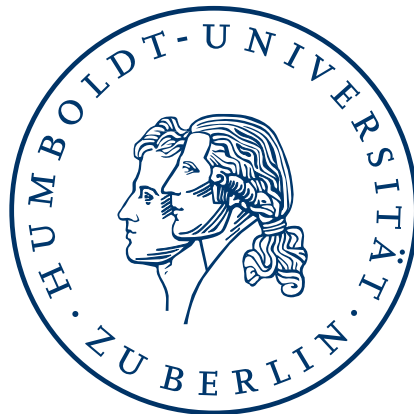


Numerische Behandlung einer linearen Reaktions-Diffusions-Gleichung mit einer Anwendung in der interzellulären Kommunikation bei Hefezellen

Diplomarbeit



Mathematisch Naturwissenschaftliche Fakultät II
Institut für Mathematik

Betreuer:

Prof. Dr. Andreas Schröder

Prof. Dr. Dr. h.c. Edda Klipp

Eingereicht von: Wolfgang Giese

geb.: 09.08.1983

Humboldt-Universität zu Berlin

Unter den Linden 6

D-10099 Berlin

Das ist eben das Hohe der Mathematik, daß ihre Methode gleich zeigt, wo ein Anstoß ist. [...] In diesem Sinne kann man die Mathematik als die höchste und sicherste Wissenschaft ansprechen. Aber wahr kann sie nichts machen, als was wahr ist.

(Johann Wolfgang von Goethe
– Maximen und Reflexionen)

Zusammenfassung

Diese Diplomarbeit behandelt die Lösung von Reaktions-Diffusions-Systemen auf komplexen Gebieten. Es werden allgemeine Grundlagen für die Formulierung von parabolischen partiellen Differentialgleichungen eingeführt und Aussagen über Existenz und Eindeutigkeit von Lösungen getroffen. Dabei werden wir auf eine lineare Reaktions-Diffusions-Gleichung mit gemischten Randwerten eingehen und dann mit Hilfe des Banach'schen Fixpunktsatzes Existenz und Eindeutigkeit nichtlinearer Systeme untersuchen. Als Beispiel werden wir das Brüsselator-System betrachten. Zur numerischen Approximation der erwähnten Gleichungen wird das unstetige Galerkin-Verfahren vorgestellt. Insbesondere werden wir uns mit den Interior-Penalty-Methoden befassen. Wir erhalten Fehler-Abschätzungen für die $\ell^2(H^1)$ - sowie die $\ell^\infty(L^2)$ -Norm. Zur Implementierung der Approximationsverfahren wird das Programm DUNE benutzt. Insbesondere wird auf Parallelisierungsaspekte eingegangen und es werden Tests auf einem Cluster durchgeführt. Diese Methoden werden genutzt, um die interzelluläre Kommunikation von Hefezellen physikalisch zu beschreiben und verschiedene Aspekte der Interaktion graphisch darzustellen.

Danksagung

Der Abschluss dieser Diplomarbeit bedeutet für mich zugleich Anfang und Ende eines wichtigen Lebensabschnittes. Ende deshalb, weil mein Studium bald beendet ist und Anfang, weil ich demnächst einen neuen und spannenden Lebensabschnitt beginnen darf.

Besonders im letzten Teil meines Studiums haben mich zwei wichtige Menschen begleitet. Ich möchte mich herzlich bei Prof. Edda Klipp und Prof. Andreas Schröder bedanken. Sie ermöglichten mir dieses interessante, weil interdisziplinäre Thema. Es war sicherlich ein erhöhter kommunikativer und organisatorischer Aufwand erforderlich, um die Fachbereiche Mathematik und Biologie erfolgreich in Einklang zu bringen. Ihr Vertrauen und die wichtigen gemeinsamen Gespräche waren für mich eine große Motivation.

Es gab aber auch andere Menschen, die mir mit Rat und Tat zur Seite standen. Zunächst meinen Dank an Christian Diener. Mit ihm konnte ich zahlreiche anregende und hilfreiche Diskussionen führen, die mir an mancher Stelle wichtige Impulse vermittelten.

Dr. Martin Eigel möchte ich danken für seine Unterstützung beim Programm DUNE und der Arbeit auf dem Cluster in Adlershof. Er stand mir oft mit seinen C++ Kenntnissen zur Seite.

Dem DUNE-Team möchte ich für die überaus schnellen Reaktionen auf meine Anfragen, für die ständige Hilfsbereitschaft und die raschen Anpassungen der Software bei entstandenen Problemen danken.

Insbesondere möchte ich hier Dr. Christian Engwer für seine zahlreichen Ratschläge danken, die mir an entscheidenden Stellen weiter geholfen haben.

Der Berlin Mathematical School möchte ich meinen Dank bekunden für die finanzielle Unterstützung, damit ich an der "Winter School and ICOR 2010 on PDE and Mathematical Biology" und dem "DUNE-User Meeting" teilnehmen konnte.

Meinen Freunden Isabella Thiesen und Christoph Wollstein möchte ich für ihre Aufmunterungen und für das Korrekturlesen danken.

Meinem Vater danke ich für seine Ratschläge und seine Unterstützung.

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Einleitung | 1 |
| 2 | Grundlagen | 4 |
| 2.1 | Schwache Ableitungen und Sobolew-Räume | 4 |
| 2.2 | Das Bochner-Integral und der Lebesgue-Raum $L_p(0, T; X)$ von vektorwertigen Funktionen | 7 |
| 2.3 | Evolutionstrippel und verallgemeinerte Ableitungen | 10 |
| 2.4 | Modellproblem für eine lineare Reaktions-Diffusions-Gleichung | 12 |
| 2.5 | Schwache Formulierung des Modellproblems | 13 |
| 3 | Analytische Untersuchung einer linearen Reaktions-Diffusions-Gleichung | 16 |
| 3.1 | Stetigkeit der Linear- und Bilinearform sowie Koerzivität | 16 |
| 3.2 | Existenz und Eindeutigkeit für eine Reaktions-Diffusions-Gleichung | 18 |
| 3.3 | Positivität, Beschränktheit und eine Erhaltungsgleichung | 22 |
| 4 | Nichtlineare Reaktions-Diffusions-Gleichungen und Systeme | 25 |
| 4.1 | Der Banach'sche Fixpunktsatz | 25 |
| 4.2 | Ein Beispiel: Der Brüsselator | 28 |
| 5 | Das unstetige Galerkin-Verfahren | 31 |
| 5.1 | Idee | 31 |
| 5.2 | Eine Übersicht | 33 |
| 5.3 | Motivation | 34 |
| 5.4 | Diskretisierung im Ort | 35 |
| 5.4.1 | Notationen | 35 |
| 5.4.2 | Gebrochene Sobolew-Räume | 36 |
| 5.4.3 | Durchschnitte und Sprünge | 37 |
| 5.4.4 | Variationsformulierung | 38 |
| 5.5 | Fehleranalyse für die Semidiskretisierung auf nicht konvexen Gebieten und nicht konformen Gittern | 42 |
| 5.5.1 | Annahmen | 42 |
| 5.5.2 | Approximationssatz, inverse Ungleichung und Abschätzung für den Strafterm | 43 |
| 5.5.3 | Koerzivität | 53 |
| 5.5.4 | ODE-Formulierung | 57 |
| 5.5.5 | Existenz und Eindeutigkeit | 57 |

| | | |
|-----------|--|------------|
| 5.5.6 | A priori-Schranken | 57 |
| 5.5.7 | Konvergenzbeweis | 59 |
| 5.6 | Erhaltungsgleichungen | 63 |
| 5.6.1 | Lokale Massenerhaltung | 63 |
| 5.6.2 | Globale Erhaltungsgleichung | 63 |
| 5.7 | Zeitdiskretisierung durch das Theta-Verfahren | 64 |
| 5.7.1 | Existenz und Eindeutigkeit | 64 |
| 5.8 | Fehleranalysis für das voll diskretisierte Problem auf nicht konvexen Gebieten und nicht konformen Gittern | 65 |
| 6 | Implementierung | 75 |
| 6.1 | Das Programm: DUNE | 75 |
| 6.2 | Parallelisierung | 76 |
| 6.2.1 | Der klassische additive überlappende Schwarz-Algorithmus für zwei Gebiete | 76 |
| 6.2.2 | Der abstrakte additive überlappende Schwarz-Algorithmus | 77 |
| 6.2.3 | Anwendung des Algorithmus auf das DG-Verfahren | 79 |
| 6.3 | Systeme von Reaktions-Diffusions-Gleichungen | 79 |
| 7 | Numerische Experimente | 81 |
| 7.1 | Konvergenztests für eine parabolische Gleichung | 81 |
| 7.2 | Tests auf einem Linux-Cluster mit 256 Kernen | 87 |
| 8 | Biologische Anwendung in der interzellulären Kommunikation bei Hefezellen | 92 |
| 8.1 | Biologischer Hintergrund | 92 |
| 8.2 | Grundlagen der Modellierung | 93 |
| 8.2.1 | Physikalische Grundlagen von Reaktions-Diffusions-Gleichungen | 93 |
| 8.2.2 | Randbedingungen | 94 |
| 8.3 | Mathematisches Modell | 96 |
| 8.4 | Numerische Resultate und Interpretation des Modells | 99 |
| 9 | Schlussbetrachtung und Ausblick | 107 |
| 10 | Erklärung | 115 |

Abkürzungen

| Abkürzung | Bedeutung |
|-------------------------|--|
| MAT | Indikator für den Paarungstyp |
| ODE | Gewöhnliche Differentialgleichung |
| DG | Discontinuous Galerkin |
| FE-Methode | Finite Elemente-Methode |
| DUNE | Distributed and Unified Numerics Environment |
| MPI | Message Passing Interface |
| IP | Interior Penalty |
| SIPG | Symmetric Interior Penalty Galerkin |
| NIPG | Nonsymmetric Interior Penalty Galerkin |
| IIPG | Incomplete Interior Penalty Galerkin |
| RD-Gleichung | Reaktions-Diffusions-Gleichung |
| \mathcal{T}_h | Triangulierung bzgl. eines Gebietes Ω |
| K | Element der Triangulierung - hier meist ein Dreieck (2D) oder ein Tetraeder (3D) |
| Γ | Eine Kante (2D) oder eine Fläche (3D) der Triangulierung |
| $C_0^\infty(\Omega)$ | Raum der unendlich oft differenzierbaren Funktionen mit kompaktem Träger in Ω |
| p | Polynomgrad der verwendeten Approximationspolynome |
| $\mathcal{P}^p(\Omega)$ | Raum der Polynome vom Grad $\leq p$ auf einem Gebiet Ω |
| S_h^p | Raum der gebrochenen Polynome vom Grad $\leq p$ bzgl. einer Triangulierung \mathcal{T}_h |
| d | Dimension des betrachteten Gebietes Ω |
| h | Maß für die Feinheit des Gitters |
| τ | Zeitschrittweite |

Tabellenverzeichnis

| | | |
|-----|---|----|
| 5.1 | Ein Vergleich der vorgestellten DG-Verfahren. | 33 |
| 7.1 | Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Verfahren für P1-Polynome auf dem Einheitsquadrat. | 83 |
| 7.2 | Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Zeitschritt für P1-Polynome auf einem Einheitsquadrat mit Löchern. | 83 |
| 7.3 | Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Verfahren für P2-Polynome auf dem Einheitsquadrat. | 83 |
| 7.4 | Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Verfahren für P2-Polynome auf einem Einheitsquadrat mit Löchern. | 84 |

Abbildungsverzeichnis

| | | |
|-----|--|----|
| 4.1 | Trapping-Region für den Brüsselator. | 29 |
| 5.1 | Im Bild sieht man eine Approximation auf einem groben Gitter, erzeugt durch das unstetige Galerkin-Verfahren. Unstetigkeiten werden bestraft, sind aber möglich. Wie in der Abbildung zu erkennen ist, werden insbesondere bei starken Gradienten kleine Sprünge zwischen den Elementen erlaubt. | 32 |
| 5.2 | Elemente mit hängenden Knoten [Fei09]. | 36 |
| 5.3 | Zwei benachbarte Gitterelemente [Fei09]. | 37 |
| 5.4 | Das betrachtete Referenzelement ist ein rechtwinkliges Dreieck mit $h_{K_{\text{ref}}} = \sqrt{2}$ und $\rho_{K_{\text{ref}}} = (2 + \sqrt{2})^{-1}$ | 43 |
| 7.1 | Lösung der Wärmeleitungsgleichung auf dem Einheitsquadrat. | 84 |
| 7.2 | Konvergenzgraph für die $\ell^\infty(L^2)$ -Norm. | 85 |
| 7.3 | Lösung der Wärmeleitungsgleichung auf einem Gebiet mit Löchern. | 85 |
| 7.4 | Konvergenzgraph für die $\ell^2(H^1)$ -Norm. | 86 |
| 7.5 | Lösung der Helmholtz-Gleichung auf der Einheitskugel. | 87 |
| 7.6 | Ergebnis einer parallelen Rechnung auf 10 Prozessoren. 4 von 10 Teilgebiete wurden entfernt. Man erkennt auch die einzelnen Tetraederelemente. | 88 |
| 7.7 | Eine 3-dimensionale Kugel wird in 24 Teilgebiete auf dem Cluster zerlegt. Die Teilgebiete werden durch unterschiedliche Farben dargestellt. | 88 |
| 7.8 | Abbildung der Rechenzeiten für die Lösung der Helmholtz-Gleichung auf einer Kugel mit 66560 Elementen unter Verwendung von MVAPICH. Die grünen Balken stellen die reine Zeit zum parallelen Lösen des Gleichungssystems dar. Die roten Balken beinhalten den gesamten Prozess vom Verteilen des Gitters, der Assemblierung der Matrix bis zum Versenden der vtu-Bilddateien auf den Head-Knoten. | 89 |
| 7.9 | Abbildung der Rechenzeiten für die Lösung der Helmholtz-Gleichung auf einer Kugel mit 66560 Elementen unter Verwendung von OpenMPI. Die grünen Balken stellen die reine Zeit zum parallelen Lösen des Gleichungssystems dar. Die roten Balken beinhalten den gesamten Prozess vom Verteilen des Gitters, der Assemblierung der Matrix bis zum Versenden der vtu-Bilddateien auf den Head-Knoten. | 90 |

| | | |
|-----|--|-----|
| 8.1 | Vereinigungsprozess der <i>Saccharomyces cerevisiae</i> . Die roten Punkte stellen α -Faktor dar, die blauen Punkte stehen für Bar1p. Abbildung übernommen und verändert aus http://en.wikipedia.org/wiki/File:Yeast_mating_scheme.svg | 93 |
| 8.2 | Schematische Darstellung des Gebietes Ω | 97 |
| 8.3 | Verteilung der MATa/ α -Zellen. | 100 |
| 8.4 | Verteilung von Alpha-Faktor nach 2s, 10s und 20s. | 102 |
| 8.5 | Verteilung von Bar1p nach 5 und 15s. | 103 |
| 8.6 | Mikroskopbild. Das Bild wurde von Dr. Gabriele Schreiber (AG Prof. Andreas Herrmann) und Christian Diener bereitgestellt. | 104 |
| 8.7 | Ein Gitter mit 45.489 Elementen. Die Gitter-Erzeugung basiert auf einem Skript, welches in Zusammenarbeit mit Christian Diener erstellt wurde. . . | 104 |
| 8.8 | Verteilung von α -Faktor ohne Bar1p-Sekretion. | 106 |
| 8.9 | Verteilung von α -Faktor mit Einfluss von Bar1p. | 106 |
| 9.1 | Mikroskopbild als Grundlage der Gittergenerierung. (Bild von Dr. Gabriele Schreiber) | 109 |
| 9.2 | Gitter mit natürlichen Zellformen. Die Gittergenerierung auf Basis von Mikroskopdaten basiert auf auf einem Skript von Christian Diener. | 109 |
| 9.3 | Verteilung von α -Faktor mit Einfluss von Bar1. | 110 |
| 9.4 | Verteilung von Bar1p. | 110 |

1 Einleitung

In den Lebenswissenschaften werden quantitative Modelle immer wichtiger um die Funktionsweise von Organismen zu verstehen. So besitzen zelluläre Netzwerke ihre ganz eigene Dynamik - welche sich mit Hilfe von mathematischen Modellen beschreiben lassen.

Um wichtige Parameter zu schätzen oder die Wirkungsweise verschiedener Stoffe zu verstehen werden Simulationen im Computer benötigt. Komplexe biologische Prozesse lassen sich dadurch veranschaulichen und können mit experimentellen Untersuchungen abgeglichen werden.

Räumliche Modelle gewinnen dabei immer mehr an Bedeutung. Alan Turing war einer der Pioniere, welcher in den 50-er Jahren Reaktions-Diffusions-Prozesse als grundlegend für viele zelldynamische Vorgänge erkannte. In seinem Werk "The chemical basis of morphogenesis" [Tur52] untersucht er Systeme von Reaktions-Diffusions(RD)-Gleichungen und zeigt, dass man mit diesen Gleichungen Phänomene wie Struktur und Musterbildung erklären kann.

Auch zahlreiche interzelluläre Prozesse werden durch Diffusion getragen: Moleküle breiten sich in einem Medium aus, bilden Gradienten und interagieren mit anderen Stoffen. Ihre Bewegungen können durch die Brown'sche Molekularbewegung beschrieben werden. Liegt eine größere Zahl von Molekülen vor, so können wir Diffusionsprozesse auf makroskopischer Ebene durch RD-Gleichungen beschreiben [Mur02].

Wir werden in dieser Arbeit ein System von partiellen Differentialgleichungen untersuchen, welches ein Modell für die interzelluläre Kommunikation von Hefezellen liefern soll. Unsere Untersuchungen beginnen wir mit einer linearen RD-Gleichung

$$\partial_t u = D\Delta u - k(x, t)u + f(x, t) \text{ in } (0, T) \times \Omega. \quad (1.1)$$

Dabei ist u die Konzentration der betrachteten Moleküle, D die Diffusionskonstante, k die Degradationsrate und f der Quellterm. Die Gleichung beschreibt die ablaufenden Prozesse in einem beschränkten Gebiet Ω über dem Zeitintervall $(0, T)$. Auf den Gebietsgrenzen geben wir Dirichlet- und Neumann-Randbedingungen vor

$$u = g_D \quad \text{auf } (0, T) \times \Gamma_D, \quad (1.2)$$

$$\frac{\partial u}{\partial n} = g_N \quad \text{auf } (0, T) \times \Gamma_N, \quad (1.3)$$

sowie die Konzentration zum Startzeitpunkt $t = 0$

$$u = u^0 \quad \text{in } \{0\} \times \Omega. \quad (1.4)$$

Die Lösung der Gleichungen (1.1) - (1.4) beschreibt die Konzentrationsverteilung der Spe-

zies u im Gebiet Ω zu einem bestimmten Zeitpunkt.

Verwenden wir den herkömmlichen Lösungsbegriff aus der Theorie der gewöhnlichen Differentialgleichungen, so erwarten wir, dass eine Lösung zweimal stetig differenzierbar in jeder Ortskoordinate bzw. einmal stetig differenzierbar in der Zeit ist. Das Ergebnis bezeichnen wir als eine Lösung im *klassischen* Sinne.

Dieser Ansatz ist jedoch zu restriktiv für die gegebene Problemstellung. In den 30-er Jahren führte Sergei Lwowitsch Sobolew einen neuen Lösungsbegriff ein. Ableitungen werden nicht mehr über den Differenzenquotienten definiert, sondern über Integrale durch partielle Integration. Diese Ableitungen nennt man *schwache* Ableitungen. Das oben aufgeführte Problem kann damit im schwachen Sinne, also unter Verwendung der schwachen Ableitungen, formuliert werden. Dies führt zu einer *Variationsformulierung*, welche an die Stelle der Ausgangsgleichung tritt. Hierzu werden im Kapitel 2 die exakten Definitionen vorgestellt.

Im Kapitel 3 werden Existenz- und Eindeutigkeitsaussagen zu dem oben genannten Problem getroffen. Zudem werden unter bestimmten Annahmen physikalische Eigenschaften wie die Existenz einer positiven Lösung nachgewiesen. Anschließend werden wir diese Aussagen in Kapitel 4 mit Hilfe des Banach'schen Fixpunktsatzes auf nichtlineare Systeme übertragen und am Beispiel des bekannten Brüsselatorsystems untersuchen.

Auf der Grundlage der erwähnten Variationsformulierung beruht auch die *Finite Elemente*(FE)-Methode. Sie wird seit den 40-er Jahren entwickelt und erprobt und hat eine Vielzahl an Verfahren für ein großes Spektrum von Problemen hervorgebracht.

In dieser Arbeit wollen wir das *unstetige Galerkin*(DG)-Verfahren untersuchen. Dieses Verfahren ist aufgrund seiner physikalischen Exaktheit besonders für die Approximation von Diffusionsprozessen geeignet [BG99]. Wir verwenden hier die *Nonsymmetric Interior Penalty Galerkin*(NIPG)-Methode. Ähnlich wie beim Finite Volumen-Verfahren erfüllt dieses Verfahren eine lokale Erhaltungsgleichung und ist deshalb auch besonders für zeitabhängige Probleme geeignet [RW00].

Wie bei den FE-Methoden wird das Gebiet Ω in eine endliche Anzahl von Polygonen, meist Dreiecke, zerlegt. Die Polygone können sich in Größe und Form unterscheiden. Auf jedem dieser Polygone definieren wir ein *Element*, bei dem es sich um eine lineare Funktion oder auch ein Polynom höherer Ordnung handeln kann. Die Lösung der Variationsformulierung können wir mit Hilfe dieser Elemente approximieren. Bei dem unstetigen Galerkin-Verfahren sind Unstetigkeiten entlang der Gitterkanten erlaubt, daher auch der Name.

Die Diskretisierung basiert auf Flüssen entlang der Kanten in unserem Finite Elemente-Gitter, wobei die entstehenden Gleichungen lokal bezüglich der Gitterelemente aufgestellt werden. Das Verfahren eignet sich deshalb besonders gut zur Parallelisierung auf mehreren Prozessoren. Wir verwenden dafür das abstrakte additive Schwarz-Verfahren [TW04, S.35ff] und zeigen in numerischen Tests, dass sich durch die Parallelisierung erhebliche Geschwindigkeitsvorteile ergeben.

Verschiedene Aspekte des Verfahrens wie Konvergenzordnung und Erhaltungsgleichungen werden in Kapitel 5 vorgestellt. Zur Diskretisierung verwenden wir die Liniennmethode und richten uns bei der Untersuchung der Konvergenz im Ort nach Feistauer [Fei09]. Für die Zeitdiskretisierung verwenden wir Approximationssätze aus Révère

[RW00]. Unter den Annahmen von Feistauer kommen wir mit leicht veränderten Konstanten und Parametern zu einem ähnlichen Resultat wie Revère [Riv08, S.82-87]. Wir leiten die Konvergenzordnung in der $\ell^\infty(L^2(\Omega))$ - und $\ell^2(H^1(\Omega))$ -Norm her. Wir bezeichnen mit h den größten Durchmesser der Elemente des verwendeten FE-Gitters, τ der grösste Zeitschritt in der Zeitdiskretisierung und p der Polynomgrad der verwendeten Approximationspolynome. In den genannten Normen ist der Fehler von der Größenordnung $\mathcal{O}(h^p) + \mathcal{O}(\tau^2)$ für das Crank-Nicolson-Zeitschrittverfahren und $\mathcal{O}(h^p) + \mathcal{O}(\tau)$ für das implizite Euler-Zeitschrittverfahren.

Die Implementierung erfolgt mit dem Open Source-Programm DUNE - Distributed and Unified Numerics Environment. Es vereint verschiedene Konzepte in einem Programm und wird unter anderem an den Universitäten Heidelberg, Freiburg, Stuttgart und an der FU Berlin entwickelt [BDE⁺05]. Es ist international verbreitet und wird von verschiedenen Institutionen und Unternehmen genutzt. Wir werden insbesondere das Tool DUNE-PDELAB verwenden, welches zahlreiche Methoden zur Lösung von partiellen Differentialgleichungen bereitstellt sowie Schnittstellen zur Parallelisierung auf Basis von MPI(Message Passing Interface) beinhaltet. Im Kapitel 6 werden wir unter anderem auch auf die Implementierung der Parallelisierung nach [BB08] und das numerische Lösen von Systemen eingehen.

Im Kapitel 7 werden einige numerische Tests durchgeführt, um die theoretischen Resultate aus Kapitel 5 an einem Beispiel zu überprüfen und aufzeigen, dass sich durch den parallelisierten Einsatz von DUNE auf einem Linux-Cluster erhebliche Geschwindigkeitsvorteile ergeben.

Im Kapitel 8 wird ein Modell für die interzelluläre Kommunikation während der Paarung der Hefespezies *Saccharomyces cerevisiae* entwickelt. Diese Hefespezies ist einer der am besten untersuchten Modellorganismen. Die Signalerkennung ist ein komplexes dynamisches Netzwerk, welches jedoch evolutionär in der Entwicklung von komplexen eukaryotischen Zellen (im Speziellen bei Säugetierzellen) konstant geblieben ist [Bar05]. Für den Paarungsprozess ist die Richtung und die Lokalisierung des Signals wichtig. Inhomogenitäten in der Verteilung der beteiligten Botenstoffe spielen deshalb eine entscheidende Rolle und machen räumliche Modelle für Signalerkennung und Signalverarbeitung unabdingbar. Eine Anwendung von RD-Gleichungen für Prozesse wie Signalverstärkung oder das Aufrechterhalten von Gradienten im Zellinneren findet sich in [CSQC08]. In dieser Arbeit wird ein Modell für die räumliche Signalverteilung im extrazellulären Bereich erstellt. Die RD-Gleichungen werden durch die vorgestellte Methode auf komplexen Gebieten berechnet, welche Zellen anhand von Mikroskopdaten abbilden sollen. Zunächst werden qualitative Aussagen aus dem Modell abgeleitet. Das Modell soll später als Grundlage dienen, um in Experimenten auch quantitative Aussagen zu erhalten.

