. 149–161.
- [112] ———, *Efficient parallel algorithms for solving initial-boundary value and time-periodic parabolic partial differential equations*, SIAM J. Sci. Statist. Comp., 13 (1992), pp. 1330–1346.
- [113] P. VASSILEVSKI AND L. ZIKATANOV, *Multiple vector preserving interpolation mappings in algebraic multigrid*, tech. rep., Lawrence Livermore National Laboratory, 2004.
- [114] R. VERFÜRTH, *A posteriori error estimation and adaptive mesh-refinement techniques*, J. Comput. Appl. Math., 50 (1994), pp. 67–83.
- [115] ———, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, 1996.

- 
- [116] J. WEIDMAN, *Lineare Operatoren in Hilberträumen*, B.G. Teubner Stuttgart, 1976.
- [117] J. WEISSINGER, *Verallgemeinerung des Seidelschen Verfahrens*, Z. Angew. Math. Mech., 33 (1953), pp. 155–163.
- [118] P. WESSELING, *A robust and efficient multigrid method*, in Multigrid Methods, Lecture Notes in Mathematics, Springer, 1982.
- [119] —, *Theoretical and practical aspects of a multigrid method*, SIAM J. Sci. Stat. Comp., 3 (1982), pp. 387–402.
- [120] G. WITTUM, *On the robustness of ILU-smoothing*, SIAM J. Sci. Stat. Comput., 10 (1989), pp. 699–717.
- [121] J. WLOKA, *Partielle Differentialgleichungen*, Teubner, 1982.
- [122] H. WOŹNIAKOWSKI, *A survey of information-based complexity*, J. Complexity, 1 (1985), pp. 11–44.
- [123] H. YSERENTANT, *On the multi-level splitting of finite element spaces*, Numer. Math., 58 (1986), pp. 379–412.
- [124] —, *Large Scale Scientific Computing*, Birkhäuser, 1987, ch. Hierarchical Bases in the Numerical Solution of Parabolic Problems, pp. 22–36.
- [125] —, *On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives*, Numer. Math., 98 (2004), pp. 731–759.
- [126] C. ZENGER, *Sparse grids*, in Parallel Algorithms for Partial Differential Equations, W. Hackbusch, ed., Notes on Numerical Fluid Mechanics 31, Vieweg, Braunschweig, 1991, pp. 241–251.