2 Grundlagen

In diesem Kapitel sollen die wichtigsten grundlegenden Begriffe und Theoreme eingeführt werden. Der Begriff der schwachen Ableitung und die damit verbundenen Normen, Räume und Theoreme werden im Zentrum stehen. Diese Begriffe werden in den Kapiteln 3, 4 und teilweise auch in Kapitel 5 verwendet. Die schwachen Ableitungen werden hier nur im Ort betrachtet. Für die Menge der Funktionen mit schwachen Ableitungen und gewissen weiteren Einschränkungen werden wir die Sobolew-Räume einführen. Dies sind Vektorräume, die mit einer Norm ausgestattet werden und damit Banachräume bilden. Wir erhalten also einen Abstandsbegriff zwischen den Funktionen. Die definierte Norm induziert insbesondere ein Maß für Fehler und Konvergenz, die wir unter anderem für das numerische Verfahren in Kapitel 5 benötigen.

Die auftretende Zeitableitung wird gesondert behandelt. Wir betrachten Funktionen, die auf einem Zeitintervall leben und Werte in den erwähnten Sobolew-Räumen annehmen. Man erzeugt damit eine gewisse Analogie zu den gewöhnlichen Differentialgleichungen (ODEs). Dies wird in dem Existenzbeweis in Kapitel 3 deutlich sowie an dem später vorgestellten numerischen Verfahren, welches die partielle Differentialgleichung zunächst auf eine ODE reduziert.

Wir definieren einen neuen Integralbegriff, das *Bochner-Integral* für Funktionen mit Werten in Banachräumen. Analog zur Definition der schwachen Ableitung in den Ortskoordinaten definieren wir Ableitungen von Funktionen in Banachräumen. Diese Ableitung wird dann *verallgemeinerte* Ableitung genannt.

Diese Begriffe sind Grundlage für das Aufstellen einer Variationsformulierung, welche zusammen mit den Eingangsdaten auch als *schwache Formulierung* bezeichnet wird. Sie ist weniger restriktiv als die sogenannte *klassische* Formulierung, welche starke Differenzierbarkeit im Sinne von Differenzenquotienten voraussetzt. Auf Grundlage der schwachen Formulierung können wir dann im folgenden Kapitel verschiedene Eigenschaften des Modellproblems wie Existenz, Eindeutigkeit, Beschränktheit und Positivität untersuchen.

2.1 Schwache Ableitungen und Sobolew-Räume

Wir wählen zwei beliebige Funktionen f und φ aus dem Raum der unendlich oft differenzierbaren Funktionen mit kompaktem Träger in Ω . Diesen Raum bezeichnen wir mit $C_0^\infty(\Omega)$. Damit gilt insbesondere, dass die beiden Funktionen auf dem Rand von Ω verschwinden. Wir nehmen hier an, dass Ω beschränkt ist. Wir erhalten durch partielle

Integration die Beziehung

$$\int_{\Omega} \frac{\partial f}{\partial x^i} \varphi dx = - \int_{\Omega} f \frac{\partial \varphi}{\partial x^i} dx.$$

Diese Relation verwenden wir nun um schwache Differenzierbarkeit zu definieren.

Definition 2.1.1 (Schwache Ableitung). Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt, $i \in \{1, \dots, d\}$, $f \in L^2(\Omega)$. Dann heißt $g \in L^2(\Omega)$ *schwache i -te Ableitung* von f genau dann, wenn

$$\int_{\Omega} g \varphi dx = - \int_{\Omega} f \frac{\partial \varphi}{\partial x^i} dx \text{ für alle } \varphi \in C_0^\infty(\Omega).$$

Offensichtlich ist für differenzierbare Funktionen $f \in L^2(\Omega) \cap C^1(\Omega)$ die schwache Ableitung identisch zur klassischen Ableitung. Diese Definition kann rekursiv auf Ableitungen höherer Ordnung ausgeweitet werden. Zur Beschreibung von Ableitungen in unterschiedliche Richtungen und von höherer Ordnung werden wir Multiindizes einführen.

Definition 2.1.2 (Multiindex). Sei $I \subset \mathbb{N}$ eine Indexmenge. $\alpha = (\alpha_1, \dots, \alpha_d) \in I^d$ heißt *Multiindex* und wir bezeichnen mit $|\alpha| = \sum_{i=1}^d \alpha_i$ dessen *Ordnung*.

Zu dem Multiindex definieren wir außerdem einen Differentialoperator.

Definition 2.1.3. Sei α ein Multiindex. Dann bezeichnet D^α den Differentialoperator $D^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \dots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}}$.

Wir wollen die Funktionen, deren schwache Ableitungen existieren, in einem Raum zusammenfassen.

Definition 2.1.4. Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt. Dann definieren wir den *Sobolew-Raum* $H^m(\Omega)$ als

$$H^m(\Omega) = \{f \in L^2(\Omega) \mid D^\alpha f \in L^2(\Omega) \text{ existiert für alle } |\alpha| \leq m\}$$

mit dem Skalarprodukt

$$(f, g)_{H^m(\Omega)} = \sum_{|\alpha| \leq m} \int_{\Omega} D^\alpha f \cdot D^\alpha g \, dx \text{ für alle } f, g \in H^m(\Omega)$$

und der dazugehörigen Norm

$$\|f\|_{H^m(\Omega)} = \sum_{|\alpha| \leq m} \|D^\alpha f\|_{L^2(\Omega)}.$$

Satz 2.1.1. Der Sobolew-Raum $(H^m(\Omega), \|\cdot\|_{H^m(\Omega)})$ ist ein Hilbertraum.

Beweis. Siehe [Alt06, S.62-63]. □

Wir möchten aus Sobolew-Funktionen im Gebietsinneren Funktionen auf dem Gebietsrand konstruieren können und umgekehrt. Dazu wollen wir hier den Spursatz zitieren. Dieser gilt jedoch nicht für beliebige Gebiete mit beliebigen Rändern. Wir wollen uns deshalb auf Gebiete mit Lipschitz-Rand beschränken.

Definition 2.1.5 (Lipschitz-Rand). Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt. Wir sagen Ω hat einen *Lipschitz-Rand*, falls sich $\partial\Omega$ durch endlich viele offene Mengen U^1, \dots, U^l , überdecken lässt, so dass $\partial\Omega \cap U^j$ für $j = 1, \dots, l$ der Graph einer Lipschitz-stetigen Funktion ist und $\Omega \cap U^j$ auf jeweils einer Seite dieses Graphen liegt.

Satz 2.1.2 (Spursatz). Sei Ω offen und beschränkt mit Lipschitz-Rand. Dann gibt es genau eine stetige lineare Abbildung

$$\gamma : H^1(\Omega) \rightarrow L^2(\partial\Omega) \text{ (Spuroperator)}$$

so dass

$$\gamma u = u|_{\partial\Omega} \text{ für } u \in H^1(\Omega) \cap C^0(\bar{\Omega}).$$

Wir bezeichnen mit γu die Spurwerte oder schwachen Randwerte von u auf $\partial\Omega$.

Beweis. Siehe [Alt06, S.265]. □

Notation. Wir werden im Folgenden in der Regel $u(x)$ anstelle von $\gamma u(x)$ für $x \in \partial\Omega$ schreiben.

Satz 2.1.3 (Green'sche Formel). Sei Ω offen und beschränkt mit Lipschitz-Rand. Mit n bezeichnen wir die äußere Einheitsnormale zum Gebietsrand $\partial\Omega$. Seien $f \in H^2(\Omega)$, $g \in H^1(\Omega)$, dann gilt:

$$\int_{\Omega} \Delta f g dx = - \int_{\Omega} \nabla f \cdot \nabla g dx + \int_{\partial\Omega} \frac{\partial f}{\partial n} g dS.$$

Satz 2.1.4 (Sobolew'scher Einbettungssatz). Für $\Omega \subset \mathbb{R}^d$ gilt

$$H^m(\Omega) \subset C^r(\Omega), \text{ falls } \frac{1}{2} < \frac{m-r}{d}.$$

Beweis. Siehe [Alt06, S.333-334]. □

Korollar. Sei Ω offen, beschränkt und mit Lipschitz-Rand. Dann ist:

$$H^1(\Omega) \subset C^0(\Omega) \text{ für } d = 1,$$

und

$$H^2(\Omega) \subset C^0(\Omega) \text{ für } d = 2, 3.$$

Satz 2.1.5 (Rellichscher Auswahlssatz). Sei $k \geq 0$ und $\Omega \subset \mathbb{R}^d$ ein beschränktes Gebiet mit Lipschitz-Rand und erfülle die Kegelbedingung, d.h., die Innenwinkel an den Ecken seien positiv, so dass man einen Kegel mit positivem Scheitelwinkel so in Ω verschieben kann, dass er die Ecken berührt. Dann ist die Einbettung $H^{k+1}(\Omega) \hookrightarrow H^k(\Omega)$ kompakt.

Beweis. Ein Beweis findet sich in [Wer09, S.220 & Anm. S.222]. \square

2.2 Das Bochner-Integral und der Lebesgue-Raum $L_p(0, T; X)$ von vektorwertigen Funktionen

Im Folgenden wollen wir klären, wie das Integral über vektorwertige Funktionen verstanden werden soll. Dies führt uns zur Definition des *Bochner-Integrals*. Wir folgen den Ausführungen nach [Gaj74, S. 124ff].

Definition 2.2.1 (Bochner-Integral). Sei X ein Banachraum über \mathbb{R} .

- (a) Jede Abbildung $u : [0, T] \rightarrow X$ der Form $u(t) = \sum_{i=0}^k \chi_{B_i}(t)x_i$ mit $x_i \in X$, $1 \leq i \leq k$ und paarweise disjunkten Lebesgue-messbaren Mengen $B_i \subset [0, T]$, $1 \leq i \leq k$, wird eine *Treppenfunktion* genannt. Für diese Funktionen definieren wir das *Bochner-Integral* durch

$$\int_0^T u(t)dt := \sum_{i=1}^k \lambda^1(B_i)x_i,$$

wobei $\lambda^1(B_i)$ das eindimensionale Lebesgue-Maß bezeichnet.

- (b) $u : [0, T] \rightarrow X$ wird genau dann *Bochner-messbar* genannt, falls eine Folge von Treppenfunktionen $\{u_n\}$ existiert mit

$$u(t) = \lim_{n \rightarrow \infty} u_n(t), \text{ für fast alle } t \in (0, T). \quad (2.1)$$

Wenn für die Folge $\{u_n\}$ gilt

$$\lim_{n \rightarrow \infty} \int_0^T \|u_n(t) - u(t)\|_X dt = 0, \quad (2.2)$$

dann wird u *Bochner-integrierbar* genannt und

$$\int_0^T u(t)dt = \lim_{n \rightarrow \infty} \int_0^T u_n(t)dt \quad (\in X). \quad (2.3)$$

Notation (Duale Paarung). Sei X ein reeller Vektorraum und X^* sein Dualraum. Sei $f \in X^*$. Dann notieren wir die duale Paarung durch:

$$\langle f, v \rangle := f(v) \text{ für alle } v \in X.$$

Wir wollen einige Eigenschaften des Bochner-Integrals festhalten.

Satz 2.2.1. (a) Ist X separabel, dann gilt. $u : [0, T] \rightarrow X$ Bochner-messbar $\Leftrightarrow t \mapsto \langle f, u(t) \rangle$ ist messbar auf $[0, T]$ für alle $f \in X^*$. (Pettis's Theorem)

(b) $X = \mathbb{R} \Rightarrow$ (Bochner-messbar \Leftrightarrow Lebesgue-messbar)

(c) $u : [0, T] \rightarrow X$ ist fast überall stetig $\Rightarrow u$ ist Bochner-messbar.

(d) Die Summe und das Produkt von Bochner-messbaren Funktionen sind Bochner-messbar.

(e) $u : [0, T] \rightarrow X$ Bochner-messbar $\Rightarrow t \mapsto \|u(t)\|_X$ ist Lebesgue-messbar auf $[0, T]$.

(f) Der Grenzwert in Gl. (2.3) existiert unter der Bedingung in Gl. (2.2) und führt für alle Folgen von Treppenfunktionen mit (2.1) und (2.2) zum selben Wert. Damit ist das Bochner-Integral wohldefiniert.

Beweis. Siehe [Gaj74, S. 124ff]. □

Mit dem eingeführten Integralbegriff können wir nun auch Lebesgue-Räume über vektorwertige Funktionen einführen.

Definition 2.2.2. Sei X ein Banachraum und $0 < T < \infty$.

(a) Die Änderung von u auf einer Menge mit λ^1 -Maß Null ändert weder die Messbarkeit noch den Wert des Bochner-Integrals. Im Folgenden bezeichnet $\mathcal{L}^p(0, T; X)$, $1 \leq p < +\infty$ die Bochner-messbaren Funktionen auf $(0, T)$ mit:

$$\mathcal{N}_p(u) := \left(\int_0^T \|u(t)\|_X^p dt \right)^{\frac{1}{p}} < +\infty. \quad (2.4)$$

Man nennt zwei Elemente von $\mathcal{L}^p(0, T; X)$ genau dann äquivalent, wenn sie fast überall auf $(0, T)$ übereinstimmen. Dann ist \mathcal{N}_p konstant in jeder dieser Äquivalenzklassen. Mit $L^p(0, T; X)$ bezeichnen wir die Menge aller Äquivalenzklassen bezüglich dieser Äquivalenzrelation.

(b) Analog können wir einen Raum $L^\infty(0, T; X)$ einführen:

$$\mathcal{L}^\infty(0, T; X) = \{u : [0, T] \rightarrow X; u \text{ Bochner-messbar, } \exists M > 0 : \|u(t)\|_X \leq M \text{ fast überall in } (0, T)\}.$$

Wenn wir wieder alle Funktionen identifizieren, die fast überall auf $[0, T]$ übereinstimmen, erhalten wir den Raum $L^\infty(0, T; X)$.

Analog zu den klassischen Ableitungen für reellwertige Funktionen können wir nun auch die Ableitung in der Zeit über einen Differenzenquotienten definieren und mit einer Norm ausstatten:

Definition 2.2.3. (a) Eine Abbildung $u : [0, T] \rightarrow X$ wird genau dann differenzierbar in $t \in [0, T]$ mit der Ableitung $u'(t) \in X$ bezeichnet, wenn

$$\lim_{h \rightarrow 0, t+h \in [0, T]} \frac{1}{|h|} \|u(t+h) - u(t) - hu'(t)\|_X = 0.$$

Wie gewöhnlich werden höhere Ableitungen induktiv definiert.

(b) Der Raum $\mathcal{C}^m([0, T], X)$ mit $m = 0, 1, \dots$ besteht aus allen stetigen Funktionen $u : [0, T] \rightarrow X$ mit stetiger Ableitung bis zur Ordnung m auf $[0, T]$ mit der Norm

$$\|u\| = \sum_{i=0}^m \max_{0 \leq t \leq T} \|u^{(i)}(t)\|_X. \quad (2.5)$$

Proposition 2.2.2. (Eigenschaften von Lebesgue-Räumen) Sei $m = 0, 1, \dots$ und $1 \leq p < \infty$. Seien X und Y Banachräume über \mathbb{R} . Dann gilt

(a) $\mathcal{C}^m([0, T], X)$ ist ein Banachraum über \mathbb{R} mit der Norm (2.5)

(b) $L_p(0, T; X)$ ist ein Banachraum über \mathbb{R} mit der Norm (2.4)

(c) Wenn X ein Hilbertraum mit dem Skalarprodukt $(\cdot, \cdot)_X$ ist, dann ist $L^2(0, T; X)$ auch ein Hilbertraum mit dem Skalarprodukt

$$(u, v) = \int_0^T (u(t), v(t))_X dt.$$

(d) Hölder-Ungleichung: Sei $1 < p < +\infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt:

$$\begin{aligned} u &\in L^p(0, T; X), v \in L^q(0, T; X^*) \\ \Rightarrow \int_0^T \langle u(t), v(t) \rangle_X dt &\leq \|u\|_{L^p(0, T; X)} \|v\|_{L^q(0, T; X^*)} \end{aligned}$$

(e) Falls X separabel, reflexiv und $1 \leq p < \infty$ ist, so ist auch $L_p(0, T; X)$ separabel und reflexiv.

Beweis. Siehe [Gaj74, S.127ff]. □

2.3 Evolutionstripel und verallgemeinerte Ableitungen

Definition 2.3.1 (Evolutionstripel). Seien V und H zwei Vektorräume mit den Eigenschaften

- (i) V ist ein reeller, separabler und reflexiver Banach-Raum,
- (ii) H ist ein reeller, separabler Hilbert-Raum,
- (iii) die Einbettung $V \subseteq H$ ist stetig, das heißt

$$\|v\|_H \leq \text{const} \|v\|_V \text{ für alle } v \in V,$$

und V ist dicht in H .

Dann bezeichnen wir

$$V \subseteq H \subseteq V^*$$

als *Evolutionstripel*.

Beispiel 2.3.1. In unserem Fall bildet

$$V := \{v \in H^1(\Omega); Tr(v) = 0\} \text{ auf } \Gamma_D$$

zusammen mit

$$H := L^2(\Omega)$$

ein Evolutionstripel.

Definition 2.3.2 (Verallgemeinerte Zeitableitung). Sei $V \subseteq H \subseteq V^*$ ein Evolutionstripel, $1 < p < \infty$, $n \in \mathbb{N}$. Eine Funktion $w \in L^q(0, T; V^*)$ wird genau dann *n-te verallgemeinerte Ableitung* von einer Funktion $u \in L^p(0, T; V)$ genannt, wenn $\frac{1}{p} + \frac{1}{q} = 1$ gilt und wenn

$$\int_0^T \varphi^{(n)}(t) u(t) dt = (-1)^n \int_0^T \varphi(t) w(t) dt \text{ für alle } \varphi \in \mathcal{C}_0^\infty(0, T). \quad (2.6)$$

Dabei hat φ Werte in \mathbb{R} , das heißt mit $\varphi^{(n)}(t)u(t)$ bzw. $\varphi(t)w(t)$ ist hier die skalare Multiplikation gemeint. Wir schreiben dann: $w = u^{(n)}$.

Definition 2.3.3 (Sobolew-Raum von vektorwertigen Funktionen). Sei $V \subseteq H \subseteq V^*$ ein Evolutionstripel, dann setzen wir:

$$H^1(0, T; V, H) = \{u \in L^2(0, T; V) \text{ und } u \text{ hat eine verallgemeinerte Ableitung } u' \in L^2(0, T; V^*)\}$$

Bemerkung 2.3.1. (i) $u^{(n)}$ ist (bis auf eine Menge vom Maß Null) eindeutig bestimmt.

(ii) Sei $V \subseteq H \subseteq V^*$ ein Evolutionstriplet und $u \in \mathcal{C}^0([0, T]; H)$. Für alle $t \in [0, T]$ existiere $w(t) := \lim_{h \rightarrow 0} (u(t+h) - u(t))$ als Grenzwert in H . Außerdem sei $w \in \mathcal{C}^0([0, T]; H)$. Dann gilt: $w = u'$

(iii) Für die Existenz der Integrale in (2.6) können wir die Young'sche Ungleichung verwenden. Offensichtlich ist die Abbildung $[0, T] \rightarrow V^*$; $t \mapsto w(t)\varphi(t)$ Bochner-messbar. Die quadratische Ungleichung $|ab| \leq \frac{1}{2}|a|^2 + \frac{1}{2}|b|^2$ liefert dann

$$\int_0^T \|\varphi(t)w(t)\|_{V^*} dt \leq \int_0^T \left(\frac{1}{2}|\varphi(t)|^2 + \frac{1}{2}\|w(t)\|_{V^*}^2 \right) dt < +\infty. \quad (2.7)$$

Für die Bochner-Integrierbarkeit haben wir nun das Normkriterium

$$\int_0^T f(t) dt \text{ existiert} \Leftrightarrow \int_0^T \|f(t)\|_X dt < +\infty.$$

Wegen Ungleichung (2.7) existiert das Integral auf der rechten Seite von Gl. (2.6) mit Werten in V^* . Das Gleiche gilt auch für das Integral auf der rechten Seite. Für fast alle $t \in (0, T)$ gilt $u(t) \in V \subset H$, also $u(t) \in V^*$ und damit $u \in L^p(0, T; V^*)$.

Beweis. Für (i) und (ii) siehe [Zei90, S.419-421]. □

Satz 2.3.1 (Charakterisierung von $u^{(n)}$). *Sei $u \in L^p(0, T; V)$. Dann gilt $w = u^{(n)}$ genau dann, wenn*

$$\int_0^T \varphi^{(n)}(t)(u(t), v)_H dt = (-1)^n \int_0^T \varphi(t)\langle w(t), v \rangle dt \text{ für alle } v \in V, \varphi \in \mathcal{C}_0^\infty(0, T). \quad (2.8)$$

Wenn Gl. (2.8) gilt, dann haben wir für fast alle $t \in (0, T)$ Sinne:

$$\frac{d^n}{dt^n}(u(t), v)_H = \langle u^{(n)}(t), v \rangle_V \text{ für alle } v \in V. \quad (2.9)$$

Beweis. Gleichung (2.9) folgt direkt aus Gl. (2.8) entsprechend der Definition der verall-

gemeinerten Ableitung. Nun gilt:

$$\begin{aligned}
 w = u^{(n)} &\Leftrightarrow \text{Gl. (2.6) wird erfüllt} \\
 &\Leftrightarrow \left\langle \int_0^T u(t) \varphi^{(n)}(t) dt, v \right\rangle_V = (-1)^{(n)} \left\langle \int_0^T w(t) \varphi^{(n)}(t) dt, v \right\rangle_V \quad \text{für alle } v \in V \\
 &\Leftrightarrow \int_0^T \left\langle u(t) \varphi^{(n)}(t), v \right\rangle_V dt = (-1)^{(n)} \int_0^T w(t) \left\langle \varphi^{(n)}(t), v \right\rangle_V dt.
 \end{aligned}$$

Die Behauptung folgt nun aus der Identität

$$\langle u(t), v \rangle_V = (u(t), v)_H \quad \text{für alle } v \in V,$$

da es sich bei $\varphi^{(n)}(t)$ nur um einen skalaren Faktor handelt. \square

Satz 2.3.2. (i) $H^1(0, T; V, H)$ ist ein Banachraum mit der Norm:

$$\|u\| = \left(\int_0^T \|u(t)\|_V^2 dt \right)^{\frac{1}{2}} + \left(\int_0^T \|u'(t)\|_{V^*}^2 dt \right)^{\frac{1}{2}}.$$

(ii) $H^1(0, T; V, H)$ kann stetig in $C^0(0, T; H)$ eingebettet werden. Das heißt also, dass ein $u \in H^1(0, T; V, H)$ bis auf eine Menge vom Maß Null mit einer Funktion aus $C^0(0, T; H)$ übereinstimmt.

(iii) *Partielle Integration:* Für alle $u, v \in H^1(0, T; V, H)$ und $0 \leq s < t \leq T$ gilt die Identität

$$(u(t), v(t))_H - (u(s), v(s))_H = \int_s^t \langle u'(\tau), v(\tau) \rangle + \langle v'(\tau), u(\tau) \rangle d\tau.$$

Beweis. Siehe [Zei90, S.422-423]. \square

Notation. Wir fixieren im Folgenden $H := L^2(\Omega)$. Wir können dann verkürzt $H^1(0, T; V)$ für $H^1(0, T; V, H)$ schreiben.

2.4 Modellproblem für eine lineare Reaktions-Diffusions-Gleichung

Die erste Gleichung, die in dieser Arbeit betrachtet wird, ist eine lineare Reaktions-Diffusions-Gleichung. Sie beschreibt einen Prozess, in dem sich eine Spezies in einem Gebiet durch Diffusion ausbreitet und linear mit der Reaktionsrate k zerfällt. Das Gebiet, in dem

die Gleichung definiert ist, bezeichnen wir mit Ω . Wir nehmen an, dass das Gebiet Ω eine beschränkte Teilmenge des \mathbb{R}^d mit $d = 2, 3$ ist. Es ergibt sich damit ein Randwertproblem. In unserem Fall betrachten wir gemischte Randwerte, dass heißt Dirichlet und Neumann-Rand. Der Rand $\partial\Omega$ wird also in zwei disjunkte Mengen, den Dirichlet-Rand Γ_D und den Neumann-Rand Γ_N unterteilt. Die Modellgleichung lautet dann:

$$\partial_t u - D\Delta u + k(x, t)u = f(x, t) \quad \text{in } (0, T) \times \Omega, \quad (2.10)$$

$$u = g_D \quad \text{auf } (0, T) \times \Gamma_D, \quad (2.11)$$

$$\frac{\partial u}{\partial n} = g_N \quad \text{auf } (0, T) \times \Gamma_N, \quad (2.12)$$

$$u = u^0 \quad \text{auf } \{0\} \times \Omega. \quad (2.13)$$

Wir suchen in dem Zeitintervall $[0, T]$ nach Lösungen. Mit n bezeichnen wir die nach außen zeigende Einheitsnormale zum Rand $\partial\Omega$. Es werden folgende Annahmen gemacht:
Annahmen (A)

1. Das Gebiet Ω hat Lipschitz-Rand.
2. Für den Neumann-Rand gilt $g_N \in L^\infty(0, T; L^2(\Gamma_N))$.
3. Für den Dirichlet-Rand gilt $\int_{\Gamma_D} dx > 0$. Wir geben eine Funktion $w \in H^1(0, T; H^1(\Omega))$ vor und wir definieren die Dirichlet-Randfunktion durch $g_D := \gamma(w)$ auf $[0, T] \times \Gamma_D$, wobei γ der Spuroperator aus Satz 2.1.2 ist.
4. $D > 0$ bezeichnet eine Konstante.
5. $f \in L^2(0, T; L^2(\Omega))$, $u^0 \in L^2(\Omega)$ und $k \in L^\infty(0, T; L^\infty(\Omega))$, $k \geq 0$ sind Funktionen.

2.5 Schwache Formulierung des Modellproblems

Der Grundgedanke bei der Formulierung einer Variationsgleichung für parabolische Differentialgleichungen ist die Trennung der zeitlichen und räumlichen Variablen. Wir multiplizieren die Gleichung (2.10) mit $v \in H^1(\Omega)$ und erhalten durch partielle Integration:

$$\frac{d}{dt} \int_{\Omega} uv dx + \int_{\Omega} \nabla u \nabla v dx - \int_{\partial\Omega} \frac{\partial u}{\partial n} \text{Tr}(v) dS + \int_{\Omega} kuv dx = \int_{\Omega} f v dx$$

für alle $v \in H^1(\Omega)$ für fast alle $t \in (0, T)$.

Wir können nun das Randintegral in einen Neumann-Rand und einen Dirichlet-Rand trennen. Für das Integral über den Neumann-Rand können wir den Gradienten $\frac{\partial u}{\partial n}$ durch die Randfunktion g_N ersetzen. Die schwache Formulierung unseres Problems im Ort lautet

also:

$$\frac{d}{dt} \int_{\Omega} u v dx + \int_{\Omega} \nabla u \nabla v dx - \int_{\Gamma_N} g_N \text{Tr}(v) dS - \int_{\Gamma_D} \frac{\partial u}{\partial n} \text{Tr}(v) dS + \int_{\Omega} k u v dx = \int_{\Omega} f v dx$$

für alle $v \in H^1(\Omega)$ für fast alle $t \in (0, T)$. (2.14)

Wir wollen die Dirichlet-Randbedingung ebenfalls in die Variationsgleichung einbringen. Dies tun wir, indem wir von u eine Funktion subtrahieren, die in Ω definiert ist und den Dirichlet-Rand erfüllt. Die entstehenden Terme bringen wir auf die rechte Seite der Variationsformulierung. Wir ersetzen u durch $\tilde{u} + w$ und u^0 durch $\tilde{u}^0 + w$. Das heißt also, \tilde{u} erfüllt entlang des Dirichlet-Randes Γ_D eine Dirichlet-Nullbedingung. Wir definieren einen entsprechenden Funktionenraum:

$$V := \{v \in H^1(\Omega); \text{Tr}(v) = 0 \text{ auf } \Gamma_D\}. \quad (2.15)$$

Wir erhalten als schwache Formulierung des Problems im Ort:

$$\begin{aligned} \frac{d}{dt} \int_{\Omega} \tilde{u} v dx + D \int_{\Omega} \nabla \tilde{u} \nabla v dx + \int_{\Omega} k(t, x) \tilde{u} v dx = \\ \int_{\Omega} f v dx + \int_{\Gamma_N} g_N \text{Tr}(v) dS - \int_{\Omega} w_t v dx - D \int_{\Omega} \nabla w \nabla v dx - \int_{\Omega} k w v dx \end{aligned}$$

für alle $v \in V$, für fast alle $t \in (0, T)$ und $\tilde{u} = \tilde{u}^0$ auf $\{0\} \times \Omega$. (2.16)

Wir wollen zu einer Formulierung mit Hilfe von Operatoren gelangen. Für eine einfachere Notation wird \tilde{u} durch u ersetzt. Wir schreiben die Gleichung (2.16) in der Form

$$\begin{aligned} \frac{d}{dt} (u(t), v)_{L^2} + a(u(t), v; t) &= \langle b(t), v \rangle \text{ auf } (0, T) \text{ für alle } v \in V \\ u &= u_0 \in L^2(\Omega). \end{aligned}$$

Wir setzen für alle $w, v \in V$

$$\begin{aligned} a(w, v; t) &= D \int_{\Omega} \nabla w \nabla v dx + \int_{\Omega} k(t, x) w v dx, \\ \langle b(t), v \rangle_V &= \int_{\Omega} f v dx + \int_{\Gamma_N} g_N \text{Tr}(v) dS - \int_{\Omega} w_t v dx - D \int_{\Omega} \nabla w \nabla v dx - \int_{\Omega} k w v dx, \end{aligned}$$

und erhalten eine Bilinearform $a : V \times V \times [0; T] \rightarrow \mathbb{R}$ und ein Funktional $b(t) \in V^*$.

Notation (Schwache Formulierung). Die Formulierung des Modellproblems in der schwa-

chen Form lautet letztendlich

$$\begin{aligned} \frac{d}{dt}(u(t), v)_{L^2} + a(u(t), v; t) &= \langle b(t), v \rangle \text{ in } (0, T) \text{ für alle } v \in V \\ u &= u_0 \in L^2(\Omega), \end{aligned} \tag{2.17}$$

wobei $u \in H^1(0, T; V, H)$ gesucht wird.

3 Analytische Untersuchung einer linearen Reaktions-Diffusions-Gleichung

In diesem Kapitel sollen wichtige analytische Aspekte von Reaktions-Diffusions-Gleichungen behandelt werden. Zunächst betrachten wir eine einzelne lineare RD-Gleichung mit gemischten Randwerten. Es werden fundamentale Eigenschaften wie Existenz, Regularität, Positivität und Beschränktheit bewiesen. Dabei werden auch die Grundlagen für die Theorie von parabolischen Differentialgleichungen eingeführt.

Die Resultate für die einzelnen linearen RD-Gleichungen können dann zum Beispiel durch Anwendung des Banach'schen Fixpunktsatzes verwendet werden, um Existenz und Eindeutigkeit von nichtlinearen RD-Gleichungen und RD-Systemen zu beweisen.

Der Fokus des folgenden Abschnittes liegt auf einem Existenzbeweis für die Variationsgleichung (2.17). In der Literatur findet sich eine generelle Behandlung von parabolischen partiellen Differentialgleichungen zum Beispiel in [Eva98], [Gaj74] oder [Zei90]. Für Probleme in dieser allgemeinen Form werden dort zwei verschiedene generelle Lösungsansätze vorgeschlagen: die *Halbgruppentheorie* [Zei90] sowie die *Galerkin-Methode* [Eva98], [Gaj74] und [Zei90].

Die Galerkin-Methode stellt zum Beweis der Existenz und Eindeutigkeit geringere Regularitätsanforderungen an die Eingangsdaten als die Halbgruppentheorie [Zei90][S.405], insbesondere bzgl. Quellterm und Dirichlet-Randbedingungen. Als Preis dafür ergeben sich bei der Galerkin-Methode geringfügig schwächere Regularitätseigenschaften der Lösung, jedoch sind diese hinreichend um physikalische Eigenschaften wie die Positivität der Lösung oder die Erhaltungsgleichung zu zeigen. Das Problem kann damit allgemeiner formuliert werden und wir erhalten einen größeren Spielraum bei der Modellierung von physikalischen Problemen. Wir verwenden aus diesem Grunde die Galerkin-Methode als Beweisverfahren.

3.1 Stetigkeit der Linear- und Bilinearform sowie Koerzivität

Lemma 3.1.1 (Stetigkeit der Linearform). *Es gilt tatsächlich $L \in L^2(0, T; V^*)$.*

Beweis.

$$\begin{aligned}
|L(t)[v]| &\leq \left| \int_{\Gamma_N} g_N \text{Tr}(v) dS \right| + \left| \int_{\Omega} w_t v dx \right| + D \left| \int_{\Omega} \nabla w \nabla v dx \right| + \left| \int_{\Omega} k w v dx \right| \\
&\leq \|g_N\|_{L^2(\partial\Omega)} \|v\|_{L^2(\partial\Omega)} + (\|w_t\|_{V^*} + D \|\nabla w\|_{L^2(\Omega, \mathbb{R}^d)} + \|k\|_{L^\infty(0, T; L^\infty(\Omega))} \|w\|_{L^2(\Omega)}) \|v\|_{L^2(\Omega)} \\
&\text{Stetigkeit des Spuoperators:} \\
&\leq C \|g_N\|_{L^2(\partial\Omega)} \|v\|_{L^2(\Omega)} + (\|w_t\|_{V^*} + D \|\nabla w\|_{L^2(\Omega)} + \|k\|_{L^\infty(0, T; L^\infty(\Omega))} \|w\|_{L^2(\Omega)}) \|v\|_{L^2(\Omega)} \\
&\leq (C \|g_N\|_{L^2(\partial\Omega)} + \|w_t\|_{V^*} + D \|\nabla w\|_{L^2(\Omega, \mathbb{R}^d)} + \|k\|_{L^\infty(0, T; L^\infty(\Omega))} \|w\|_{L^2(\Omega)}) \|v\|_V.
\end{aligned}$$

Es folgt also insbesondere $\|L(t)\|_{V^*} \leq C \|g_N\|_{L^2(\partial\Omega)} + \|w_t\|_{V^*} + D \|\nabla w\|_{L^2(\Omega)} + \|k\|_{L^\infty(0, T; L^\infty(\Omega))} \|w\|_{L^2(\Omega)}$. Für die $L^2(0, T; V^*)$ -Norm ergibt sich damit:

$$\begin{aligned}
\|L\|_{L^2(0, T; V^*)} &= \left(\int_0^T \left(C \|g_N\|_{L^2(\partial\Omega)} + \|w_t\|_{V^*} + D \|\nabla w\|_{L^2(\Omega, \mathbb{R}^d)} + \|k\|_{L^\infty(0, T; L^\infty(\Omega))} \|w\|_{L^2(\Omega)} \right)^2 dt \right)^{\frac{1}{2}} \\
&\leq \left(\int_0^T 4(C^2 \|g_N\|_{L^2(\partial\Omega)}^2 + \|w_t\|_{V^*}^2 + D^2 \|\nabla w\|_{L^2(\Omega, \mathbb{R}^d)}^2 + \|k\|_{L^\infty(0, T; L^\infty(\Omega))}^2 \|w\|_2^2) dt \right)^{\frac{1}{2}} \\
&\leq 2 \left(\int_0^T C \|g_N\|_{L^2(\partial\Omega)}^2 dt + \int_0^T \|w_t\|_{V^*}^2 dt + D \int_0^T \|\nabla w\|_{L^2(\Omega, \mathbb{R}^d)}^2 dt + \int_0^T \|k\|_{L^\infty(0, T; L^\infty(\Omega))}^2 \|w\|_2^2 dt \right)^{\frac{1}{2}} \\
&\leq 2C \|g_N\|_{L^2(0, T; L^2(\partial\Omega))} + 2\|w_t\|_{L^2(0, T; V^*)} \\
&\quad + 2D \|\nabla w\|_{L^2(0, T; L^2(\Omega, \mathbb{R}^d))} + 2\|k\|_{L^\infty(0, T; L^\infty(\Omega))} \|w\|_{L^2(0, T; L^2(\Omega))}. \tag{3.1}
\end{aligned}$$

□

Lemma 3.1.2. Die Form $a : V \times V \times [0, T] \rightarrow \mathbb{R}$ ist in den ersten zwei Argumenten stetig und koerziv mit Konstanten unabhängig von $t \in [0, T]$. Das heißt

$$|a(u, v; t)| \leq C \|u\|_V \|v\|_V \quad (\text{Stetigkeit}), \tag{3.2}$$

$$|a(u, v; t)| \geq \alpha \|v\|_{H^1(\Omega)}^2 \quad (\text{Koerzivitat}), \tag{3.3}$$

fur Konstanten $C > 0$ und $\alpha > 0$ unabhangig von $t \in [0, T]$.

Beweis. (i) Stetigkeit:

$$\begin{aligned}
|a(u, v; t)| &= \left| D \int_{\Omega} \nabla u \nabla v dx + k(t) \int_{\Omega} u v dx \right| \\
&\leq D \|\nabla u\|_{L^2(\Omega, \mathbb{R}^d)} \|\nabla v\|_{L^2(\Omega, \mathbb{R}^d)} + \|k\|_{L^\infty(0, T; L^\infty(\Omega))} \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\
&\leq \max\{D, k\} \|u\|_V \|v\|_V.
\end{aligned}$$

(ii) Koerzivitat: Wir wenden die allgemeine Poincar -Ungleichung [Alt06, S. 240] auf V an:

$$\begin{aligned} |a(v, v; t)| &= D \|\nabla v\|_{L^2(\Omega, \mathbb{R}^d)}^2 + k(t) \|u\|_2^2 \\ &\geq D \|\nabla v\|_{L^2(\Omega, \mathbb{R}^d)}^2 \\ &\geq C \|v\|_{H^1(\Omega)}^2 \text{ f r eine Konstante } C > 0. \end{aligned}$$

F r Details siehe [Alt06, S. 246], Beispiel 5. □

3.2 Existenz und Eindeutigkeit f r eine Reaktions-Diffusions-Gleichung

Im Folgenden wollen wir einen Satz zur Existenz und Eindeutigkeit von Evolutionsgleichungen formulieren und diesen mit Hilfe einer Galerkin-Approximation beweisen. Ein ausf hrlicher Existenz- und Eindeutigkeitsbeweis f r das Variationsproblem (2.17) findet sich in [Zei90, S.430ff]. Wir nehmen hier der Einfachheit halber an, dass die Funktion f r den Reaktionsterm k (s. Annahmen(A), Kap. 2.4) eine Konstante ist mit $k \geq 0$.

Satz 3.2.1. *Sei $V \subset H \subset V^*$ ein Evolutionstriplet und $a : V \times V \rightarrow \mathbb{R}$ eine koerzive und stetige Bilinearform. Desweiteren sei $b \in L^2(0, T; V^*)$. Dann hat die Variationsgleichung*

$$\frac{d}{dt}(u(t), v)_H + a(u(t), v) = \langle b(t), v \rangle_V \text{ f r alle } v \in V, \text{ f r fast alle } t \in (0, T), \quad (3.4)$$

$$u(0) = u_0 \in H \quad (3.5)$$

genau eine L sung $u \in H^1(0, T; V, H)$.

Beweis. (i) **Anwendung der Galerkin-Approximation.** Da V separabel ist, gibt es eine abzahlbare Teilmenge $\{w_i | i \in \mathbb{N}\} \subset V$, die dicht in V liegt und linear unabhangig ist. Mit V_m bezeichnen wir die endlichen Mengen $V_m := \text{span}\{w_1, \dots, w_m\}$. Da V dicht in H liegt, gibt es ein $u_{0m} \in V_m$, $u_{0m} := \sum_{k=1}^m \alpha_{km} w_k$ (mit $\alpha_{km} \in \mathbb{R}$), so dass $u_{0m} \rightarrow u_0$ in H f r $m \rightarrow \infty$. Wir suchen Funktionen $u_m(t) := \sum_{k=1}^m c_{km}(t) w_k$, so dass f r fast alle $t \in (0, T)$ gilt

$$\frac{d}{dt}(u_m(t), v)_H + a(u_m(t), v) = \langle b(t), v \rangle_V \text{ f r alle } v \in V_m, \text{ f r fast alle } t \in (0, T), \quad (3.6)$$

$$u_m(0) = u_{0m}. \quad (3.7)$$

Formulieren wir (3.6), (3.7) explizit so erhalten wir

$$\sum_{k=1}^m [c'_{km}(w_k, w_j)_H + c_{km}(t)a(w_k, w_j)] = \langle b(t), v \rangle_V \text{ für fast alle } t \in (0, T), \quad (3.8)$$

$$c_{jm}(0) = \alpha_{jm}, \quad 1 \leq j \leq m. \quad (3.9)$$

Die Gleichungen in (3.8) und (3.9) bilden ein Anfangswertproblem für ein implizites System von m linearen gewöhnlichen Differentialgleichungen für die unbekannten Funktionen c_{km} , $1 \leq k \leq m$. Die Funktionen $t \mapsto \langle b(t), w_j \rangle$ auf der rechten Seite liegen in $L^2(0, T)$. Da $\{w_1, \dots, w_m\}$ linear unabhängig sind, gilt für die Gram-Determinante $\det\{(w_k, w_j)_H\} \neq 0$. Damit können die Gleichungen in (3.8) nach c'_{km} aufgelöst werden und wir erhalten ein explizites System. Nach dem Satz von Carathéodory für Systeme linearer gewöhnlicher Differentialgleichungen mit Lebesgue-integrierbarer rechter Seite existiert eine eindeutige Lösung $(c_{1m}, \dots, c_{mm}) \in (H^1(0, T_m))^m$ für gewisse $T_m \in (0, T]$. Dann löst $u_m(t) := \sum_{k=1}^m c_{km}(t)w_k$ die Galerkin-Gleichungen (3.6), (3.7) und es gilt $u_m \in H^1(0, T; V, H)$, $m \in \mathbb{N}$.

(ii) **A priori-Schranken für $\{u_m\}$.** Durch Einsetzen von $v = u_m(t) \in V_m$ in (3.6) erhalten wir nach Integration über $[0, t]$, $0 < t < T$

$$\|u_m(t)\|_H^2 + \int_0^t a(u_m(s), u_m(s))ds = \|u_{0m}\|_H^2 + \int_0^t \langle b(s), u_m(s) \rangle_V ds.$$

Da die Bilinearform $a(., .)$ koerziv ist, gibt es ein $\alpha > 0$ mit $a(v, v; t) \geq \alpha \|v\|_V^2$. Es folgt also

$$\|u_m(t)\|_H^2 + \alpha \int_0^t \|u_m(s)\|_V^2 ds \leq \|u_{0m}\|_H^2 + \int_0^t \|b(s)\|_{V^*} \|u_m(s)\|_V ds.$$

Anwendung der Young'schen Ungleichung ergibt

$$\|u_m(t)\|_H^2 + \alpha \int_0^t \|u_m(s)\|_V^2 ds \leq \|u_{0m}\|_H^2 + \frac{1}{2\alpha} \int_0^t \|b(s)\|_{V^*}^2 ds + \frac{\alpha}{2} \int_0^t \|u_m(s)\|_V^2 ds.$$

Nun ist $\|u_{0m}\|$ als konvergente Folge beschränkt und $b \in L^2(0, T; V^*)$. Es gibt also eine Konstante $C_1 > 0$ mit

$$\|u_m(t)\|_H^2 + \frac{\alpha}{2} \int_0^t \|u_m(s)\|_V^2 ds \leq C_1.$$

Folglich gilt $T_m = T$ für alle $m \in \mathbb{N}$ und es gibt eine Konstante $C_2 > 0$ mit

$$\|u_m\|_{L^\infty(0, T; H)} + \|u_m\|_{L^2(0, T; V)} \leq C_2. \quad (3.10)$$

Sei nun $A : V \rightarrow V^*$, $\langle Au, v \rangle_V := a(u, v)$ für $v \in V$, die Abbildung $A \in \mathcal{L}(V, V^*)$, welche

durch a induziert wird. Wir betrachten dann die Abbildung $\tilde{A} : L^2(0, T; V) \rightarrow L^2(0, T; V^*)$, definiert durch

$$\langle \tilde{A}u, v \rangle_{L^2(0, T; V)} := \int_0^T \langle Au(t), v(t) \rangle_V dt, \quad v \in L^2(0, T; V).$$

Dann ist $\tilde{A} \in \mathcal{L}(L^2(0, T; V), L^2(0, T; V^*))$, denn es gilt

$$|\langle \tilde{A}u, v \rangle_L| \leq \int_0^T |a(u(t), v(t))| dt \leq C_3 \|u\|_{L^2(0, T; V)} \|v\|_{L^2(0, T; V)}.$$

Mit Ungleichung (3.10) erhalten wir

$$\|\tilde{A}u_m\|_{L^2(0, T; V^*)} \leq C_4 \quad \text{für alle } m \in \mathbb{N}.$$

Aufgrund der Reflexivität der involvierten Räume gibt es $u \in L^2(0, T; V)$, $w \in L^2(0, T; V^*)$, $z \in H$, so dass nach Übergang zu einer entsprechenden Teilfolge gilt

$$u_m \rightarrow u \text{ schwach in } L^2(0, T; V), \quad (3.11)$$

$$u_m(T) \rightarrow z \text{ schwach in } H, \quad (3.12)$$

$$\tilde{A}u_m \rightarrow w \text{ schwach in } L^2(0, T; V^*), \quad (3.13)$$

$$u_m(0) \rightarrow u_0 \text{ stark in } H. \quad (3.14)$$

(iii) Wir zeigen, dass $u \in H^1(0, T; V, H)$ und

$$\begin{aligned} u' + w &= b, \\ u(0) &= u_0, u(T) = z. \end{aligned}$$

Zunächst gilt für alle $\psi \in C^\infty[0, T]$, $v \in V_m$, dass $\psi v_m \in H^1(0, T; V, H)$. Es folgt

$$\begin{aligned} & (u_m(T), \psi(T)v)_H - (u_m(0), \psi(0)v)_H \\ &= \int_0^T \{ \langle w'_m(t), \psi(t)v \rangle_V + \langle \psi'(t)v, u_m(t) \rangle_V \} dt \\ &= \int_0^T \{ \langle b(t) - Au_m(t), \psi(t)v \rangle + \langle \psi'(t)v, u_m(t) \rangle_V \} dt. \end{aligned}$$

Aus (3.11)-(3.14) können wir folgern

$$(z, \psi(T)v)_H - (u_0, \psi(T)v)_H = \int_0^T \{ \langle b(t) - w(t), \psi(t)v \rangle_V + \langle \psi'(t)v, u(t) \rangle_V \} dt. \quad (3.15)$$

Nun ist $\bigcup_{m=1}^{\infty} V_m$ dicht in V . Für jedes $v \in V$ gibt es also ein $v_m \in V$, so dass

$$\psi v_m \rightarrow \psi v \text{ in } L^2(0, T; V), \quad \psi' v_m \rightarrow \psi' v \text{ in } L^2(0, T; V^*).$$

Folglich gilt Gl. (3.15) auch für alle $\psi \in \mathcal{C}_0^\infty(0, T)$, $v \in V$:

$$\int_0^T \langle b(t) - w(t), v \rangle \psi(t) dt = - \int_0^T (v, u(t))_H \psi'(t) dt.$$

Es gibt also ein $u' \in V^*$ mit $u' = b - w$. Außerdem gibt es wegen Gl. (3.15) ein $v \in V$, $\psi \in \mathcal{C}^\infty[0, T]$ mit $\psi(T) = 1$ und $\psi(0) = 0$, also:

$$\begin{aligned} (z, v)_H &= \int_0^T \{ \langle u'(t), \psi(t)v \rangle_V + \langle \psi'(t)v, u(t) \rangle_V \} \\ &= (u(T), \psi(T)v)_H - (u(0), \psi(0)v)_H \\ &= (u(T), v)_H. \end{aligned}$$

Demnach gilt $z = u(T)$. Setzen wir $\psi(0) = 1$ und $\psi(T) = 0$, so erhalten wir analog $u(0) = u_0$.

(iv) **Existenz.** Wir zeigen, dass $\tilde{A}u = w$. Dann folgt $u' + \tilde{A}u = b$ und die Existenz einer Lösung ist bewiesen. Es gilt

$$\begin{aligned} u_m &\rightarrow u \text{ schwach in } L^2(0, T; V) \\ \Rightarrow \tilde{A}u_m &\rightarrow \tilde{A}u \text{ schwach in } L^2(0, T; V^*) \\ \Rightarrow \tilde{A}u &= w. \end{aligned}$$

(v) **Eindeutigkeit.** Seien $u_1, u_2 \in H^1(0, T; V, H)$ zwei Lösungen zu (3.4), (3.5). Dann gilt für alle $\phi \in V$ und für fast alle $t \in (0, T)$

$$\begin{aligned} \frac{d}{dt}(u_1(t) - u_2(t), \phi)_H + a(u_1(t) - u_2(t), \phi) &= 0, \\ (u_1 - u_2)(0) &= 0. \end{aligned}$$

Setzen wir $\phi = u_1(t) - u_2(t)$ und integrieren über $[0, T]$, so erhalten wir

$$\begin{aligned} 0 &= \|u_1(T) - u_2(T)\|_H^2 + \int_0^T a(u_1(t) - u_2(t), u_1(t) - u_2(t)) dt \\ &\geq \|u_1(T) - u_2(T)\|_H^2 + \alpha \|u_1 - u_2\|_{L^2(0,T;V)}^2. \end{aligned} \quad (3.16)$$

Somit gilt $u_1(t) = u_2(t)$ fast überall auf $(0, T)$. □

Korollar. *Unter den Annahmen (A), Kap. 2.4, liefert der vorangehende Satz Existenz und Eindeutigkeit für das Modellproblem (2.17) in der schwachen Form.*

3.3 Positivität, Beschränktheit und eine Erhaltungsgleichung

Eine Lösung $u \in H^1(0, T; H^1(\Omega))$ des Modellproblems (2.17) in der schwachen Form ist nicht mehr unbedingt zweimal stetig differenzierbar, wie im klassischen Sinne in den Gleichungen (2.10)-(2.13) angenommen. Sie muss nur noch einmal schwach differenzierbar sein. Es kann daher durchaus Lösungen der Gleichung (2.17) geben, auch wenn die Gleichungen ((2.10)-(2.13) im klassischen Sinne nicht lösbar sind. Dennoch ist eine schwache Lösung auch physikalisch sinnvoll. Einige wichtige Eigenschaften sollen deshalb hier bewiesen werden. Wir beginnen mit der Positivität der Lösungen unserer linearen RD-Gleichung. Der Beweis folgt der Methode von Stampacchia [Per10] auf unbeschränkten Gebieten. Der Beweis wurde hier auf den Fall eines beschränkten Gebietes mit Dirichlet- und Neumann-Randbedingungen angepasst.

Satz 3.3.1 (Positivität). *Es gelten die Annahmen (A), Kap. 2.4, und die Randdaten erfüllen $g_D \geq 0$ fast überall auf $\Gamma_D \times (0, T)$ sowie $g_N \geq 0$ fast überall auf $\Gamma_N \times (0, T)$. Für den Quellterm gilt $f \geq 0$ in $\Omega \times (0, T)$ und die Startfunktion u^0 ist ebenfalls positiv in Ω . Dann folgt für die Lösung u von (2.17)*

$$u \geq 0 \text{ in } \Omega \times [0, T]. \quad (3.17)$$

Beweis. (nach Stampacchia):

Wir setzen $p = -u$. Dann gilt für p die schwache Formulierung

$$\begin{aligned} \frac{d}{dt} \int_{\Omega} p v dx + \int_{\Omega} \nabla p \nabla v dx - \int_{\partial\Omega} \frac{\partial p}{\partial n} \text{Tr}(v) dS + \int_{\Omega} k p v dx &= - \int_{\Omega} f v dx \\ \text{für alle } v \in H^1(\Omega) \text{ und für fast alle } t \in (0, T). \end{aligned}$$

Die Testfunktion v substituieren wir durch $(p)_+ := \max(0, p)$ und erhalten

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} (p(x, t))_+^2 dx + \int_{\Omega} |\nabla(p)_+|^2 dx - \int_{\partial\Omega} \frac{\partial p}{\partial n} \text{Tr}((p)_+) dS + \int_{\Omega} k \cdot (p)_+^2 dx = - \int_{\Omega} f \cdot (p)_+ dx$$

für fast alle $t \in (0, T)$.

Offensichtlich gilt fast überall auf $\Gamma_D \times (0, T)$: $\text{Tr}((p)_+) = 0$. Wir ersetzen zudem $\frac{\partial p}{\partial n}$ durch $-g_N$

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} (p(x, t))_+^2 dx + \int_{\Omega} |\nabla(p)_+|^2 dx + \int_{\Gamma_N} g_N \text{Tr}((p)_+) dS + \int_{\Omega} k \cdot (p)_+^2 dx = - \int_{\Omega} f \cdot (p)_+ dx$$

für fast alle $t \in (0, T)$.

Nun gilt fast überall auf $\Gamma_N \times (0, T)$: $g_N \geq 0$, also

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} (p(x, t))_+^2 dx \leq -\|k\|_{\infty} \int_{\Omega} (p)_+^2 dx \text{ für fast alle } t \in (0, T).$$

Es folgt

$$\int_{\Omega} (p(x, t))_+^2 dx \leq e^{-2 \int_0^t \|k(s)\|_{\infty} ds} \int_{\Omega} (p^0(x))_+^2 dx \text{ für } t \in (0, T).$$

Nach Annahme ist $u^0 \geq 0$, woraus $\int_{\Omega} (p^0(x))_+^2 dx = 0$ folgt und damit auch $\int_{\Omega} (p(x, t))_+^2 dx = 0$ für alle $t \in (0, T)$. Das heißt $u(t, x) \geq 0$ in $\Omega \times [0, T]$. □

Neben der wichtigen Positivitätseigenschaft können wir noch eine Erhaltungsgleichung ableiten.

Bemerkung 3.3.1 (Erhaltungsgleichung). Sei u eine Lösung von (2.17). Wir setzen in (2.14) $v \equiv 1$ in Ω und bekommen die Erhaltungsgleichung

$$\frac{d}{dt} \int_{\Omega} u dx = \int_{\partial\Omega} \frac{\partial u}{\partial n} dS + \int_{\Omega} (f - ku) dx. \quad (3.18)$$

$$(3.19)$$

Zuletzt wollen wir in diesem Abschnitt noch eine Beschränktheitseigenschaft zeigen.

Satz 3.3.2 (Beschränktheit). *Es gelten die Annahmen (A), Kap. 2.4, und die Randdaten erfüllen $g_D \leq C$ fast überall auf $\Gamma_D \times (0, T)$ und $g_N \leq 0$ fast überall auf $\Gamma_N \times (0, T)$. Für den Quellterm gilt $f \leq 0$ in $\Omega \times (0, T)$ und die Startfunktion ist beschränkt durch $u^0 \leq C$*

in Ω . Dann folgt für die Lösung u von (2.17):

$$u \leq C \text{ in } \Omega \times [0, T].$$

Beweis. Wir wenden den Satz über die Positivität von Lösungen an. $\tilde{u} = -u + C$ ist eine Lösung des Problems

$$\frac{d}{dt} \int_{\Omega} \tilde{u} v dx + \int_{\Omega} \nabla \tilde{u} \nabla v dx - \int_{\partial\Omega} \frac{\partial \tilde{u}}{\partial n} \text{Tr}(v) dS + \int_{\Omega} k \tilde{u} v dx = - \int_{\Omega} f v dx$$

für alle $v \in H^1(\Omega)$ und für fast alle $t \in (0, T)$,

$$\tilde{u} = -g_D + C \quad \text{auf } (0, T) \times \Gamma_D,$$

$$\frac{\partial \tilde{u}}{\partial n} = -g_N \quad \text{auf } (0, T) \times \Gamma_N,$$

$$\tilde{u} = -u^0 + C \quad \text{auf } \{0\} \times \Omega.$$

Der Satz über Positivität liefert uns dann $\tilde{u} = -u + C \geq 0$, also $u \leq C$. □

4 Nichtlineare Reaktions-Diffusions-Gleichungen und Systeme

4.1 Der Banach'sche Fixpunktsatz

Im Folgenden wollen wir Systeme von Reaktions-Diffusions-Gleichungen mit Hilfe des Banach'schen Fixpunktsatzes auf Lösbarkeit untersuchen. Für den Beweis folgen wir den Ausführungen in [Eva98, S. 498ff]. Wir betrachten folgendes System von Reaktions-Diffusions-Gleichungen, der Einfachheit halber setzen wir Dirichlet-Nullrandbedingungen voraus:

$$u_t - D \cdot \Delta u = f(u) \text{ in } \Omega \times [0, T], \quad (4.1)$$

$$u = 0 \text{ auf } \partial\Omega \times [0, T], \quad (4.2)$$

$$u = u^0 \text{ auf } \Omega \times \{t = 0\}. \quad (4.3)$$

Dabei sind $u = (u_1, \dots, u_m)$, $u^0 = (u_1^0, \dots, u_m^0)$ Vektoren und $D = \text{diag}(D_1, \dots, D_m)$. Die Diffusionskonstanten D_i sind alle positiv, also $D_i > 0, i = 1, \dots, m$. Das Gebiet $\Omega \in \mathbb{R}^n$ ist offen und beschränkt mit Lipschitz-Rand. Die Zeit $T > 0$ wird fixiert. Wir nehmen an, dass $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$ Lipschitz-stetig ist und die Anfangswertfunktion u^0 in $H_0^1(\Omega; \mathbb{R}^m)$ liegt.

Satz 4.1.1. *Das oben gegebene System hat genau eine Lösung.*

Beweis. (i) Wir wollen den Banach'schen Fixpunktsatz auf den folgenden metrischen Raum anwenden:

$$X = \mathcal{C}([0, T]; L^2(\Omega; \mathbb{R}^m)),$$

mit der Norm

$$\|v\| := \max_{0 \leq t \leq T} \|v(t)\|_{L^2(\Omega; \mathbb{R}^2)}.$$

Wir definieren uns einen Operator $A : X \rightarrow X$ wie folgt: Sei u eine beliebige Funktion aus X . Wir setzen $h(t) := f(u(t))$, $(0 \leq t \leq T)$. Es folgt: $h \in L^2(0, T; L^2(\Omega; \mathbb{R}^m))$. Damit hat

nach dem Existenzsatz 3.2.1 die lineare parabolische partielle Differentialgleichung

$$\begin{aligned} w_t - D \cdot \Delta w &= f(u) \text{ in } \Omega \times [0, T], \\ w &= 0 \text{ auf } \partial\Omega \times [0, T], \\ w &= u^0 \text{ auf } \Omega \times \{t = 0\}. \end{aligned}$$

eine eindeutige Lösung

$$w \in H^1(0, T; H_0^1(\Omega; \mathbb{R}^m)) \subset X.$$

Folglich erfüllt $w \in X$ die Gleichung

$$\langle w'_i, v \rangle + D_i \int_{\Omega} \nabla w_i \cdot \nabla v dx = (h_i, v)_{\Omega} \text{ für } i = 1, \dots, m, \text{ fast überall in } [0, T] \quad (4.4)$$

für alle $v \in H_0^1(\Omega)$ und der Startbedingung $w(0) = u^0$. Wir definieren $A : X \rightarrow X$ durch $A[u] = w$.

(ii) Wir wollen zeigen, dass A eine strikte Kontraktion ist, wenn T klein genug gewählt wird. Um dies zu beweisen, wählen wir $u, \tilde{u} \in X$ und definieren $w = A[u], \tilde{w} = A[\tilde{u}]$ wie oben. Damit erfüllt die Iterierte w das Gleichungssystem (4.4) für $h = f(u)$ und \tilde{w} erfüllt dieses Gleichungssystem für $\tilde{h} := f(\tilde{u})$. In der schwachen Formulierung setzen wir die Testfunktion $v = w - \tilde{w}$ und es folgt:

$$\frac{1}{2} \frac{d}{dt} \|w - \tilde{w}\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \min_{i=1, \dots, m} \{D_i\} \|\nabla w - \nabla \tilde{w}\|_{L^2(\Omega; \mathbb{R}^{n \times m})}^2 \leq (w - \tilde{w}, h - \tilde{h}).$$

Anwendung der Poincaré-Ungleichung ergibt

$$\frac{d}{dt} \|w - \tilde{w}\|_{L^2(\Omega; \mathbb{R}^m)}^2 + 2C_1 \min_{i=1, \dots, m} \{D_i\} \|w - \tilde{w}\|_{H_0^1(\Omega; \mathbb{R}^m)}^2 \leq 2(w - \tilde{w}, h - \tilde{h}).$$

Bei Anwendung der Hölder- und Young'schen Ungleichung folgt

$$\begin{aligned} \frac{d}{dt} \|w - \tilde{w}\|_{L^2(\Omega; \mathbb{R}^m)}^2 + 2C_1 \min_{i=1, \dots, m} \{D_i\} \|w - \tilde{w}\|_{H_0^1(\Omega; \mathbb{R}^m)}^2 \\ \leq \epsilon \|w - \tilde{w}\|_{L^2(\Omega; \mathbb{R}^m)}^2 + \frac{1}{\epsilon} \|f(u) - f(\tilde{u})\|_{L^2(\Omega; \mathbb{R}^m)}^2 \\ \leq \epsilon \|w - \tilde{w}\|_{H_0^1(\Omega; \mathbb{R}^m)}^2 + \frac{1}{\epsilon} \|f(u) - f(\tilde{u})\|_{L^2(\Omega; \mathbb{R}^m)}^2. \end{aligned}$$

Nach Voraussetzung ist $\min_{i=1, \dots, m} \{D_i\} > 0$ und wir können $\epsilon = 2C_1 \min_{i=1, \dots, m} \{D_i\}$ wählen. Damit erhalten wir

$$\frac{d}{dt} \|w - \tilde{w}\|_{L^2(\Omega; \mathbb{R}^m)}^2 \leq C_2 \|f(u) - f(\tilde{u})\|_{L^2(\Omega; \mathbb{R}^m)}^2.$$

Da f Lipschitz-stetig ist, gilt

$$\begin{aligned} \|w(s) - \tilde{w}(s)\|_{L^2(\Omega; \mathbb{R}^m)}^2 &\leq C_3 \int_0^s \|u - \tilde{u}\|_{L^2(\Omega; \mathbb{R}^m)}^2 dt \\ &\leq C_3 T \max_{0 \leq s \leq T} \|u(s) - \tilde{u}(s)\|^2. \end{aligned} \quad (4.5)$$

für alle $0 \leq s \leq T$. Maximieren wir auch die linke Seite bezüglich s , bekommen wir:

$$\|w - \tilde{w}\|^2 \leq C_3 T \|u - \tilde{u}\|^2.$$

Somit gilt

$$\|A[u] - A[\tilde{u}]\| \leq (C_3 T)^{\frac{1}{2}} \|u - \tilde{u}\|.$$

Gilt nun $T < \frac{1}{C_3}$ so folgt $(C_3 T)^{\frac{1}{2}} = \gamma < 1$. Damit ist A eine Kontraktion. Der Fall $T \geq \frac{1}{C_3}$ wird im nächsten Punkt diskutiert.

(iii) Bei einem beliebigen $T > 0$ können wir das Problem aufteilen. Die Konstante C_3 ist nämlich unabhängig von der Zeit und der gewählten Startfunktion. Wir wählen $\bar{T} > 0$ so klein, dass $(C_3 \bar{T})^{\frac{1}{2}} < 1$. Wir können den Banach'schen Fixpunktsatz anwenden, um eine schwache Lösung u von unserem Problem zu finden, die auf dem Intervall $[0, \bar{T}]$ existiert. Da $u \in H^1([0, \bar{T}]; H^1(\Omega; \mathbb{R}^m))$ können wir das Problem neu definieren mit $u^0 = u(\bar{T})$. Wir können dann das Argument wiederholen und die Lösung auf das Intervall $[\bar{T}, 2\bar{T}]$ erweitern. Nach endlich vielen Schritten können wir so eine Lösung auf das ganze Zeitintervall $[0, T]$ erweitern.

(iv) Eindeutigkeit:

Nehmen wir an, wir hätten zwei schwache Lösungen u und \tilde{u} unseres Problems. Es gilt dann $w = u$ und $\tilde{w} = \tilde{u}$. Dies setzen wir in in Ungleichung (4.5) ein und erhalten

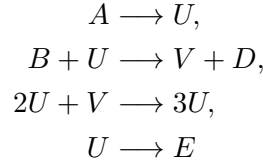
$$\|u(s) - \tilde{u}(s)\|_{L^2(\Omega; \mathbb{R}^m)}^2 \leq C_3 \int_0^s \|u - \tilde{u}\|_{L^2(\Omega; \mathbb{R}^m)}^2 dt.$$

für alle $0 \leq s \leq T$. Aus dem Lemma von Gronwall folgt dann: $u = \tilde{u}$.

□

4.2 Ein Beispiel: Der Brüsselator

Der Brüsselator beschreibt einen Fall, in dem die chemischen Reaktionen dem Schema



folgen. Dabei sind A , B , D , E , U und V chemische Verbindungen. Seien $u(t, x)$ und $v(t, x)$ Konzentrationen von U und V . Wir nehmen an, dass die Konzentrationen der Verbindungen A und B während des Reaktionsprozesses konstant gehalten werden. Sie werden mit den Konstanten a bzw. b bezeichnet. Aus dem oben stehenden Reaktions-Schema lässt sich dann folgendes System von zwei nichtlinearen gekoppelten Reaktions-Diffusions-Gleichungen ableiten:

$$\frac{\partial u}{\partial t} = d_1 \Delta u + au^2v - (b+1)u + 1 \text{ in } (0, T) \times \Omega, \quad (4.6)$$

$$\frac{\partial v}{\partial t} = d_2 \Delta v - au^2v + bu \text{ in } (0, T) \times \Omega, \quad (4.7)$$

$$u(t, x) = v(t, x) = 0, t > 0, x \in \partial\Omega,$$

$$u(0, x) = u_0(x), v(0, x) = v_0(x), x \in \Omega.$$

Um die Existenz dieses Gleichungssystems zu zeigen, wollen wir den Banach'schen Fixpunktsatz anwenden. Wir führen eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein, welche die Reaktionsterme der Brüsselatorgleichungen beschreibt:

$$f^1(x, y) := ax^2y - (b+1)x + 1, \quad (4.8)$$

$$f^2(x, y) := -ax^2y + bx. \quad (4.9)$$

Wir versuchen eine sogenannte *Trapping-Region* (s. Abbildung 4.1) für das Brüsselator-Modell zu finden. Als Trapping-Region K definieren wir uns folgendes Gebiet

$$K := \left\{ (x, y) \in \mathbb{R}^2 \mid x > \frac{1}{b+1}, y > 0, y < \frac{b(b+1)}{a} \text{ und } x + y < 2 + \frac{b(b+1)}{a} \right\}. \quad (4.10)$$

Auf den Rändern der Trapping-Region gilt

$$f^1\left(\frac{1}{b+1}, y\right) = -\frac{a}{(b+1)^2}y + 1 > 0 \text{ für } y < \frac{b(b+1)}{a} \quad (\text{linker Rand}),$$

$$f^2(x, 0) = bx > 0 \text{ für } x > \frac{1}{b+1} \quad (\text{unterer Rand}),$$

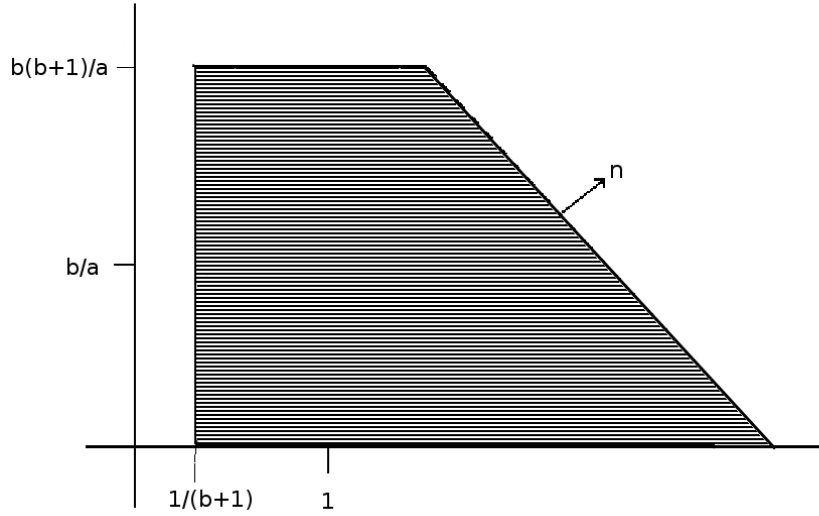


Abbildung 4.1: Trapping-Region für den Brusselator.

$$f^2(x, \frac{b(b+1)}{a}) = -x^2b(b+1) + bx = -bx \cdot (x(b+1) - 1) < 0 \text{ für } x > \frac{1}{b+1} \quad (\text{oberer Rand}),$$

$$\vec{n} \cdot (f^1(x, y), f^2(x, y)) = f^1(x, y) + f^2(x, y) = -x + 1 < 0 \text{ für } x > 1 \quad (\text{rechter Rand}).$$

Jede Lösung des Brusselatorsystems mit Startkonzentration innerhalb der Trapping-Region bleibt also dort gefangen. Wir definieren

$$f_+(x, y) = \begin{cases} f & \text{falls } (x, y) \in K, \\ 0 & \text{sonst.} \end{cases}$$

Wir sehen, dass f stetig ist. Damit ist f eingeschränkt auf K insbesondere Lipschitz-stetig. Wir wenden Theorem 4.1.1 auf

$$\frac{\partial u}{\partial t} = d_1 \Delta u + f_+^1(u, v) \text{ in } (0, T) \times \Omega, \quad (4.11)$$

$$\frac{\partial v}{\partial t} = d_2 \Delta v + f_+^2(u, v) \text{ in } (0, T) \times \Omega, \quad (4.12)$$

$$u(t, x) = v(t, x) = 0, t > 0, x \in \partial\Omega,$$

$$u(0, x) = u_0(x), v(0, x) = v_0(x), x \in \Omega,$$

an und erhalten die Existenz einer Lösung innerhalb der Trapping-Region.

Aus diesem Existenzsatz leitet sich insbesondere eine Möglichkeit zur Lösung von großen Gleichungssystemen ab. In jedem Iterationsschritt können die einzelnen Gleichungen getrennt berechnet werden. Es ergeben sich also Möglichkeiten zur Paralle-

lisierung. Eine ausführliche Behandlung der Dynamik des Brusselators und ähnlicher Reaktions-Diffusions-Systeme, wie dem *Gray-Scott-Modell*, dem *Glykolyse-Modell* oder *Schnackenberg-Modell* findet sich in [You07].

5 Das unstetige Galerkin-Verfahren

Das unstetige Galerkin-Verfahren ist eine Methode zur Approximation von Lösungen von partiellen Differentialgleichungen, welche seit über 30 Jahren verwendet und weiterentwickelt wird. Reed und Hill [RH73] führten 1973 die erste *Discontinuous Galerkin* (DG)-Methode für hyperbolische Probleme ein. Ebenfalls in den 1970ern wurden DG-Methoden für elliptische und parabolische Gleichungen unabhängig von mehreren Wissenschaftlern eingeführt. Die dort verwendeten Methoden werden *Interior Penalty* (IP)-Methoden genannt, dieser Begriff wurde 1976 erstmals von Douglas und Dupont [JD76] geprägt. Die IP-Methoden wurden unabhängig von den Methoden für hyperbolische Probleme weiterentwickelt und sind nach wie vor Forschungsgegenstand. Die IP-Methoden werden auch in dieser Arbeit untersucht. Eine allgemeine Einführung und Übersicht in die wichtigsten DG-Verfahren findet sich in der *Unified Analysis* [ABCM01] von Arnold, Brezzi, Cockburn und Marini aus dem Jahr 2001. Wir werden nachfolgend die Idee der DG-Verfahren darstellen.

5.1 Idee

Genau wie bei der klassischen FE-Methode zerlegt man das Gebiet, auf dem die partielle Differentialgleichung gelöst werden soll, in kleine Gitterelemente, meist in Dreiecke (2D) oder Tetraeder (3D). Bei den DG-Verfahren wählt man gewöhnlich für jedes Gitterelement eine lineare Funktion oder ein Polynom höherer Ordnung, welches die Lösung auf diesem Element approximieren soll. Die Koeffizienten der Polynome werden durch eine sogenannte *Variationsgleichung* bestimmt, welche verfahrensabhängig aufgestellt wird. Man erhält ein Gleichungssystem, welches nach den eben erwähnten Koeffizienten aufgelöst werden muss. Bei den DG-Methoden ist die Anzahl der Freiheitsgrade direkt proportional zur Anzahl der Gitterelemente. Es wird für gewöhnlich eine Monombasis zur Approximation der Lösung gewählt [Riv08, S. 35]. Hat man beispielsweise $N = 10000$ Gitterelemente, so beträgt die Anzahl der zu bestimmenden Koeffizienten bei der Approximation durch lineare Funktionen $N \times 3 = 30000$ (2D) bzw. $N \times 4 = 40000$ (3D) und bei quadratischen Polynomen $N \times 6 = 60000$ (2D) bzw. $N \times 10 = 100000$ (3D).

Im Gegensatz zur klassischen FE-Methode wird an den Elementübergängen keine Stetigkeit vorausgesetzt. Die Unstetigkeiten werden in der Variationsformulierung für die IP-Verfahren durch die *Interior Penalty* (IP)-Terme bestraft. Es werden drei verschiedene DG-Methoden vorgestellt: *Symmetric Interior Penalty Galerkin* (SIPG), *Incomplete Interior Penalty Galerkin* (IIPG) und *Nonsymmetric Interior Penalty Galerkin* (NIPG). Diese Verfahren unterscheiden sich in der Wahl der Strafterme. Besondere Aufmerksamkeit wer-

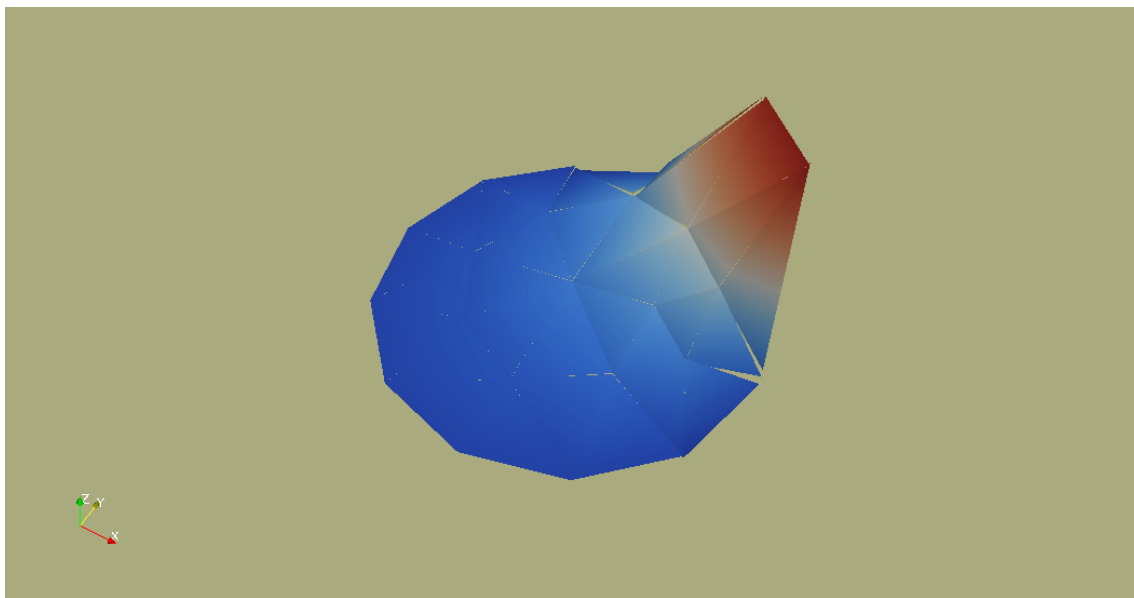


Abbildung 5.1: Im Bild sieht man eine Approximation auf einem groben Gitter, erzeugt durch das unstetige Galerkin-Verfahren. Unstetigkeiten werden bestraft, sind aber möglich. Wie in der Abbildung zu erkennen ist, werden insbesondere bei starken Gradienten kleine Sprünge zwischen den Elementen erlaubt.

den wir der NIPG-Methode widmen, welche sich besonders gut für diffusionsdominante Probleme eignet [BG99, Seite 2].

Analog zum funktionalanalytischen Teil werden wir Ort und Zeit getrennt behandeln. Zur Approximation wenden wir die sogenannte *Linienmethode* an. Das heißt, wir beginnen mit der Diskretisierung im Ort durch das unstetige Galerkin Verfahren und gehen dann zur Approximation in der Zeit über. Wir erhalten durch Aufstellung der Variationsgleichungen ein System aus gewöhnlichen Differentialgleichungen. Für die Fehleranalyse verwenden wir die Resultate aus [Fei09]. Wir werden zeigen, dass der Fehler in der $L^\infty(L^2)$ -Norm mit der Ordnung $\mathcal{O}(h^p)$ und $L^2(H^1)$ -Norm ebenfalls mit der Ordnung $\mathcal{O}(h^p)$ schrumpft. Dabei ist h der maximale Elementdurchmesser und p die Polynomordnung. Dies ist optimal für die NIPG-Methode. Wie bei der FE-Methode kann die Konvergenz für die SIPG-Methode in der $L^\infty(L^2)$ -Norm auf $\mathcal{O}(h^{p+1})$ [Riv08] verbessert werden, indem man die Strafparameter, welche die Unstetigkeiten bestrafen, geeignet wählt (*Superpenalization*) und weitere Regularitätsanforderungen an Gitter und analytische Lösung stellt. Dadurch ergeben sich jedoch zwei Nachteile. Zum Einen sind die Parameter gebietsabhängig und damit ist deren Wahl mit Schwierigkeiten behaftet; zum Anderen werden gewisse physikalische Eigenschaften der DG-Approximation zerstört, wodurch der Anwendungsvorteil auf parabolische Probleme verschwindet. Zum Schluss gehen wir über zur Analyse der Diskretisierung in der Zeit nach [RW00] über. Wir verwenden das sogenannte *Theta-Verfahren*,

| Methode | konsistent | stabil | Bedingung an Strafparameter | H^1 | L^2 |
|------------------------|------------|--------|-----------------------------|----------|------------------------|
| LDG[CS98] | ja | ja | $\sigma^0 > 0$ | $O(h^p)$ | $\mathcal{O}(h^{p+1})$ |
| SIPG [Arn82] | ja | ja | $\sigma^0 > \sigma^*$ | $O(h^p)$ | $\mathcal{O}(h^{p+1})$ |
| NIPG[BG99] | ja | ja | $\sigma^0 > 0$ | $O(h^p)$ | $\mathcal{O}(h^p)$ |
| IIPG [DSW04] | ja | ja | $\sigma^0 > \sigma^*$ | $O(h^p)$ | $\mathcal{O}(h^p)$ |
| Brezzi et al. [BMMP00] | nein | ja | $\sigma^0 \approx h^{-2p}$ | $O(h^p)$ | $\mathcal{O}(h^{p+1})$ |
| OBB(p=1) | ja | nein | - | - | - |
| OBB(p=2) [BO99] | ja | nein | - | $O(h^p)$ | $\mathcal{O}(h^{p+1})$ |
| Bassi-Rebay [BR97] | ja | nein | - | $O(h^p)$ | $\mathcal{O}(h^{p+1})$ |

Tabelle 5.1: Ein Vergleich der vorgestellten DG-Verfahren.

welches für verschiedene Parameter θ als Spezialfälle das *implizite Euler-Verfahren* und das *Crank-Nicolson-Verfahren* enthält. Die Fehleranalyse wird zeigen, dass wir generell in der $\ell^\infty(L^2)$ -Norm eine Konvergenzordnung von $\mathcal{O}(h^p) + \mathcal{O}(\tau)$ bzw. mit erhöhten Regularitätsbedingungen sogar $\mathcal{O}(h^p) + \mathcal{O}(\tau^2)$ mit der Crank-Nicolson-Methode erhalten.

5.2 Eine Übersicht

Es gibt DG-Methoden, die stabil und konsistent sind und in der L^2 -Norm, wie auch in der H^1 -Norm optimal konvergieren. Die Methoden unterscheiden sich vor allem in ihrer Anwendung auf bestimmte Problemklassen. Bassi und Rebay [BR97] führten 1990 eine DG-Methode ein, die speziell auf die kompressiblen Navier-Stokes-Gleichungen zugeschnitten wurde. Aufbauend auf seinen Ideen, wurde dann von Cockburn und Shu [CS98] die *Local Discontinuous Galerkin* (LDG)-Methode als Verallgemeinerung entwickelt. Aktuelle Untersuchungen beschäftigen sich mit der Lösung von rein elliptischen Gleichungen durch DG-Verfahren, hier sind einerseits die eben erwähnte LDG-Methode und auch die Methoden von Brezzi et al. [BMMP00] aufzuführen. Im Jahr 1999 wurde von Baumann und Oden [BO99] eine spezielle Methode für diffusionsdominante Probleme entwickelt. Wie das *Finite Volumen* (FE)-Verfahren gilt bei diesem Verfahren auf jedem Gitterelement die Massenerhaltung, jedoch ist diese Methode nicht stabil. Die stabile Version des OBB-Verfahrens ist das *Nonsymmetric Interior Penalty Galerkin* (NIPG)-Verfahren, welches wir in dieser Arbeit näher untersuchen. In Rivière [BG99] werden beide Methoden ausführlich behandelt. Das NIPG-Verfahren gehört zu den IP-Verfahren, zu denen auch das *Incomplete Interior Penalty Galerkin* (IIPG)-Verfahren [DSW04] und das *Symmetric Interior Penalty Galerkin* (SIPG)-Verfahren [Arn82] zu zählen sind. Eine Gegenüberstellung der erwähnten Verfahren findet sich in der folgenden Tabelle. Als vergleichendes Problem dient hier die Anwendung auf die Poisson-Gleichung

$$-\Delta u = f \text{ in } \Omega, u = 0 \text{ auf } \partial\Omega. \quad (5.1)$$

Ω ist hier ein konvexes polygonales Gebiet und f eine gegebene Funktion in $L^2(\Omega)$. Die ersten fünf Verfahren sind abhängig von einem Strafparameter σ^0 , welcher in die Bestrafung

der Unstetigkeiten an den Elementkanten einfließt. Bei dem SIPG- und IIPG-Verfahren hängt dieser insbesondere von dem betrachteten Gebiet ab. Deshalb ist die Berechnung dieser Parameter für diese Verfahren in der Praxis problematisch. Um die Konvergenzordnung und Stabilität bei der LDG- und NIPG-Methode einzuhalten, müssen die Strafparameter nur größer als Null gewählt werden. Wir konzentrieren uns hier auf das NIPG-Verfahren, welches insbesondere für diffusionsdominante Probleme geeignet ist. Die Analysis umfasst jedoch auch andere IP-Methoden, wie das SIPG- oder IIPG-Verfahren und stützt sich auf eine aktuelle Arbeit von Feistauer [DFKS08] aus dem Jahr 2008.

5.3 Motivation

Aufgrund ihrer physikalischen Genauigkeit sind die unstetigen Galerkin-Verfahren insbesondere für Diffusionsprobleme geeignet [Riv08, Seite 67]. Wir werden zeigen, dass diese Verfahren auf jedem einzelnen Gitterelement eine Form der Kontinuitätsgleichung erfüllen. Die DG-Verfahren wird wegen dieser Eigenschaft auch gern als Hybridmethode von gewöhnlichen FE-Methoden und *Finite Volumen* (FV)-Verfahren betrachtet. Mit der FV-Methode teilt das DG-Verfahren die erwähnte Erhaltungsgleichung, mit der FE-Methode die Verwendung von Polynomen zur Approximation und die Methoden zur Fehleranalyse. Bei den FE-Verfahren entstehen durch die Stetigkeitsbedingung ungewollte *numerische Flüsse* entlang der Elementkanten, so dass die Massenerhaltung nicht elementweise eingehalten werden kann.

In der Anwendung auf Modelle in der Praxis lässt das DG-Verfahren zusätzliche Spielräume. Im Gegensatz zum *Finite Differenzen* (FD)-Verfahren können Dirichlet, Neumann oder Robin-Randbedingungen ohne Schwierigkeiten verwendet werden. Die Unstetigkeit des Verfahrens erweitert das Anwendungsfeld zudem auf unstetige Diffusions- oder Reaktionskonstanten und vor allem auch auf unstetige Randbedingungen.

Für die Implementierung bietet das DG-Verfahren eine Reihe von Freiheiten. Das DG-Verfahren ermöglicht die Verwendung von *unstrukturierten* Gittern, also Gittern, deren Elemente unterschiedliche Form und Größe besitzen. Desweiteren unterliegt die Verwendung von *Hanging-Nodes*, also Ecken, die auf Kanten anderer Elemente liegen, keinen Einschränkungen. *High Polynomial*-Verfeinerungsmethoden können ebenfalls leicht implementiert werden. Man verwendet für gewöhnlich eine einfache Monombasis, die man elementweise definiert. Durch die fehlenden Stetigkeitsvoraussetzungen kann man den Polynomgrad von Element zu Element anders wählen und beliebig erhöhen.

In der Praxis ist es notwendig, für die DG-Verfahren geeignete Parameter für die Bestrafung der Unstetigkeiten zu wählen. Wir werden sehen, dass dieses Problem für das hier verwendete NIPG-Verfahren weitestgehend entfällt. Vorschläge für Parameter in der Praxis finden sich in [Riv08].

Für die Verwendung von Parallelisierungsalgorithmen gibt es weitere Vorteile. Ein wichtiger Unterschied zu den gewöhnlichen Finite-Element-Methoden ist nämlich, dass die resultierenden Gleichungen lokal bzgl. des erzeugenden Elementes aufgestellt werden. Gegenüber dem herkömmlichen FE-Verfahren bedeutet dies zunächst einen Mehraufwand in der Assemblierung, da einerseits mehr Freiheitsgrade existieren und auf der anderen

Seite durch die notwendigen IP-Terme zusätzliche Integrale ausgewertet werden müssen. Da jedoch jedes Element als separate Entität angesehen werden kann und nur über die Kantenflüsse mit den anderen verbunden ist, welche bei der Assemblierung gleich mitberechnet werden, ergeben sich entscheidende Vorteile für die Parallelisierung der Berechnungen [KEP⁺00]. Die Geschwindigkeitsvorteile durch die Parallelisierung sind erheblich [Yan96]. In Kapitel 6 werden wir einige Parallelisierungsalgorithmen vorstellen, welche auf dem DG-Verfahren basieren.

5.4 Diskretisierung im Ort

In den Kapiteln 2 und 3 haben wir eine lineare Reaktions-Diffusions-Gleichung funktionalanalytisch untersucht. Um die Gleichung mit numerischen Methoden zu behandeln, müssen wir das kontinuierliche Problem in geeigneter Weise diskretisieren. Analog zur Herangehensweise in der funktionalanalytischen Behandlung, werden wir Ort und Zeit getrennt betrachten. Wir verwenden die sogenannte *Linienmethode*. Das heißt, wir beginnen mit der Diskretisierung im Ort und gehen dann zur Diskretisierung in der Zeit über. Bleiben wir stetig in der Zeit und diskretisieren wir im Ort, so nennen wir diese eine *Semidiskretisierung*. Die Semidiskretisierung führt auf ein System von gewöhnlichen Differentialgleichungen. Diese wollen wir herleiten und deren Eigenschaften in den folgenden Abschnitten behandeln. Wir beginnen mit der Einführung einiger notwendiger Bezeichnungen und gehen dann zur Variationsformulierung über. Die verwendeten Bezeichnungen und die vorgestellten Verfahren halten sich an [DFKS08] und [Riv08].

5.4.1 Notationen

Das Gebiet Ω wird in geeigneter Weise in konvexe Polygone zerlegt. Der Rand $\partial\Omega$ wird dabei durch Kanten dieser Polygone approximiert. Mit \mathcal{T}_h , $h > 0$, bezeichnen wir eine solche gegebene endliche Partitionierung des Abschlusses $\overline{\Omega}$ in konvexe Polygone K mit disjunktem Inneren. Es ist darauf hinzuweisen, dass die Partitionierung keine Konformitätseigenschaften erfüllen muss, wie man es aus den stetigen FE-Methoden kennt. Als Maß für die Feinheit des Gitters dient die Größe der einzelnen Elemente. Wir definieren:

$h_K := \text{diam}(K)$ und

$$h := \max_{K \in \mathcal{T}_h} h_K$$

$\rho_K := \text{Radius der größten Kugel (bzw. Kreis in 2D), die in } K \text{ eingeschrieben werden kann.}$

Die Menge aller Kanten, aus denen die Polygone zusammengesetzt sind, bezeichnen wir mit \mathcal{F}_h . Der Rand $\partial\Omega$ wird durch die Kanten der angrenzenden Polygone beschrieben, wobei wir jede Kante entweder dem Dirichlet-Rand oder dem Neumann-Rand zuordnen. Wir bezeichnen die Menge aller Dirichlet-Kanten mit

$$\mathcal{F}_h^D := \{\Gamma \in \mathcal{F}_h; \Gamma \subset \Gamma_D\},$$

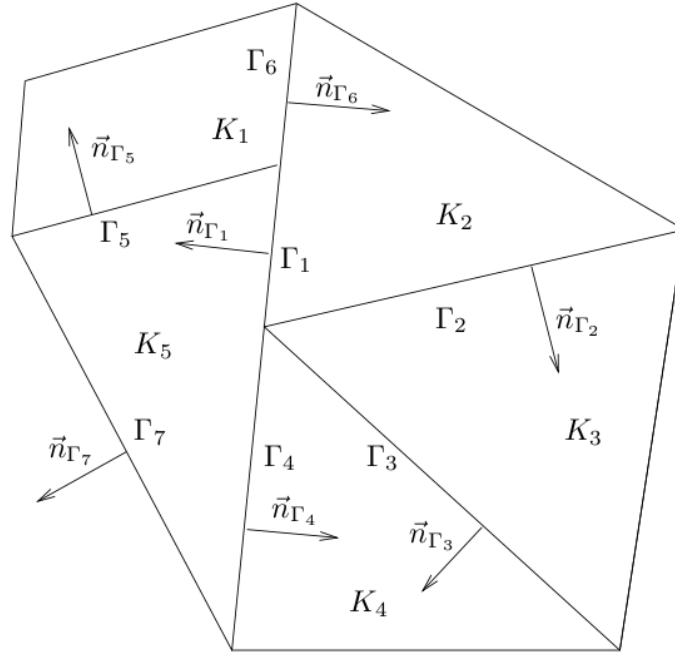


Abbildung 5.2: Elemente mit hängenden Knoten [Fei09].

die Menge aller Neumann-Kanten mit

$$\mathcal{F}_h^N := \{\Gamma \in \mathcal{F}_h; \Gamma \subset \Gamma_N\}$$

und die Menge der inneren Kanten mit

$$\mathcal{F}_h^I := \{\Gamma \in \mathcal{F}_h; \Gamma \subset \Omega\}.$$

Jeder Kante $\Gamma \in \mathcal{F}_h^I$ ordnen wir eine beliebige, aber von nun an festgewählte *Einheitsnormale* n_Γ zu. Für die äußeren Kanten $\Gamma \in \partial\Omega$ ist n_Γ die ins Gebietsäußere zeigende Einheitsnormale.

5.4.2 Gebrochene Sobolew-Räume

Das vorgestellte unstetige Galerkin-Verfahren lässt Unstetigkeiten entlang der Kanten zu. Daher können wir die gewohnten Sobolew-Räume als Funktionenräume nicht verwenden. Wir definieren deshalb den sogenannten *gebrochenen Sobolew-Raum* bezüglich einer Partitionierung \mathcal{T}_h :

$$H^k(\Omega, \mathcal{T}_h) := \{v; v|_K \in H^k(K) \text{ für alle } K \in \mathcal{T}_h\},$$

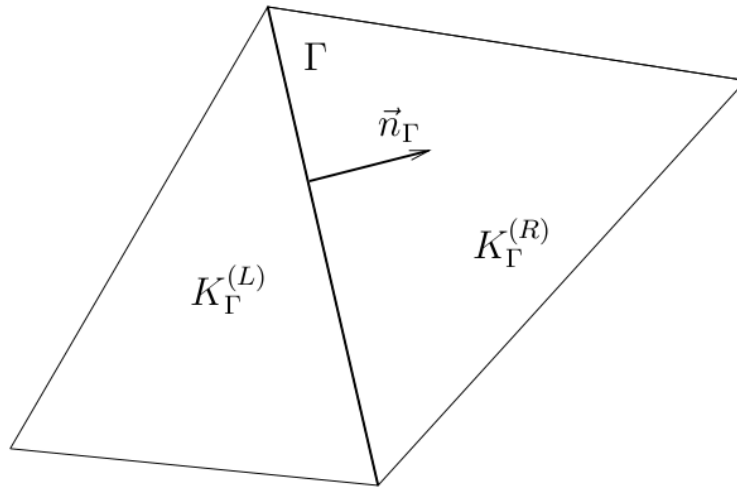


Abbildung 5.3: Zwei benachbarte Gitterelemente [Fei09].

mit der Norm

$$\|v\|_{H^k(\Omega, \mathcal{T}_h)} := \left(\sum_{K \in \mathcal{T}_h} \|v\|_{H^k(K)}^2 \right)^{\frac{1}{2}}$$

und der Seminorm

$$|v|_{H^k(\Omega, \mathcal{T}_h)} := \left(\sum_{K \in \mathcal{T}_h} |v|_{H^k(K)}^2 \right)^{\frac{1}{2}}.$$

5.4.3 Durchschnitte und Sprünge

Wir wollen die Unstetigkeiten der Funktionen aus den gebrochenen Sobolew-Räumen in geeigneter Weise beschreiben. Wir führen deshalb Durchschnitte bzw. Sprünge einer Funktion v aus dem gebrochenen Sobolew-Raum $H^k(\Omega, \mathcal{T}_h)$ entlang einer inneren Kante Γ ein. Für jedes $\Gamma \in \mathcal{F}_h^I$ gibt es zwei benachbarte Elemente $K_\Gamma^{(L)}$ und $K_\Gamma^{(R)}$, so dass $\Gamma \subset K_\Gamma^{(L)} \cap K_\Gamma^{(R)}$. Dabei ist $K_\Gamma^{(R)}$ das Polygon, in das die Einheitsnormale n_Γ zeigt und $K_\Gamma^{(L)}$ das Gebiet, von dem n_Γ wegzeigt. Zum Verständnis siehe Abbildung 5.3. Für $v \in H^k(\Omega, \mathcal{T}_h)$ sei dann $v|_\Gamma^{(R)}$ die Spur von $v|_{K_\Gamma^{(R)}}$ entlang Γ . Analog ist $v|_\Gamma^{(L)}$ die Spur von $v|_{K_\Gamma^{(L)}}$ entlang Γ . Wir definieren den “Durchschnitt”

$$\langle v \rangle = \frac{1}{2}(v|_\Gamma^L + v|_\Gamma^R)$$

und den “Sprung”

$$[v] = v|_{\Gamma}^L - v|_{\Gamma}^R$$

entlang Γ bezüglich n_{Γ} .

5.4.4 Variationsformulierung

Wir wollen eine semidiskrete Variationsformulierung nach der Linienmethode herleiten. Ausgangspunkt ist wieder das Modellproblem (2.10)-(2.13) und es wird angenommen, dass eine Lösung u in $C^1([0, T], H^2(\Omega))$ existiert. Wir multiplizieren die Gleichung (2.10) mit einer Testfunktion v aus $H^k(\Omega, \mathcal{T}_h)$ und integrieren über $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$:

$$\int_{\Omega} \frac{\partial u}{\partial t} v dx - D \sum_{K \in \mathcal{T}_h} \int_K \Delta u v dx + \int_{\Omega} k u v dx = \int_{\Omega} f v dx.$$

Die Anwendung der Green’schen Formel führt zu:

$$\begin{aligned} & \int_{\Omega} \frac{\partial u}{\partial t} v dx + D \sum_{K \in \mathcal{T}_h} \int_K \nabla u \cdot \nabla v dx + \int_{\Omega} k u v dx \\ & - D \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \nabla u \cdot n_{\Gamma} [v] dS - D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla u \cdot n_{\Gamma} v dS = \sum_{K \in \mathcal{T}_h} \int_K f v dx + D \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} \nabla u \cdot n_{\Gamma} v dS. \end{aligned}$$

Wegen der Regularität von u gilt $\nabla u \cdot n_{\Gamma} [v] = \langle \nabla u \rangle \cdot n_{\Gamma} [v]$. Einsetzen dieser Identität ergibt:

$$\int_{\Omega} \frac{\partial u}{\partial t} v dx + D \sum_{K \in \mathcal{T}_h} \int_K \nabla u \cdot \nabla v dx - D \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla u \rangle \cdot n_{\Gamma} [v] dS \quad (5.2)$$

$$- D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla u \cdot n_{\Gamma} v dS + \int_{\Omega} k u v dx \quad (5.3)$$

$$= \sum_{K \in \mathcal{T}_h} \int_K f v dx + D \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} \nabla u \cdot n_{\Gamma} v dS. \quad (5.4)$$

Zu dieser Gleichung addieren wir auf beiden Seiten:

$$-D\epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla v \cdot n_{\Gamma} u dS = -D\epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla v \cdot n_{\Gamma} g_D dS.$$

Zusätzlich addieren wir zur linken Seite der Gleichung

$$-D\epsilon \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla v \rangle \cdot n_{\Gamma} [u] dS \quad (= 0).$$

Der Parameter ϵ bestimmt die Verfahrensart. Es gibt folgende Möglichkeiten:

- $\epsilon = -1$ Nonsymmetric Interior Penalty Galerkin (NIPG)
- $\epsilon = 0$ Incomplete Interior Penalty Galerkin (IIPG)
- $\epsilon = 1$ Symmetric Interior Penalty Galerkin (SIPG)

Wir erhalten nach Addition der oben stehenden Terme:

$$\begin{aligned}
& \int_{\Omega} \frac{\partial u}{\partial t} v dx + D \sum_{K \in \mathcal{T}_h} \int_K \nabla u \cdot \nabla v dx + k \int_{\Omega} u v dx \\
& - D \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla u \rangle \cdot n_{\Gamma} [v] dS - D\epsilon \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla v \rangle \cdot n_{\Gamma} [u] dS \\
& - D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla u \cdot n_{\Gamma} v dS - D\epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla v \cdot n_{\Gamma} u dS \\
& = \sum_{K \in \mathcal{T}_h} \int_K f v dx + D \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} \nabla u \cdot n_{\Gamma} v dS - D\epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla v \cdot n_{\Gamma} g_D dS.
\end{aligned}$$

Mit Blick auf die Neumann-Randbedingung ersetzen wir den Term auf der rechten Seite durch:

$$D \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} g_N v dS.$$

Zur Stabilisierung des Verfahrens addieren wir auf der linken Seite der Gleichung einen Strafterm für das Gebietsinnere (Interior Penalty)

$$- \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \sigma [u] \cdot [v] dS = 0$$

und ebenso für den Dirichlet-Rand (Boundary Penalty):

$$\sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma u \cdot v dS = \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma g_D \cdot v dS.$$

Diese Strafterme in der Diskretisierung sollen die Sprünge an den inneren Kanten bestrafen. Der Parameter σ ist ein Gewicht, welches von der Kante Γ und vom jeweiligen Verfahren abhängt. Wir werden später bestimmen, wie σ gewählt werden muss. Wir führen nun eine Reihe von Linear- und Bilinear-Formen ein, um die Terme in der Variationsformulierung zu systematisieren. Die Bilinearform, welche die *Diffusionsterme* zusammenfasst,

definieren wir durch

$$\begin{aligned}
a_h^\epsilon : H^k(\Omega, \mathcal{T}_h) \times H^k(\Omega, \mathcal{T}_h) &\rightarrow \mathbb{R}, \\
a_h^\epsilon(u, v) := & D \sum_{K \in \mathcal{T}_h} \int_K \nabla u \cdot \nabla v dx \\
& - D \sum_{\Gamma \in \mathcal{F}_h^I} \int_\Gamma \langle \nabla u \rangle \cdot n_\Gamma [v] dS \\
& - D \epsilon \sum_{\Gamma \in \mathcal{F}_h^I} \int_\Gamma \langle \nabla v \rangle \cdot n_\Gamma [u] dS \\
& - D \sum_{\Gamma \in \mathcal{F}_h^D} \int_\Gamma \nabla u \cdot n_\Gamma v dS \\
& - D \epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_\Gamma \nabla v \cdot n_\Gamma u dS \text{ für alle } u, v \in H^k(\Omega, \mathcal{T}_h).
\end{aligned}$$

Anmerkung:

- (i) Für das SIPG-Verfahren ($\epsilon = 1$) ist $a_h^\epsilon(u, v)$ symmetrisch.
- (ii) Es ist darauf hinzuweisen, dass die Definition von a_h^ϵ unabhängig von der Wahl der Einheitsnormalen n_Γ entlang der inneren Kanten ist.

Beweis von (ii). Wir betrachten zwei entgegengesetzte Einheitsnormalen n_Γ^1 und n_Γ^2 mit $n_\Gamma^1 = -n_\Gamma^2$. Abhängig von der Wahl der Einheitsnormalen wird die Bedeutung von "links von Γ " und "rechts von Γ " vertauscht. Das heißt also $K_\Gamma^{R^1} = K_\Gamma^{L^2}$ und $K_\Gamma^{L^1} = K_\Gamma^{R^2}$. Selbstverständlich gilt dann auch $v_{|\Gamma}^{R^1} = v_{|\Gamma}^{L^2}$ und $v_{|\Gamma}^{L^1} = v_{|\Gamma}^{R^2}$. Es folgt

$$\langle \nabla w \rangle \cdot n_\Gamma^1 (v_{|\Gamma}^{L^1} - v_{|\Gamma}^{R^1}) = \langle \nabla w \rangle \cdot n_\Gamma^2 (v_{|\Gamma}^{L^2} - v_{|\Gamma}^{R^2})$$

Die Terme der $\langle \nabla w \rangle \cdot n_\Gamma^1 (v_{|\Gamma}^{L^1} - v_{|\Gamma}^{R^1})$ sind also unabhängig von der Wahl von n_Γ . \square

Für die *Stabilisierungsterme* definieren wir die Bilinearform

$$J_h^\sigma : H^k(\Omega, \mathcal{T}_h) \times H^k(\Omega, \mathcal{T}_h) \mapsto \mathbb{R},$$

mit

$$\begin{aligned}
J_h^\sigma(u, v) := & D \sum_{\Gamma \in \mathcal{F}_h^I} \int_\Gamma \sigma[u] \cdot [v] dS \\
& + D \sum_{\Gamma \in \mathcal{F}_h^D} \int_\Gamma \sigma u \cdot v dS \text{ für alle } u, v \in H^k(\Omega, \mathcal{T}_h).
\end{aligned}$$

Für die Terme auf der rechten Seite definieren wir eine Linearform

$$L_h : H^k(\Omega, \mathcal{T}_h) \rightarrow \mathbb{R},$$

mit

$$\begin{aligned} L_h(v) := & \sum_{K \in \mathcal{T}_h} \int_K f v dx + D \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} g_N v dS \\ & - D\epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla v \cdot n_{\Gamma} g_D dS + D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma v g_D dS \text{ für alle } v \in H^k(\Omega, \mathcal{T}_h). \end{aligned}$$

Mit diesen Notationen erhalten wir die Variationsgleichung

$$\begin{aligned} \left(\frac{\partial u}{\partial t}, v \right)_{\Omega} + a_h^{\epsilon}(u, v) + J_h^{\sigma}(u, v) + (ku, v)_{\Omega} &= L_h(v), \\ \text{für alle } v \in H^k(\Omega, \mathcal{T}_h) \text{ und } t \in (0, T). \end{aligned} \quad (5.5)$$

Korollar (Konsistenz). *Aus der dargelegten Herleitung der Variationsgleichung geht hervor, dass das Verfahren konsistent ist. Das heißt eine Lösung u in $\mathcal{C}^1([0, T], H^2(\Omega))$ des RD-Modells (2.10)-(2.13) erfüllt die Variationsgleichung (5.5).*

Semidiskrete Variationsformulierung

Wir wollen die Variationsgleichung diskretisieren. Die approximierte Lösung für unser Modell suchen wir in dem Raum der stückweise stetigen Polynome. Sei also $p \geq 1$ eine ganze Zahl und $P^p(K)$ der Raum der Polynome vom Grad p mit Definitionsbereich K , dann ist

$$S_h^p := \{v; v|_K \in P^p(K) \text{ für alle } K \in \mathcal{T}_h\}$$

der Raum der *gebrochenen Polynome* vom Grad p . Wir nennen u_h eine DG-approximierte Lösung des Reaktion-Diffusion-Modells (2.10)-(2.13), falls gilt

$$\begin{aligned} \text{(a)} \quad & u_h \in \mathcal{C}^1([0, T]; S_h^p), \\ \text{(b)} \quad & \left(\frac{\partial u_h}{\partial t}, v_h \right)_{\Omega} + a_h^{\epsilon}(u_h, v_h) + J_h^{\sigma}(u_h, v_h) + (u_h, v_h)_{\Omega} = L_h(v_h), \\ & \forall v_h \in S_h^p \text{ und für fast alle } t \in (0, T), \\ \text{(c)} \quad & u_h(0) = u_h^0. \end{aligned} \quad (5.6)$$

5.5 Fehleranalyse für die Semidiskretisierung auf nicht konvexen Gebieten und nicht konformen Gittern

5.5.1 Annahmen

Definition 5.5.1 (Lokal quasiuniform). Ein Gitter heißt *lokal quasiuniform*, wenn es folgende Ungleichung erfüllt:

$$h_{K_\Gamma^{(L)}} \leq C_Q h_{K_\Gamma^{(R)}}, \text{ für alle } \Gamma \in \mathcal{F}_h^I. \quad (5.7)$$

Wir müssen zum Gitter, wie auch zu den verschiedenen Parametern, die in der Diskretisierung auftauchen, einige Annahmen machen, die uns durch die weitere Fehleranalyse begleiten werden.

Definition 5.5.2 (S_h -Interpolation). Für $v \in L^2(\Omega)$ bezeichnen wir mit $\Pi_h v \in S_h^p$ die $L^2(\Omega)$ Projektion von v auf S_h^p . Sie wird über die Beziehung

$$(\Pi_h v - v, \varphi_h) = 0 \text{ für alle } \varphi_h \in S_h.$$

definiert.

Annahmen (A1):

(a) Regularität der analytischen Lösung:

$$u \in \mathcal{C}([0, T], H^{p+1}(\Omega)).$$

(b) Regularität des Gitters

\mathcal{T}_h ist quasiuniform mit der Konstante C_Q

und es existiert eine Konstante $C_T > 0$ mit

$$\frac{h_K}{\rho_K} \leq C_T \text{ für alle } K \in \mathcal{T}_h \text{ und für alle } h \in (0, h_0).$$

(c) Es werden in der Gittererzeugung nur Dreiecke oder Tetraeder verwendet.

(d) Startfunktion wird interpoliert: $u_h^0 = \Pi_h u^0$.

(e) Der Strafparameter σ wird wie folgt berechnet [Fei09]:

$$\sigma = \frac{2C_W}{h_{K_\Gamma^{(L)}} + h_{K_\Gamma^{(R)}}}, \Gamma \in \mathcal{F}_h^I,$$

$$\sigma = \frac{C_W}{h_{K_\Gamma^{(L)}}}, \Gamma \in \mathcal{F}_h^D.$$

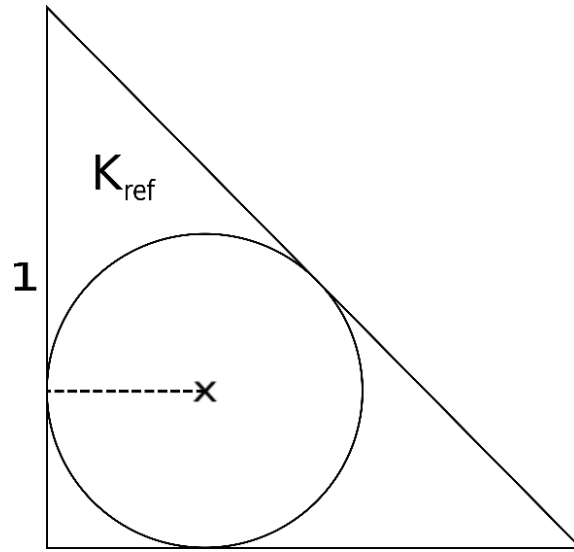


Abbildung 5.4: Das betrachtete Referenzelement ist ein rechtwinkliges Dreieck mit $h_{K_{\text{ref}}} = \sqrt{2}$ und $\rho_{K_{\text{ref}}} = (2 + \sqrt{2})^{-1}$.

Die Konstante C_W wird abhängig vom gewählten Verfahren definiert:

| | |
|-----------------------------------|-----------------------|
| $C_W > 0$ (z.B. $C_W = 1$) | für die NIPG-Version, |
| $C_W \geq 2C_M(1 + C_I)(1 + C_Q)$ | für die SIPG-Version, |
| $C_W \geq C_M(1 + C_I)(1 + C_Q)$ | für die IIPG-Version. |

5.5.2 Approximationssatz, inverse Ungleichung und Abschätzung für den Strafterm

Einige der folgenden Beweise werden auf einem sogenannten Referenzelement entwickelt. Mit Hilfe des Transformationssatzes werden die Resultate dann auf die Elemente in der Triangulierung \mathcal{T}_h angewendet. Im Fall von zwei Raumdimensionen verwenden wir das Referenzelement in Abbildung 5.4. Im Fall von drei Raumdimensionen betrachten wir einen regelmäßigen Tetraeder mit Kantenlänge $a = 1$ als Referenzelement. Es gilt dann $h_{K_{\text{ref}}} = 1$ und $\rho_{K_{\text{ref}}} = \frac{1}{12}\sqrt{6}$.

Satz 5.5.1 (Transformationsformel). *Seien K und \hat{K} zwei affin äquivalente Dreiecke oder Teraeder, d.h. es gebe eine bijektive affine Abbildung*

$$\begin{aligned} F : \hat{K} &\rightarrow K, \\ F\hat{x} &= x_0 + B\hat{x} \end{aligned}$$

mit einer nicht singulären Matrix B . Für $v \in H^k(K)$ ist dann $\hat{v} := v \circ F \in H^k(\hat{K})$, und

es gibt eine Konstante $C = C(\hat{K}, k)$, so dass für $0 \leq l \leq k$

$$|\hat{v}|_{H^l(\hat{K})} \leq C \cdot \|B\|^l \cdot |\det B|^{-\frac{1}{2}} |v|_{H^l(K)}$$

ist.

Beweis. Ein Beweis findet sich in [Bra07, S. 76]. \square

Die folgenden Lemmata und der Approximationssatz konnten weitestgehend aus [Bra07] übernommen werden. Jedoch haben wir in der genannten Literatur anstelle des hier verwendeten Projektionsoperators Π_h einen Interpolationsoperator. Die Beweise werden zudem an die hier verwendeten Referenzelemente in zwei und drei Raumdimensionen angepasst. Die Beweise mussten deshalb leicht verändert neu geführt werden.

Lemma 5.5.2. *Sei $\Omega \subset \mathbb{R}^d$ ein beschränktes Gebiet mit Lipschitz-Rand und Ω erfüllt eine Kegelbedingung. Ferner sei $k \geq 2$, und $\Pi_h : H^k(\Omega) \rightarrow \mathcal{P}^{k-1}$ die L^2 -Projektion auf Polynome vom Grad $\leq k-1$. Dann gilt mit einer Zahl $C = C(\Omega)$*

$$\|u - \Pi_h u\|_{H^k(\Omega)} \leq C |u|_{H^k(\Omega)} \text{ für } u \in H^k(\Omega). \quad (5.8)$$

Beweis. Wir führen in $H^k(\Omega)$ die Norm

$$|||v||| := |v|_{H^k(\Omega)} + \|\Pi_h v\|_{L^2(\Omega)}$$

ein und zeigen, dass die Normen $||| \cdot |||$ und $\| \cdot \|_{H^k(\Omega)}$ äquivalent sind. Dann folgt nämlich die zu beweisende Ungleichung aus

$$\begin{aligned} \|u - \Pi_h u\|_{H^k(\Omega)} &\leq C |||u - \Pi_h u||| \\ &= C \{ |u - \Pi_h u|_{H^k(\Omega)} + C \|\Pi_h(u - \Pi_h u)\|_{L^2(\Omega)} \} \\ &= C \{ |u - \Pi_h u|_{H^k(\Omega)} + C \|\Pi_h u - \Pi_h^2 u\|_{L^2(\Omega)} \} \\ &= C |u - \Pi_h u|_{H^k(\Omega)} = C |u|_{H^k(\Omega)}. \end{aligned}$$

Dabei haben wir ausgenutzt, dass $\Pi_h^2 u = \Pi_h u$ gilt und $D^k \Pi_h u = 0$ ist.

(i) Wir zeigen zunächst die Richtung $|||v||| \leq C \|v\|_{H^k(\Omega)}$.

Nach Definition des Projektionsoperators gilt

$$\begin{aligned} (\Pi_h v - v, \Pi_h v)_{L^2(\Omega)} &= 0 \\ \Rightarrow (\Pi_h v, \Pi_h v)_{L^2(\Omega)} &= (v, \Pi_h v)_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)} \|\Pi_h v\|_{L^2(\Omega)} \\ \Rightarrow \|\Pi_h v\|_{L^2(\Omega)} &\leq \|v\|_{L^2(\Omega)}. \end{aligned}$$

Die Behauptung $|||v||| \leq C \|v\|_{H^k(\Omega)}$ folgt dann direkt aus der Definition der Normen.

(ii) Die Umkehrung $\|v\|_{H^k(\Omega)} \leq C |||v|||$ beweisen wir indirekt.

Angenommen

$$\|v\|_{H^k(\Omega)} \leq C |||v||| \text{ für alle } v \in H^k(\Omega)$$

sei für jede positive Zahl C falsch. Dann gibt es eine Folge $(v_n)_{n \in \mathbb{N}}$ in $H^k(\Omega)$ mit

$$\|v_n\|_{H^k(\Omega)} = 1, \quad \|v_n\| \leq \frac{1}{n}, \quad n = 1, 2, \dots$$

Nach dem Rellickschen Auswahlssatz konvergiert eine Teilfolge von $(v_n)_{n \in \mathbb{N}}$ in $H^{k-1}(\Omega)$. Wir können o.B.d.A. annehmen, dass es sich dabei um die ganze Folge handelt. Insbesondere ist (v_n) eine Cauchy-Folge in $H^{k-1}(\Omega)$. Aus $|v_n| \rightarrow 0$ und $\|v_n - v_m\|_{H^k(\Omega)}^2 \leq \|v_n - v_m\|_{H^{k-1}(\Omega)}^2 + (\|v_n\|_{H^k(\Omega)} + \|v_m\|_{H^k(\Omega)})^2$ schließen wir, dass (v_n) sogar eine Cauchy-Folge in $H^k(\Omega)$ ist. Wegen der Vollständigkeit des Raumes liegt Konvergenz im Sinne von H^k gegen ein Element $v^* \in H^k(\Omega)$ vor. Aus Stetigkeitsgründen haben wir:

$$\|v^*\|_{H^k(\Omega)} = 1 \text{ und } \|v^*\| = 0.$$

Aus $\|v^*\| = 0$ folgt, dass $|v^*|_{H^k(\Omega)} = 0$ und $\|\Pi_h v^*\|_{L^2} = 0$ gilt. Aus $|v^*|_{H^k(\Omega)} = 0$ können wir schlussfolgern, dass v^* ein Polynom vom Grad $k-1$ ist. Damit gilt aber $\|v^*\|_{L^2(\Omega)} = \|\Pi_h v^*\|_{L^2(\Omega)} = 0$. Also ist v^* das Nullpolynom. Dies ist jedoch ein Widerspruch zu $\|v^*\|_{H^k(\Omega)} = 1$. \square

Lemma 5.5.3. *Seien K_1 und K_2 zwei affin äquivalente Dreiecke bzw. Tetraeder. Sei $F : K_1 \rightarrow K_2$ eine bijektive affin lineare Abbildung wie in Satz 5.5.1. Dann gilt:*

$$\|B\| \leq \frac{h_{K_2}}{\rho_{K_1}}.$$

Falls zudem $K_1, K_2 \in \mathcal{T}_h \cup \{K_{ref}\}$ gilt folgt

$$\|B\| \|B\|^{-1} \leq C_T^2.$$

wobei C_T die in den Annahmen (A1) erwähnte Konstante ist, welche $\frac{h_K}{\rho_K} \leq C_T$ für alle $K \in \mathcal{T}_h$ (und auch $\frac{h_{K_{ref}}}{\rho_{K_{ref}}} \leq C_T$) erfüllt.

Beweis. Sei $x \in \mathbb{R}^d$ mit $\|x\| \leq \rho_{K_1}$. Dann finden wir zwei Punkte $y_1, z_1 \in K_1$ mit $x = y_1 - z_1$. Wegen $F(y_1), F(z_1) \in K_2$ gilt $\|Bx\| \leq h_{K_2}$. Also folgt:

$$\|B\| = \sup_{\|x\|=\rho_{K_1}} \frac{\|Bx\|}{\|x\|} \leq \frac{h_{K_2}}{\rho_{K_1}}.$$

Vertauschen wir K_1 und K_2 , so erhalten wir für die inverse Matrix B^{-1} analog: $\|B^{-1}\| \leq \frac{h_{K_1}}{\rho_{K_2}}$ und insgesamt:

$$\|B\| \cdot \|B^{-1}\| \leq \frac{h_{K_1} h_{K_2}}{\rho_{K_1} \rho_{K_2}} \leq C_T^2.$$

\square

Satz 5.5.4 (Approximationssatz). *Sei \mathcal{T}_h eine Triangulierung von Ω , welche die Annahmen (A1) erfüllt. Sei weiterhin $k \geq 2$ und $0 \leq l \leq k$. Für $K \in \mathcal{T}_h$. Mit $\Pi_h : H^k(K) \rightarrow \mathcal{P}^{k-1}$ bezeichnen wir die L^2 -Projektion. Dann gibt es eine Konstante C_A , welche unabhängig von h, K und u folgende Ungleichung erfüllt:*

$$\|u - \Pi_h u\|_{H^l(K)} \leq C_A h^{k-l} |u|_{H^k(K)} \text{ für alle } K \in \mathcal{T}_h, u \in H^k(K), h \in (0, h_0).$$

Beweis. Mit $K_{\text{ref}} \subset \mathbb{R}^d$ bezeichnen wir das in (5.4) eingeführte Referenzelement, mit $h_{K_{\text{ref}}}$ und $\rho_{K_{\text{ref}}}$ seinen Diameter bzw. Innkreisradius. Sei $F : K_{\text{ref}} \rightarrow K$ eine bijektive affin lineare Abbildung wie in Satz 5.5.1 und $K \in \mathcal{T}_h$. Einer gegebenen Funktion $u \in H^k(K)$ ordnen wir eine Funktion $v \in H^k(K_{\text{ref}})$ gemäß

$$v(x) := u(Bx) \text{ für alle } x \in K_{\text{ref}}$$

zu. Indem wir Lemma (5.5.2) auf das Referenzdreieck anwenden und die Transformationsformel in beiden Richtungen verwenden, erhalten wir für $0 \leq m \leq l$:

$$\begin{aligned} |u - \Pi_h u|_{m,K} &\leq C \|B^{-1}\|^{-m} \cdot |\det B|^{\frac{1}{2}} |v - \Pi_h v|_{m,K_{\text{ref}}} \\ &\leq \tilde{C} \|B^{-1}\|^{-m} \cdot |\det B|^{\frac{1}{2}} \cdot |v|_{k,K_{\text{ref}}} \\ &\leq \hat{C} \|B^{-1}\|^{-m} \cdot |\det B|^{\frac{1}{2}} \cdot \|B\|^k \cdot |\det B|^{-\frac{1}{2}} |u|_{k,K} \\ &\leq \hat{C} (\|B\| \cdot \|B^{-1}\|)^m \|B\|^{k-m} |u|_{k,K}. \end{aligned}$$

Wegen Lemma (5.5.3) gilt $\|B\| \leq \frac{h_K}{\rho_{K_{\text{ref}}}}$ und $\|B\| \|B^{-1}\| \leq C_T^2$. Es ergibt sich also:

$$|u - \Pi_h u|_{m,K} \leq \hat{C} C_T^{2m} \left(\frac{h}{\rho_{K_{\text{ref}}}} \right)^{k-m} |u|_{k,K}.$$

Nach dem Quadrieren liefert die Summation über m von 0 bis l die Behauptung. □

Der nachfolgende Beweis wurde aus [Bra07][S.79] übertragen.

Satz 5.5.5 (Inverse Ungleichung). *Wir fixieren ein $p \leq 1$. Es gibt eine (von p abhängige Konstante) $C_I > 0$, welche unabhängig von v, h und K ist und folgende Ungleichung erfüllt:*

$$\|v\|_{H^1(K)} \leq C_I h_K^{-1} \|v\|_{L^2(K)} \text{ für alle } K \in \mathcal{T}_h, v \in P^p(K), h \in (0, h_0).$$

Beweis. Wegen der endlichen Dimension von $P^p(K_{\text{ref}})$ sind die Normen $\|\cdot\|_{H^1(K_{\text{ref}})}$ und $\|\cdot\|_{L^2(K_{\text{ref}})}$ äquivalent. Es folgt:

$$|v|_{H^1(K_{\text{ref}})} \leq C \|v\|_{L^2(K_{\text{ref}})}$$

Mit der Transformationsformel und der Äquivalenz der Normen erhalten wir:

$$\begin{aligned}
|u|_{H^1(K)} &\leq \tilde{C} \|B^{-1}\| \cdot |\det B|^{\frac{1}{2}} |v|_{H^1(K_{\text{ref}})} \\
&\leq \hat{C} \|B^{-1}\| \cdot |\det B|^{\frac{1}{2}} \|v\|_{L^2(K_{\text{ref}})} \\
&\leq \bar{C} \|B^{-1}\| \cdot |\det B|^{\frac{1}{2}} |\det B|^{-\frac{1}{2}} \|u\|_{L^2(K)} \\
&= \bar{C} \|B^{-1}\| \|u\|_{L^2(K)}.
\end{aligned}$$

Wegen $\|B^{-1}\| \leq \frac{h_{K_{\text{ref}}}}{\rho_K} \leq \frac{C_T h_{K_{\text{ref}}}}{h_K}$ folgt:

$$|u|_{H^1(K)} \leq C_I h_K^{-1} \|u\|_{L^2(K)}.$$

□

Die Beweisidee für die folgende multiplikative Spurgleichung stammt aus einer Vorlesung von Feistauer in Berlin 2009.

Satz 5.5.6 (Multiplikative Spurgleichung). *Es gibt eine Konstante $C_M > 0$, welche unabhängig von v, h und K ist und folgende Ungleichung erfüllt:*

$$\begin{aligned}
\|v\|_{L^2(\partial K)}^2 &\leq C_M (\|v\|_{L^2(K)} |v|_{H^1(K)} + h_K^{-1} \|v\|_{L^2(K)}), \\
K &\in \mathcal{T}_h, v \in H^1(K), h \in (0, h_0).
\end{aligned}$$

Beweis. Sei $K \in \mathcal{T}_h$ beliebig gewählt. Wir legen o.B.d.A. den Koordinatenursprung in den Mittelpunkt der größten Kugel (Kreis in 2D), welche in K eingeschrieben ist. \vec{n} ist die nach außen zeigende Einheitsnormale auf ∂K . Sei $x \in \partial K$ beliebig gewählt. Dann gilt $x \cdot \vec{n} = \rho_K$. Sei $v \in H^1(K)$, dann folgt:

$$\begin{aligned}
\rho_K \|v\|_{L^2(\partial K)}^2 &= \int_{\partial K} v^2 \cdot x \cdot \vec{n} \stackrel{\text{Green}}{=} \int_K \nabla \cdot (v^2 x) dx \\
&= \int_K v^2 \underbrace{\nabla \cdot x}_{=d} dx + \int_K x \cdot \underbrace{\nabla(v^2)}_{=2v \nabla v} dx \\
&\leq d \int_K v^2 dx + 2 \sup_{x \in K} |x| \int_K |v| |\nabla v| dx \\
&\leq d \|v\|_{L^2(K)}^2 + 2h_K \|v\|_{L^2(K)} |v|_{H^1(K)}.
\end{aligned}$$

Wir erhalten:

$$\|v\|_{L^2(\partial K)}^2 \leq \frac{d}{\rho_K} \|v\|_{L^2(K)}^2 + \frac{2h_K}{\rho_K} \|v\|_{L^2(K)} |v|_{H^1(K)}$$

Mit der Beziehung $\frac{h_K}{\rho_K} \leq C_T$ aus den Annahmen(A1) folgt also:

$$\|v\|_{L^2(\partial K)}^2 \leq dC_T \left(h_K^{-1} \|v\|_{L^2(K)}^2 + \|v\|_{L^2(K)} |v|_{H^1(K)} \right)$$

Mit $C_M = dC_T$ folgt die Behauptung. \square

Das folgende Lemma stammt aus [Fei09]. Die Beweisidee stammt aus einer Vorlesung von Feistauer in Berlin 2009.

Lemma 5.5.7.

$$(1) |J_h^\sigma(\eta, \xi)| \leq |J_h^\sigma(\eta, \eta)|^{\frac{1}{2}} |J_h^\sigma(\xi, \xi)|^{\frac{1}{2}} \quad (5.9)$$

$$(2) |J_h^\sigma(\eta, \eta)| \leq Ch^p |u|_{H^{p+1}(\Omega)} \quad (5.10)$$

$$(3) |a_h^\xi(\eta, \xi)| \leq Ch^p \|\xi\|_\varepsilon |u|_{H^{p+1}(\Omega, \mathcal{T}_h)}^2. \quad (5.11)$$

Beweis. (1) Durch Anwendung der Cauchy-Schwarz'schen Ungleichung erhalten wir folgende Abschätzung:

$$\begin{aligned} |J_h^\sigma(\eta, \xi)| &= D \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \sigma[\eta] \cdot [\xi] dS + D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma \eta \cdot \xi dS \\ &\leq D \sum_{\Gamma \in \mathcal{F}_h^I} \left(\int_{\Gamma} \sigma[\eta]^2 dS \right)^{\frac{1}{2}} \left(\int_{\Gamma} \sigma[\xi]^2 dS \right)^{\frac{1}{2}} \\ &\quad + D \sum_{\Gamma \in \mathcal{F}_h^D} \left(\int_{\Gamma} \sigma \eta^2 dS \right)^{\frac{1}{2}} \left(\int_{\Gamma} \sigma \xi^2 dS \right)^{\frac{1}{2}} \end{aligned}$$

Wiederholte Anwendung der Cauchy-Schwarz'schen Ungleichung ergibt dann das gewünschte Resultat:

$$\begin{aligned} |J_h^\sigma(\eta, \xi)| &\leq D \left(\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \sigma[\eta]^2 dS + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma \eta^2 dS \right)^{\frac{1}{2}} \\ &\quad \cdot \left(\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \sigma[\xi]^2 dS + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma \xi^2 dS \right)^{\frac{1}{2}} \\ &= |J_h^\sigma(\eta, \eta)|^{\frac{1}{2}} |J_h^\sigma(\xi, \xi)|^{\frac{1}{2}}. \end{aligned}$$

(2) Es gilt zunächst

$$\begin{aligned} J_h^\sigma(\eta, \eta) &= D \left(\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \sigma[\eta]^2 dS + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma \eta^2 dS \right) \\ &= D \left(\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \frac{2C_W}{h_{K_\Gamma^{(L)}} + h_{K_\Gamma^{(R)}}} [\eta]^2 dS + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \frac{C_W}{h_{K_\Gamma^{(L)}}} \eta^2 dS \right). \end{aligned}$$

Nun haben wir für die inneren Kanten $\Gamma \in \mathcal{F}_h^I$ entweder $h_K = h_{K_\Gamma^{(L)}}$ oder $h_K = h_{K_\Gamma^{(R)}}$. Für die Dirichlet-Kanten $\Gamma \in \mathcal{F}_h^D$ gilt immer $h_K = h_{K_\Gamma^{(L)}}$. Es folgt also weiter:

$$\begin{aligned} J_h^\sigma(\eta, \eta) &\leq C \sum_{K \in \mathcal{T}_h} h_K^{-1} \int_{\partial K} |\eta|^2 dS \\ &= C \sum_{K \in \mathcal{T}_h} h_K^{-1} \|\eta\|_{L^2(\partial K)}^2. \end{aligned}$$

Anwendung der multiplikativen Spurungleichung (siehe Theorem 5.5.6) führt zu:

$$J_h^\sigma(\eta, \eta) \leq C \sum_{K \in \mathcal{T}_h} h_K^{-1} C_M \left(\|\eta\|_{L^2(K)} \|\eta\|_{H^1(K)} + h_K^{-1} \|\eta\|_{L^2(K)}^2 \right).$$

Anwendung des Approximationssatzes (siehe Theorem 5.5.4) führt weiter zu:

$$\begin{aligned} J_h^\sigma(\eta, \eta) &\leq C \sum_{K \in \mathcal{T}_h} h^{2p} |u|_{H^{p+1}(K)} \\ &= C h^{2p} |u|_{H^{p+1}(\Omega)}. \end{aligned}$$

(3) Anwendung der Dreiecksungleichung führt zu:

$$\begin{aligned}
a_h^\epsilon(\eta, \xi) \leq & D \underbrace{\left| \sum_{K \in \mathcal{T}_h} \int_K \nabla \eta \cdot \nabla \xi dx \right|}_{T_1} \\
& + D \underbrace{\left| \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla \eta \rangle \cdot n_{\Gamma} [\xi] dS + \epsilon \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla \xi \rangle \cdot n_{\Gamma} [\eta] dS \right|}_{T_2} \\
& + D \underbrace{\left| \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla \eta \cdot n_{\Gamma} \xi dS + \epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla \xi \cdot n_{\Gamma} \eta dS \right|}_{T_3}.
\end{aligned}$$

Wir schätzen die Terme einzeln ab.

$$\begin{aligned}
|T_1| &= \left| \sum_{K \in \mathcal{T}_h} \int_K \nabla u \cdot \nabla v dx \right| \\
&\leq \left(\sum_{K \in \mathcal{T}_h} \int_K |\nabla \eta|^2 dx \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}_h} \int_K |\nabla \xi|^2 dx \right)^{\frac{1}{2}} \\
&= |\eta|_{H^1(\Omega, \mathcal{T}_h)} |\xi|_{H^1(\Omega, \mathcal{T}_h)}.
\end{aligned}$$

Der Approximationssatz und die Definition für die Energienorm liefern dann:

$$|T_1| \leq C \|\xi\|_{\mathcal{E}} h_K^k |u|_{H^{k+1}(K)}.$$

Schauen wir uns den zweiten Term an:

$$\begin{aligned}
|T_2| &= \left| \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla \eta \rangle \cdot n_{\Gamma} [\xi] dS + \epsilon \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \langle \nabla \xi \rangle \cdot n_{\Gamma} [\eta] dS \right| \\
&\leq \left(\sum_{\Gamma \in \mathcal{F}_h^I} d(\Gamma) \int_{\Gamma} |\langle \nabla \eta \rangle|^2 dS \right)^{\frac{1}{2}} \left(\sum_{\Gamma \in \mathcal{F}_h^I} d(\Gamma)^{-1} \int_{\Gamma} [\xi]^2 dS \right)^{\frac{1}{2}} \\
&\quad + \left(\sum_{\Gamma \in \mathcal{F}_h^I} d(\Gamma) \int_{\Gamma} |\langle \nabla \xi \rangle|^2 dS \right)^{\frac{1}{2}} \left(\sum_{\Gamma \in \mathcal{F}_h^I} d(\Gamma)^{-1} \int_{\Gamma} [\eta]^2 dS \right)^{\frac{1}{2}}.
\end{aligned}$$

Mit der Definition für den Fehlerterm $J_h^\sigma(.,.)$ erhalten wir die Abschätzung:

$$\begin{aligned}
|T_2| &\leq C \left(\sum_{\Gamma \in \mathcal{F}_h^I} d(\Gamma) \int_{\Gamma} |\langle \nabla \eta \rangle|^2 dS \right)^{\frac{1}{2}} J_h^\sigma(\xi, \xi)^{\frac{1}{2}} \\
&\quad + C \left(\sum_{\Gamma \in \mathcal{F}_h^I} d(\Gamma) \int_{\Gamma} |\langle \nabla \xi \rangle|^2 dS \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}} \\
&\leq C \left(\sum_{K \in \mathcal{T}_h} h_K \int_{\partial K} |\nabla \eta|^2 \right)^{\frac{1}{2}} J_h^\sigma(\xi, \xi)^{\frac{1}{2}} \\
&\quad + C \left(\sum_{K \in \mathcal{T}_h} h_K \int_{\partial K} |\nabla \xi|^2 \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}}. \tag{5.12}
\end{aligned}$$

Mit Anwendung der multiplikativen Spurungleichung ergibt sich dann:

$$\begin{aligned}
|T_2| &\leq C \left(\sum_{K \in \mathcal{T}_h} h_K \left(\|\nabla \eta\|_{L^2(K)} |\nabla \eta|_{H^1(K)} + h_K^{-1} \|\nabla \eta\|_{L^2(K)}^2 \right) \right)^{\frac{1}{2}} J_h^\sigma(\xi, \xi)^{\frac{1}{2}} \\
&\quad + C \left(\sum_{K \in \mathcal{T}_h} h_K \left(\|\nabla \xi\|_{L^2(K)} |\nabla \xi|_{H^1(K)} + h_K^{-1} \|\nabla \xi\|_{L^2(K)}^2 \right) \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}} \\
&= C \left(\sum_{K \in \mathcal{T}_h} \left(h_K |\eta|_{H^1(K)} |\eta|_{H^2(K)} + |\eta|_{H^1(K)}^2 \right) \right)^{\frac{1}{2}} J_h^\sigma(\xi, \xi)^{\frac{1}{2}} \\
&\quad + C \left(\sum_{K \in \mathcal{T}_h} \left(h_K |\xi|_{H^1(K)} |\xi|_{H^2(K)} + |\xi|_{H^1(K)}^2 \right) \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}}.
\end{aligned}$$

Anwendung des Approximationssatzes auf den ersten Summanden und Anwendung

der inversen Ungleichung auf den zweiten Summanden ergeben:

$$\begin{aligned}
|T_2| &\leq C \left(\sum_{K \in \mathcal{T}_h} \left(C_A h_K^{2k} |u|_{H^{k+1}(K)}^2 + h_K^{2k} |u|_{H^{k+1}(K)}^2 \right) \right)^{\frac{1}{2}} \|\xi\|_{\mathcal{E}} \\
&\quad + C \left(\sum_{K \in \mathcal{T}_h} \left(|\xi|_{H^1(K)} \frac{1}{C_I} |\xi|_{H^1(K)} + |\xi|_{H^1(K)}^2 \right) \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}} \\
&\leq \bar{C} h^k \left(\sum_{K \in \mathcal{T}_h} |u|_{H^{k+1}(K)}^2 \right)^{\frac{1}{2}} \|\xi\|_{\mathcal{E}} + \tilde{C} \left(\sum_{K \in \mathcal{T}_h} |\xi|_{H^1(K)}^2 \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}} \\
&= \bar{C} h^k |u|_{H^{k+1}(\Omega, \mathcal{T}_h)}^2 \|\xi\|_{\mathcal{E}} + \tilde{C} |\xi|_{H^1(\Omega, \mathcal{T}_h)} J_h^\sigma(\eta, \eta)^{\frac{1}{2}}.
\end{aligned}$$

Anwendung von (2) aus Lemma 5.5.7 führt zu:

$$\begin{aligned}
|T_2| &\leq \bar{C} h^k |u|_{H^{k+1}(\Omega, \mathcal{T}_h)}^2 \|\xi\|_{\mathcal{E}} + \tilde{C} \|\xi\|_{\mathcal{E}} h^k |u|_{H^{k+1}(\Omega, \mathcal{T}_h)}^2 \\
&= \hat{C} h^k \|\xi\|_{\mathcal{E}} |u|_{H^{k+1}(\Omega, \mathcal{T}_h)}^2.
\end{aligned}$$

Zum Ende der dritte Term:

$$\begin{aligned}
|T_3| &= \left| \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla \eta \cdot n_{\Gamma} \xi dS + \epsilon \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla \xi \cdot n_{\Gamma} \eta dS \right| \\
&\leq \left(\sum_{\Gamma \in \mathcal{F}_h^D} d(\Gamma) \int_{\Gamma} |\nabla \eta|^2 dS \right)^{\frac{1}{2}} \left(\sum_{\Gamma \in \mathcal{F}_h^D} d(\Gamma)^{-1} \int_{\Gamma} \xi^2 dS \right)^{\frac{1}{2}} \\
&\quad + \left(\sum_{\Gamma \in \mathcal{F}_h^D} d(\Gamma) \int_{\Gamma} |\nabla \xi|^2 dS \right)^{\frac{1}{2}} \left(\sum_{\Gamma \in \mathcal{F}_h^D} d(\Gamma)^{-1} \int_{\Gamma} \eta^2 dS \right)^{\frac{1}{2}}
\end{aligned}$$

Mit der Definition für den Fehlerterm $J_h^\sigma(\cdot, \cdot)$ erhalten wir die Abschätzung:

$$\begin{aligned}
|T_3| &\leq C \left(\sum_{\Gamma \in \mathcal{F}_h^D} d(\Gamma) \int_{\Gamma} |\nabla \eta|^2 dS \right)^{\frac{1}{2}} J_h^\sigma(\xi, \xi)^{\frac{1}{2}} \\
&\quad + C \left(\sum_{\Gamma \in \mathcal{F}_h^D} d(\Gamma) \int_{\Gamma} |\nabla \xi|^2 dS \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}} \\
&\leq C \left(\sum_{K \in \mathcal{T}_h} h_K \int_{\partial K} |\nabla \eta|^2 \right)^{\frac{1}{2}} J_h^\sigma(\xi, \xi)^{\frac{1}{2}} \\
&\quad + C \left(\sum_{K \in \mathcal{T}_h} h_K \int_{\partial K} |\nabla \xi|^2 \right)^{\frac{1}{2}} J_h^\sigma(\eta, \eta)^{\frac{1}{2}}.
\end{aligned}$$

An dieser Stelle können wir wie in (5.12) fortfahren und erhalten:

$$|T_3| \leq \check{C} h^k \|\xi\|_{\mathcal{E}} |u|_{H^{k+1}(\Omega, \mathcal{T}_h)}^2.$$

Insgesamt folgt also:

$$\begin{aligned}
|a_h^\epsilon(\eta, \xi)| &\leq |T_1| + |T_2| + |T_3| \\
&\leq \check{C} h^k \|\xi\|_{\mathcal{E}} |u|_{H^{k+1}(\Omega, \mathcal{T}_h)}^2.
\end{aligned}$$

□

5.5.3 Koerzivität

Wir wollen die Koerzivität der bilinearen Terme in unserer Variationsformulierung beweisen. Wir definieren eine weitere Bilinearform und eine Energienorm.

Definition 5.5.3 (Bilinearform, Energienorm). Die folgende Bilinearform setzt sich aus dem Diffusionsterm, dem Reaktionsterm und dem Strafterm zusammen:

$$A_h^{\epsilon, \sigma}(v, w) := a_h^\epsilon(v, w) + (kv, w)_\Omega + J_h^\sigma(v, w)$$

Wir definieren eine *Energienorm* auf $H^k(\Omega, \mathcal{T}_h)$ wie folgt:

$$\|v\|_{\mathcal{E}} = \left(\sum_{K \in \mathcal{T}_h} \int_K D(\nabla v)^2 dx + (kv, v)_\Omega + J_h^\sigma(v, v) \right)^{\frac{1}{2}}.$$

Satz 5.5.8. Für die Menge der Dirichlet-Kanten nehmen wir an: $\mathcal{F}_h^D \neq \emptyset$. Dann ist die

Energienorm tatsächlich eine Norm auf $v \in H^k(\Omega, \mathcal{T}_h)$.

Beweis. Wir beginnen mit der positiven Definitheit, das heißt $\|v\|_{\mathcal{E}} = 0 \Rightarrow v = 0$ für alle $v \in H^k(\Omega, \mathcal{T}_h)$. Aus

$$\|v\|_{\mathcal{E}} = \left(\sum_{K \in \mathcal{T}_h} \int_K D(\nabla v)^2 dx + (kv, v)_{\Omega} + J_h^{\sigma}(v, v) \right)^{\frac{1}{2}} = 0 \quad (5.13)$$

folgt zunächst:

$$\begin{aligned} \int_K (\nabla v)^2 &= 0 \text{ für alle } K \in \mathcal{T}_h, \\ \Rightarrow v &\equiv C \text{ auf } K, \text{ für alle } K \in \mathcal{T}_h. \end{aligned}$$

Wegen (5.13) gilt außerdem $0 = J_h^{\sigma}(v, v)$. Es folgt:

$$\begin{aligned} 0 &= \int_{\Gamma} \sigma[v] \cdot [v] dS \text{ für alle } \Gamma \in \mathcal{F}_h^I, \\ \text{und } 0 &= \int_{\Gamma} \sigma u \cdot v dS \text{ für alle } \Gamma \in \mathcal{F}_h^D. \end{aligned}$$

Das heißt $\mathcal{F}_h^D \neq \emptyset$ gibt es ein $K \in \mathcal{T}_h$ mit $v \equiv 0$ auf K . Weiterhin sind alle Sprungterme Null. Es folgt also $v \equiv 0$ auf Ω .

Die absolute Homogenität, also $\|\lambda v\|_{\mathcal{E}} = |\lambda| \|v\|_{\mathcal{E}}$ für alle $\lambda \in \mathbb{R}, v \in H^k(\Omega, \mathcal{T}_h)$ folgt direkt aus der Definition der Norm. Das gleiche gilt für die Dreiecksungleichung. \square

Wir werden nun folgenden Satz beweisen:

Satz 5.5.9 (Koerzivität). *Unter den Voraussetzungen (A1) ist die Bilinearform $A_h^{\epsilon, \sigma}$ koerziv, d.h. es gibt eine Konstante $\kappa > 0$ mit:*

$$A_h^{\epsilon, \sigma}(v, v) \geq \kappa \|v\|_{\mathcal{E}}^2 \text{ für alle } v \in H^k(\Omega, \mathcal{T}_h), \epsilon \in \{-1, 0, 1\}.$$

Insbesondere können wir für die NIPG-Methode $\kappa = 1$ wählen.

Beweis. Für die Bilinearform gilt zunächst:

$$\begin{aligned} A_h^{\epsilon, \sigma}(v, v) &= \sum_{K \in \mathcal{T}_h} \int_K D(\nabla v)^2 dx + (kv, v)_{\Omega} + J_h^{\sigma}(v, v) \\ &\quad - (\epsilon + 1) \left(\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} D \langle \nabla v \rangle \cdot n_{\Gamma} [v] dS + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} D \nabla v \cdot n_{\Gamma} v dS \right). \end{aligned}$$

Im Falle des NIPG-Verfahrens mit $\epsilon = -1$ verschwindet der letzte Term und es folgt direkt:

$$A_h^{\epsilon, \sigma}(v, v) = \|v\|_{\mathcal{E}}^2.$$

Die Koerzivittsbedingung wird also mit $\kappa = 1$ erfllt.

Im Folgenden werden wir den Fall $\epsilon = 1$ (SIPG) oder $\epsilon = 0$ (NIPG) betrachten.

Dazu schauen wir uns zunchst die Bilinearform a_h^ϵ an. Es gilt:

$$a_h^\epsilon(v, v) \geq D|v|_{H^1(K)} - (1 + \epsilon) \left(\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} D \langle \nabla v \rangle \cdot n_{\Gamma} [v] dS + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} D \nabla v \cdot n_{\Gamma} v dS \right).$$

Anwendung der Young'schen Ungleichung fr ein beliebiges $\delta > 0$:

$$\begin{aligned} a_h^\epsilon(v, v) &\geq D|v|_{H^1(K)} \\ &\quad - (1 + \epsilon) D \frac{1}{2\delta} \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \frac{h_{K_{\Gamma}^{(L)}} + h_{K_{\Gamma}^{(R)}}}{2} (\langle \nabla v \rangle \cdot n_{\Gamma})^2 dS - (1 + \epsilon) D \frac{\delta}{2C_W} \sigma \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} [v]^2 dS \\ &\quad - (1 + \epsilon) D \frac{1}{2\delta} \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} h_{K_{\Gamma}^{(L)}} (\nabla v \cdot n_{\Gamma})^2 dS - (1 + \epsilon) D \frac{\delta}{2C_W} \sigma \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} v^2 dS \\ &= D|v|_{H^1(K)} - D \frac{1}{2\delta} \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \frac{h_{K_{\Gamma}^{(L)}} + h_{K_{\Gamma}^{(R)}}}{2} (\langle \nabla v \rangle \cdot n_{\Gamma})^2 dS \\ &\quad - (1 + \epsilon) D \frac{1}{2\delta} \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} h_{K_{\Gamma}^{(L)}} (\nabla v \cdot n_{\Gamma})^2 dS - D(1 + \epsilon) \frac{\delta}{2C_W} J_h^\sigma(v, v). \end{aligned}$$

Anwendung von Cauchy-Schwarz:

$$\begin{aligned} a_h^\epsilon(v, v) &\geq D|v|_{H^1(K)} - (1 + \epsilon) D \frac{1}{2\delta} \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \frac{h_{K_{\Gamma}^{(L)}} + h_{K_{\Gamma}^{(R)}}}{2} |\langle \nabla v \rangle|^2 dS \\ &\quad - D(1 + \epsilon) \frac{1}{2\delta} \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} h_{K_{\Gamma}^{(L)}} |\nabla v|^2 dS - (1 + \epsilon) D \frac{\delta}{2C_W} J_h^\sigma(v, v). \end{aligned}$$

Wir nutzen die Quasiuniformitt des Gitters und erhalten weiter:

$$\begin{aligned} a_h^\epsilon(v, v) &\geq D|v|_{H^1(K)} - (1 + \epsilon) D \frac{1}{2\delta} \frac{1 + C_Q}{2} \left(\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} |\langle \nabla v \rangle|^2 dS + \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} |\nabla v|^2 dS \right) \\ &\quad - (1 + \epsilon) D \frac{\delta}{2C_W} J_h^\sigma(v, v). \end{aligned}$$

Mit der quadratischen Ungleichung $|\langle \nabla v \rangle|^2 \leq 2(|\nabla v_{\Gamma}^L|^2 + |\nabla v_{\Gamma}^R|^2)$ knnen wir weiter ver-

einfachen zu:

$$\begin{aligned} a_h^\epsilon(v, v) &\geq D|v|_{H^1(K)} - (1 + \epsilon)D \frac{1}{2\delta} \frac{1 + C_Q}{2} \sum_{K \in \mathcal{T}_h} h_K \int_{\partial K} |\nabla v|^2 dS \\ &\quad - (1 + \epsilon)D \frac{\delta}{2C_W} J_h^\sigma(v, v). \end{aligned}$$

Die Anwendung der multiplikativen Spurungleichung (siehe Theorem 5.5.6) führt zu:

$$\begin{aligned} a_h^\epsilon(v, v) &\geq D|v|_{H^1(K)} - (1 + \epsilon)D \frac{1}{2\delta} \frac{1 + C_Q}{2} C_M \sum_{K \in \mathcal{T}_h} \left(h_K \|v\|_{L_K^2} |v|_{H^1(K)} + \|v\|_{L^2(K)}^2 \right) \\ &\quad - (1 + \epsilon)D \frac{\delta}{2C_W} J_h^\sigma(v, v). \end{aligned}$$

Durch weitere Anwendung der inversen Ungleichung (siehe Theorem 5.5.5) erhalten wir dann:

$$\begin{aligned} a_h^\epsilon(v, v) &\geq D|v|_{H^1(K)} - (1 + \epsilon)D \frac{1}{2\delta} \frac{1 + C_Q}{2} C_M (1 + C_I) |v|_{H^1(\Omega, \mathcal{T}_h)}^2 \\ &\quad - (1 + \epsilon)D \frac{\delta}{2C_W} J_h^\sigma(v, v). \end{aligned}$$

SIPG($\epsilon = 1$): Wir setzen $\delta := C_M(1 + C_I)(1 + C_Q)$ und $C_W \geq 2C_M(1 + C_I)(1 + C_Q)$. Es folgt:

$$a_h^\epsilon(v, v) \geq D|v|_{H^1(K)} - D \frac{1}{2} |v|_{H^1(\Omega, \mathcal{T}_h)}^2 - D \frac{1}{2} J_h^\sigma(v, v).$$

und damit:

$$A_h^{\epsilon, \sigma}(v, v) \geq \min\left\{\frac{D}{2}, 1\right\} \left(|v|_{H^1(K)} + k(v, v)_\Omega + J_h^\sigma(v, v) \right).$$

IIPG($\epsilon = 0$): Wir setzen $\delta := \frac{1}{2}C_M(1 + C_I)(1 + C_Q)$ und $C_W \geq C_M(1 + C_I)(1 + C_Q)$. Damit folgt dann:

$$a_h^\epsilon(v, v) \geq D|v|_{H^1(K)} - D \frac{1}{2} |v|_{H^1(\Omega, \mathcal{T}_h)}^2 - D \frac{1}{2} J_h^\sigma(v, v)$$

und letztendlich:

$$A_h^{\epsilon, \sigma}(v, v) \geq \min\left\{\frac{D}{2}, 1\right\} \left(|v|_{H^1(K)} + k(v, v)_\Omega + J_h^\sigma(v, v) \right).$$

□

5.5.4 ODE-Formulierung

Das semidiskrete Problem (5.6) ist äquivalent zu einem nichtlinearen System gewöhnlicher Differentialgleichungen. Sei $\{v_1, \dots, v_N\}$ eine Basis für S_h^p . Wir können dann $u_h \in \mathcal{C}^1([0, T]; S_h^p)$ in der Form $u_h(t) = b_1(t) \cdot v_1 + \dots + b_N(t) \cdot v_N$ darstellen mit zeitabhängigen Koeffizienten $b_i \in \mathcal{C}^1([0, T])$. Wir erhalten dann das Gleichungssystem:

$$\sum_{i=1}^N b'_i(t) (v_i, v_j)_\Omega + \sum_{i=1}^N b_i(t) (a_h^\epsilon(v_i, v_j) + J_h^\sigma(v_i, v_j) + (kv_i, v_j)_\Omega) = L_h(v_j),$$

$$\vec{b}(0) = \text{Koeffizientenvektor bezüglich } v_i \text{ von } u_h^0. \text{ für alle } j = 1, \dots, N.$$

Das gleiche kann in Matrixschreibweise ausgedrückt werden:

$$\mathcal{M} \cdot \vec{b}'(t) + \mathcal{S} \cdot \vec{b}(t) = \vec{L} \text{ für alle } j = 1, \dots, N, \quad (5.14)$$

$$\vec{b}(0) = \text{Koeffizientenvektor bezüglich der Basis } \{v_1, \dots, v_N\} \text{ von } u_h^0. \quad (5.15)$$

Dabei ist $\mathcal{M} = ((v_i, v_j)_\Omega)_{i,j}$ die *Massematrix*, $\mathcal{S} = ((a_h^\epsilon(v_i, v_j) + J_h^\sigma(v_i, v_j) + (kv_i, v_j)_\Omega))_{i,j}$ die *Steifigkeitsmatrix* und $\vec{L} = (L_h(v_i))_i$ ein Vektor. Dies ist eine gewöhnliche Differentialgleichung erster Ordnung, die mit den üblichen Methoden gelöst werden kann. Die Diskretisierung in der Zeit folgt in Abschnitt 5.7.

5.5.5 Existenz und Eindeutigkeit

Die Steifigkeitsmatrix \mathcal{S} in (5.14) ist wegen der Koerzivitätsbeziehung (s. Theorem(5.5.9)) positiv definit. Die Massematrix ist symmetrisch und ebenfalls positiv definit. Nach dem bekannten Existenz- und Eindeutigkeitsatz von Peano [Ama95] ist (5.14) eindeutig lösbar.

5.5.6 A priori-Schranken

Das folgende Lemma, in dem wir A priori-Schranken herleiten stammt aus [Riv08, S. 75] und wurde hier an die Methode von Fesitauer angepasst.

Lemma 5.5.10. *Es gibt eine positive Konstante C unabhängig von h mit:*

$$\begin{aligned} & \max_{t \in [0, T]} \|u_h(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \|u_h(s)\|_{\mathcal{E}}^2 \\ & \leq C \|u_0\|_{L^2(\Omega)}^2 + C \|f\|_{L^2(0, T; L^2(\Omega))}^2 + C \sum_{\Gamma \in \partial\Omega} \sigma(\Gamma) \|g_D\|_{L^2(0, T; L^2(\Omega))} \end{aligned}$$

Beweis. Wir setzen $v = u_h$ und erhalten dann aus der Variationsformulierung und Koerzivität der Bilinearform

$$\frac{1}{2} \frac{d}{dt} \|u_h(t)\|_{L^2(\Omega)}^2 + \kappa \|u_h(t)\|_{\mathcal{E}}^2 \leq |L(t, u_h)|.$$

Anwendung der Cauchy-Schwarz-Ungleichung auf den rechten Term ergibt

$$\begin{aligned} |L_h(t, u_h(t))| &\leq \|f(t)\|_{L^2(\Omega)} \|u_h(t)\|_{L^2(\Omega)} + D \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t)\|_{L^2(\Gamma)} \|u_h(t)\|_{L^2(\Gamma)} \\ &\quad + D \sum_{\Gamma \in \mathcal{F}_h^D} (\|\nabla u_h(t) \cdot n_\Gamma\|_{L^2(\Gamma)} + \|u_h(t)\|_{L^2(\Gamma)}) \|g_D(t)\|_{L^2(\Gamma)}. \end{aligned}$$

Wir wenden die Spurungleichung an

$$\begin{aligned} |L_h(t, u_h(t))| &\leq \|f(t)\|_{L^2(\Omega)} \|u_h(t)\|_{L^2(\Omega)} + C \|u_h(t)\|_{L^2(\Omega)} \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t)\|_{L^2(\Gamma)} \\ &\quad + C \sum_{\Gamma \in \mathcal{F}_h^D} (D \|\nabla u_h(t)\|_{L^2(\Omega)} + D \|u_h(t)\|_{L^2(\Gamma)}) \|g_D(t)\|_{L^2(\Gamma)} \\ &\stackrel{\text{Defintion Energienorm}}{\leq} \|u_h(t)\|_{L^2(\Omega)} \left(\|f(t)\|_{L^2(\Omega)} + \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t)\|_{L^2(\Gamma)} \right) \\ &\quad + C \|u_h(t)\|_{\mathcal{E}} \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(t)\|_{L^2(\Gamma)}. \end{aligned}$$

Anwendung der Young'schen Ungleichung führt dann zu:

$$\begin{aligned} |L_h(t, u_h(t))| &\leq \frac{1+C^2}{2} \|u_h(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|f(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t)\|_{L^2(\Gamma)}^2 \\ &\quad + \frac{\kappa}{2} \|u_h(t)\|_{\mathcal{E}}^2 + \bar{C} \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(t)\|_{L^2(\Gamma)}^2. \end{aligned}$$

Wir erhalten also:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u_h(t)\|_{L^2(\Omega)}^2 + \kappa \|u_h(t)\|_{\mathcal{E}}^2 &\leq \frac{1+C^2}{2} \|u_h(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|f(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t)\|_{L^2(\Gamma)}^2 \\ &\quad + \frac{\kappa}{2} \|u_h(t)\|_{\mathcal{E}}^2 + \bar{C} \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(t)\|_{L^2(\Gamma)}^2. \end{aligned}$$

Integration von 0 bis $t \in [0, T]$ und Multiplikation mit 2 liefert:

$$\begin{aligned} \|u_h(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \|u_h(s)\|_{\mathcal{E}}^2 ds &\leq (1 + C^2) \int_0^t \|u_h(s)\|_{L^2(\Omega)}^2 ds + \int_0^t \frac{1}{2} \|f(s)\|_{L^2(\Omega)}^2 ds \\ &\quad + \frac{1}{2} \sum_{\Gamma \in \mathcal{F}_h^N} \int_0^t \|g_N\|_{L^2(\Gamma)}^2 ds + 2\bar{C} \sum_{\Gamma \in \mathcal{F}_h^D} \int_0^t \|g_D(s)\|_{L^2(\Gamma)}^2 ds. \end{aligned}$$

Anwendung des Lemmas von Gronwall:

$$\begin{aligned} \|u_h(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \|u_h(s)\|_{\mathcal{E}}^2 ds &\leq e^{(1+C^2)t} \left(\int_0^t \|f(s)\|_{L^2(\Omega)}^2 ds + \sum_{\Gamma \in \mathcal{F}_h^N} \int_0^t \|g_N(s)\|_{L^2(\Gamma)}^2 ds \right. \\ &\quad \left. + \hat{C} \int_0^t \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(s)\|_{L^2(\Gamma)}^2 ds \right) \\ &\leq \tilde{C} \left(\|f\|_{L^2(0,T;L^2(\Omega))} + \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N\|_{L^2(0,T;L^2(\Gamma))} \right. \\ &\quad \left. + \hat{C} \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D\|_{L^2(0,T;L^2(\Gamma))} \right), \end{aligned}$$

wobei $\tilde{C} = e^{(1+C^2)T}$. □

5.5.7 Konvergenzbeweis

Der folgende Satz findet sich in [Fei09]. Die dort vorgestellten Beweisideen werden hier ausgeführt.

Satz 5.5.11. *Unter den Annahmen (A1) erfüllt der Fehler $e_h = u - u_h$ die Abschätzung:*

$$\max_{t \in [0, T]} \|e_h(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \|e_h(s)\|_{\mathcal{E}}^2 ds \leq Ch^{2p}, \text{ für } h \in (0, h_0)$$

mit einer Konstanten C unabhängig von h .

Beweis. Sind u und u_h die exakte bzw. approximierte Lösung, so setzen wir

$$\begin{aligned} \eta(t) &:= \Pi_h u(t) - u(t), \\ \xi(t) &:= u_h(t) - \Pi_h u(t) \text{ für alle } t \in (0, T). \end{aligned}$$

Wegen der Konsistenz des Verfahrens gilt:

$$\begin{aligned} & \left(\frac{\partial(u_h - u)}{\partial t}, \xi(t) \right)_\Omega + a_h^\epsilon(u_h(t) - u(t), \xi(t)) + k(u_h(t) - u(t), \xi(t)) \\ & \quad + J_h^\sigma(u_h(t) - u(t), \xi(t)) = 0 \end{aligned}$$

Wir können $u_h(t) - u(t)$ durch $\xi(t) + \eta(t)$ ersetzen:

$$\begin{aligned} & \left(\frac{\partial(\xi(t) + \eta(t))}{\partial t}, \xi(t) \right)_\Omega + a_h^\epsilon(\xi(t) + \eta(t), \xi(t)) + k(\xi(t) + \eta(t), \xi(t)) \\ & \quad + J_h^\sigma(\xi(t) + \eta(t), \xi(t)) = 0 \end{aligned}$$

Es ergibt sich also mit Hilfe der Koerzivittsungleichung:

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) \\ & \leq - \left(\frac{\partial \eta(t)}{\partial t}, \xi(t) \right)_\Omega - a_h^\epsilon(\eta(t), \xi(t)) - k(\eta(t), \xi(t)) - J_h^\sigma(\eta(t), \xi(t)) \end{aligned}$$

Cauchy-Schwarz und die Dreiecksungleichung liefert:

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) \\ & \leq \left\| \frac{\partial \eta}{\partial t} \right\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} + |a_h^\epsilon(\eta(t), \xi(t))| + k \|\eta\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} + |J_h^\sigma(\eta, \xi)| \end{aligned}$$

Nun knnen wir die einzelnen Terme mit Hilfe von (5.9)-(5.11) aus Lemma 5.5.7 abschtzen:

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) \\ & \leq C \left\{ \left\| \frac{\partial \eta}{\partial t} \right\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} + k \|\eta\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} + h^p |u|_{H^{p+1}(\Omega)} \left(J_h^\sigma(\xi, \xi)^{\frac{1}{2}} + |\xi|_{H^1(\Omega, \mathcal{T}_h)} \right) \right\} \end{aligned}$$

Anwendung des Approximationstheorems 5.5.4 fhrt dann zu:

$$\begin{aligned} & \frac{d}{dt} \|\xi(t)\|_{L^2(\Omega)}^2 + 2\kappa \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) \\ & \leq \tilde{C} \left\{ h^{p+1} \left| \frac{\partial u}{\partial t} \right|_{H^{p+1}(\Omega)} \|\xi\|_{L^2(\Omega)} + h^p |u|_{H^{p+1}(\Omega)} \left(J_h^\sigma(\xi, \xi)^{\frac{1}{2}} + hk \|\xi\|_{L^2(\Omega)} + |\xi|_{H^1(\Omega, \mathcal{T}_h)} \right) \right\} \end{aligned}$$

Mit der Young'schen Ungleichung ergibt sich:

$$\begin{aligned} & \frac{d}{dt} \|\xi(t)\|_{L^2(\Omega)}^2 + 2\kappa \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) \\ & \leq \kappa \left(J_h^\sigma(\xi, \xi) + kh^2 \|\xi\|_{L^2(\Omega)}^2 + |\xi|_{H^1(\Omega, \mathcal{T}_h)}^2 \right) \\ & \quad + \bar{C} \left\{ \frac{1}{4\kappa} h^{2p+2} \left| \frac{\partial u}{\partial t} \right|_{H^{p+1}(\Omega)}^2 + \kappa \|\xi\|_{L^2(\Omega)}^2 + \frac{1}{4\kappa} h^{2p} |u|_{H^{p+1}(\Omega)}^2 \right\} \end{aligned}$$

Wir nehmen an, das $h^2 < 1$ gilt und es folgt:

$$\begin{aligned} & \frac{d}{dt} \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) \\ & \leq \bar{C} \left\{ \frac{1}{4\kappa} h^{2p+2} \left| \frac{\partial u}{\partial t} \right|_{H^{p+1}(\Omega)}^2 + \kappa \|\xi\|_{L^2(\Omega)}^2 + \frac{1}{4\kappa} h^{2p} |u|_{H^{p+1}(\Omega)}^2 \right\} \end{aligned}$$

Integration von 0 bis $t \in [0, T]$ und die Beziehung $\xi = u_h^0 - \Pi_h u^0 = 0$ ergeben dann:

$$\begin{aligned} & \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) dt \\ & \leq \bar{C} \left\{ \frac{1}{4\kappa} h^{2p+2} \int_0^t \left| \frac{\partial u}{\partial t} \right|_{H^{p+1}(\Omega)}^2 dt + \kappa \int_0^t \|\xi\|_{L^2(\Omega)}^2 dt + \frac{1}{4\kappa} h^{2p} \int_0^t |u|_{H^{p+1}(\Omega)}^2 dt \right\} \end{aligned}$$

Es folgt weiter:

$$\begin{aligned} & \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) dt \\ & \leq \bar{C} \left\{ \frac{1}{4\kappa} h^{2p+2} \left\| \frac{\partial u}{\partial t} \right\|_{L^2(0,T;H^{p+1}(\Omega))}^2 + \kappa \int_0^t \|\xi\|_{L^2(\Omega)}^2 dt + \frac{1}{4\kappa} h^{2p} \|u\|_{L^2(0,T;H^{p+1}(\Omega))}^2 \right\} \end{aligned}$$

Anwendung des Lemmas von Gronwall führt weiter zu:

$$\begin{aligned} & \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) dt \\ & \leq \bar{C} h^{2p} \left\{ \frac{1}{4\kappa} h^2 \left\| \frac{\partial u}{\partial t} \right\|_{L^2(0,T;H^{p+1}(\Omega))}^2 + \frac{1}{4\kappa} \|u\|_{L^2(0,T;H^{p+1}(\Omega))}^2 \right\} \exp(\bar{C}\kappa t) \end{aligned}$$

Es gibt also eine Konstante C , die nicht von h , sondern nur von der analytischen Lösung

u abhängt, so dass gilt:

$$\max_{t \in [0, T]} \|\xi(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \left(|\xi|_{H^1(K)}^2 + k \|\xi(t)\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi(t), \xi(t)) \right) dt \leq Ch^{2p}$$

Für den richtigen Fehler folgt dann letztendlich:

$$\begin{aligned} & \max_{t \in [0, T]} \|e_h(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \left(|e_h|_{H^1(K)}^2 + k \|e_h\|_{L^2(\Omega)}^2 + J_h^\sigma(e_h, e_h) \right) dt \\ & \leq \max_{t \in [0, T]} 2 \left(\|\xi(t)\|_{L^2(\Omega)}^2 + \|\eta(t)\|_{L^2(\Omega)}^2 \right) + 2\kappa \int_0^t \left(|\xi|_{H^1(K)}^2 + k \|\xi\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi, \xi) \right) dt \\ & \quad + 2\kappa \int_0^t \left(|\eta|_{H^1(K)}^2 + k \|\eta\|_{L^2(\Omega)}^2 + J_h^\sigma(\eta, \eta) \right) dt \end{aligned}$$

Mit der Fehlerabschätzung für ξ in (5.16) und dem Approximationstheorem 5.5.4 folgt dann das gewünschte Resultat:

$$\max_{t \in [0, T]} \|e_h(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^t \left(|e_h|_{H^1(K)}^2 + k \|e_h\|_{L^2(\Omega)}^2 + J_h^\sigma(e_h, e_h) \right) dt \leq Ch^{2p}$$

□

Korollar. Insbesondere folgt damit auch die Konvergenz in der $L^2(H^1)$ -Norm, das heißt:

$$\|v\|_{L^2(0, T; H^1(\Omega, \mathcal{T}_h))} \leq Ch^p.$$

für eine Konstante C unabhängig von h .

Beweis. Es gibt eine Konstante \bar{C} mit

$$\begin{aligned} \bar{C}h^{2p} & \geq \max_{t \in [0, T]} \|e_h(t)\|_{L^2(\Omega)}^2 + \kappa \int_0^T \|e_h(t)\|_{\mathcal{E}}^2 dt \\ & \geq \frac{1}{T} \int_0^T \|e_h(t)\|_{L^2(\Omega)}^2 dt + \kappa D \int_0^T \|\nabla e_h(t)\|_{L^2(\Omega)}^2 dt \\ & \geq \min\left\{\frac{1}{T}, \kappa D\right\} \int_0^T \|e_h(t)\|_{H^1(\Omega)}^2 dt. \end{aligned}$$

Daraus folgt das gewünschte Resultat. □

5.6 Erhaltungsgleichungen

5.6.1 Lokale Massenerhaltung

Eine interessante Eigenschaft der DG-Verfahren ist die Massenerhaltung auf jedem Gitterelement. Da es keine Stetigkeitseinschränkungen zwischen den Gitterelementen im Raum der gebrochenen Sobolew-Funktionen, können wir eine Testfunktionen wählen, die auf jedem Gitterelement einer anderen Konstante gleicht. Wir fixieren ein Gitterelement $K \in \mathcal{T}_h$ und wählen uns eine Testfunktion $v \in H^k(\Omega, \mathcal{T}_h)$ mit $v \equiv 1$ auf K und $v \equiv 0$ sonst. Unsere Variationsgleichung verkürzt sich dann zu:

$$\frac{d}{dt} \int_K u_h dx - D \sum_{\Gamma \in \partial K} \int_{\Gamma} \langle \nabla u_h \rangle n_{\Gamma} dS + \sum_{\Gamma \in \partial K} \int_{\Gamma} \sigma[u_h] dS + \int_K k u_h dx = \int_K f dx.$$

Stellen wir diese Gleichung etwas um, so erhalten wir letztlich die lokale Massenerhaltungsgleichung:

$$\underbrace{\frac{d}{dt} \int_K u_h dx}_{\text{Zeitliche Änderung der Masse in } K} = D \underbrace{\sum_{\Gamma \in \partial K} \int_{\Gamma} \langle \nabla u_h \rangle n_{\Gamma} dS}_{\text{Flux über die Kanten von } K} + \underbrace{\int_K f dx}_{\text{Masseproduktion in } K} - \underbrace{\int_K k u_h dx}_{\text{Degradation von Masse in } K} - \underbrace{\sum_{\Gamma \in \partial K} \int_{\Gamma} \sigma[u_h] dS}_{\text{Numerische Masse}}.$$

Bei dem Term $\sum_{\Gamma \in \partial K} \int_{\Gamma} \sigma[u_h] dS$ handelt es sich um eine rein numerische Masse. Sie verschwindet, wenn wir σ klein wählen.

5.6.2 Globale Erhaltungsgleichung

Analog können wir auch eine globale Erhaltungsgleichung herleiten. Wir wählen diesmal $v \equiv 1$ in ganz Ω und erhalten:

$$\begin{aligned} \frac{d}{dt} \sum_{K \in \mathcal{T}_h} \int_K u_h dx &= D \sum_{\Gamma \in \mathcal{F}_h^N} \int_{\Gamma} g_N dS + D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \nabla u_h \cdot n_{\Gamma} dS + \int_K (f - k u_h) dx \\ &\quad - \sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \sigma[u_h] dS - D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma(u_h - g_D) dS. \end{aligned}$$

Die Terme $\sum_{\Gamma \in \mathcal{F}_h^I} \int_{\Gamma} \sigma[u_h] dS + D \sum_{\Gamma \in \mathcal{F}_h^D} \int_{\Gamma} \sigma(u_h - g_D) dS$ sind wie oben schon beschrieben rein numerische Massen, die verschwinden, wenn der Strafparameter σ klein gewählt wird.

5.7 Zeitdiskretisierung durch das Theta-Verfahren

Für die Zeitapproximation wollen wir das *Theta-Verfahren* verwenden, welches als Spezialfälle das implizite *Euler-Verfahren* und das *Crank-Nicolson-Verfahren* enthält. Wir betrachten eine Partition $0 = t_0 < t_1 < \dots < t_M = T$ von $[0, T]$. Wir erhalten dadurch Intervalle: $I_m = (t_{m-1}, t_m)$, $\tau_m = t_m - t_{m-1}$, $m = 1, \dots, M$.

Notation:

$$\begin{aligned}\tau &:= \max_{m=0, \dots, M} \tau_m \\ t_{m,\theta} &:= \frac{1}{2}(1+\theta)t_m + \frac{1}{2}(1-\theta)t_{m-1}, \\ v^m(x) &:= v(x, t_m), \text{ für alle } 0 \leq m \leq M \text{ und für alle } x \in \Omega \\ v^{m,\theta}(x) &:= \frac{1}{2}(1+\theta)v^m + \frac{1}{2}(1-\theta)v^{m-1}, \text{ für alle } 0 \leq m \leq M \text{ und für alle } x \in \Omega,\end{aligned}$$

wobei $\theta \in [0, 1]$. Weiterhin definieren wir folgende Normen:

$$\begin{aligned}\|v\|_{\ell^\infty(L^2)} &= \max_{m=0, \dots, M} \|v^m\|_{L^2(\Omega)}, \\ \|v\|_{l^2(H^s)} &= \left(\sum_{m=1}^M \tau_m \|v^{m,\theta}\|_{H^s}^2 \right)^{\frac{1}{2}}, \\ \|v\|_{l^2(\mathcal{E})} &= \left(\sum_{m=1}^M \tau_m \|v^{m,\theta}\|_{\mathcal{E}}^2 \right)^{\frac{1}{2}}.\end{aligned}$$

Das räumliche und zeitliche Variationsproblem hat dann folgende Gestalt: Sei $\{U_m\}_{m=0, \dots, M}$ eine Folge in S_h^p die folgendes erfüllt:

$$\left(\frac{U^m - U^{m-1}}{\tau_m}, v_h \right)_\Omega + a_h^\epsilon(U^{m,\theta}, v_h) + k \cdot (U^{m,\theta}, v_h)_\Omega + J_h^\sigma(U^{m,\theta}, v_h) = L(t_{m,\theta}, v_h) \quad (5.16)$$

$$U^0 = \Pi_h u^0, \quad (5.17)$$

wobei $\theta \in [0, 1]$. Für $\theta = 0$ erhalten wir das Crank-Nicolson-Verfahren, für $\theta = 1$ das implizite Euler-Verfahren.

5.7.1 Existenz und Eindeutigkeit

Aus der Variationsformulierung (5.17) entsteht folgendes Gleichungssystem:

$$\begin{aligned}\left(\mathcal{M} + \tau_m \frac{1}{2}(1+\theta)\mathcal{S} \right) \vec{b}^m &= \left(\mathcal{M} + \tau_m \frac{1}{2}(1-\theta)\mathcal{S} \right) \vec{b}^{m-1} + \tau_m \vec{L}, \\ \vec{b}^0 &= \text{Koeffizientenvektor bezüglich der Basis } \{v_1, \dots, v_N\} \text{ von } U_0.\end{aligned} \quad (5.18)$$

Da sowohl \mathcal{M} als auch \mathcal{S} positiv definit sind, ist es auch $(\mathcal{M} + \tau_m \frac{1}{2}(1 + \theta)\mathcal{S})$. Insbesondere ist $(\mathcal{M} + \tau_m \frac{1}{2}(1 + \theta)\mathcal{S})$ damit auch invertierbar. Das zu lösende Gleichungssystem besitzt damit genau eine Lösung.

5.8 Fehleranalyse für das voll diskretisierte Problem auf nicht konvexen Gebieten und nicht konformen Gittern

In diesem Abschnitt untersuchen wir die Diskretisierung in der Zeit. Als Zeitschrittverfahren wird dabei das Theta-Verfahren untersucht, welches für verschiedene Werte von einem Parameter θ auf das implizite Euler- bzw. Crank-Nicolson-Verfahren führt. Aufbauend auf der Fehleranalyse für das semidiskrete Problem, welches auf ein System von gewöhnlichen nichtlinearen Differentialgleichungen geführt hat, wollen wir nun die Konvergenz des voll diskretisierten Problems untersuchen.

Annahmen (A2):

- (a) Annahmen(A1)
- (b) Regularität der analytischen Lösung:
Für $\theta \in (0, 1]$:

$$u_{tt} \in L^\infty([0, T], H^1(\Omega)).$$

Für $\theta = 0$ (Crank-Nicolson):

$$u_{ttt} \in L^\infty([0, T], H^1(\Omega)).$$

Das folgende Lemma stammt aus [RW00, S.9] und wird für den Konvergenzsatz benötigt.

Lemma 5.8.1.

$$\frac{\Pi_h u^m - \Pi_h u^{m-1}}{\tau_m} = \Pi_h \frac{\partial u^{m,\theta}}{\partial t} + \tau_m \rho^{m,\theta},$$

wobei

$$\|\rho^{m,\theta}\|_2 \leq \tilde{C} \|u_{tt}\|_{L^\infty(I_m, H^1(\Omega))}$$

und für den speziellen Fall $\theta = 0$ gilt sogar:

$$\|\rho^{m,\theta}\|_2 \leq \tilde{C} \tau_m \|u_{ttt}\|_{L^\infty(I_m, H^1(\Omega))}$$

mit Konstanten C und \tilde{C} unabhängig von τ .

Beweis. Wir unterteilen den Beweis in zwei Teile $\theta = 0$ und $\theta \in (0, 1]$:

1.Fall: $\theta \in (0, 1]$

Taylor-Entwicklung um $t_{m,\theta}$ ergibt:

$$\Pi_h u^m = \Pi_h u^{m,\theta} + \frac{1-\theta}{2} \tau_m \Pi_h u_t(t_{m,\theta}) + \frac{1}{2} \left(\frac{1-\theta}{2} \right)^2 \tau_m^2 \Pi_h u_{tt}(t^*) \quad (5.19)$$

$$\Pi_h u^{m-1} = \Pi_h u^{m,\theta} - \frac{1+\theta}{2} \tau_m \Pi_h u_t(t_{m,\theta}) + \frac{1}{2} \left(\frac{1+\theta}{2} \right)^2 \tau_m^2 \Pi_h u_{tt}(t^*) \quad (5.20)$$

Subtrahieren wir (5.20) von (5.19) so erhalten wir:

$$\Pi_h u^m - \Pi_h u^{m-1} = \tau_m \Pi_h u_t(t_{m,\theta}) + \tau_m^2 \rho_{m,\theta},$$

wobei:

$$\rho_{m,\theta} = \frac{1}{2} \left(\left(\frac{1-\theta}{2} \right)^2 - \left(\frac{1+\theta}{2} \right)^2 \right) \Pi_h u_{tt}(t^*).$$

Es folgt:

$$\|\rho_{m,\theta}\|_{L^2(\Omega)} \leq C(\theta) \|\Pi_h u_{tt}\|_{L^\infty(I_m, L^2(\Omega))}$$

Für ein gegebenes $t \in I_m$ gilt nach Approximationssatz:

$$\begin{aligned} \|\Pi_h u_{tt}(t)\|_{L^2(\Omega)} &\leq \|\Pi_h u_{tt}(t) - u_{tt}(t)\|_{L^2(\Omega)} + \|u_{tt}(t)\|_{L^2(\Omega)} \\ &\leq C_A h \|u_{tt}(t)\|_{H^1(\Omega)} + \|u_{tt}(t)\|_{L^2(\Omega)} \end{aligned}$$

Es folgt also:

$$\|\rho_{m,\theta}\|_{L^2(\Omega)} \leq C_A(h+1) \|u_{tt}\|_{L^\infty(I_m; H^1(\Omega))}$$

2.Fall: $\theta = 0$:

Taylor-Entwicklung ergibt:

$$\begin{aligned} \Pi_h u^m &= \Pi_h u^{m,0} + \frac{1}{2} \tau_m \Pi_h u_t(t_{m,0}) + \frac{1}{2} \cdot \frac{1}{4} \tau_m^2 \Pi_h u_{tt}(t_{m,0}) \\ &\quad + \frac{1}{6} \cdot \frac{1}{8} \tau_m^3 \Pi_h u_{ttt}(t^*) \end{aligned} \quad (5.21)$$

$$\begin{aligned} \Pi_h u^{m-1} &= \Pi_h u^{m,0} - \frac{1}{2} \tau_m \Pi_h u_t(t_{m,0}) + \frac{1}{2} \cdot \frac{1}{4} \tau_m^2 \Pi_h u_{tt}(t_{m,0}) \\ &\quad - \frac{1}{6} \cdot \frac{1}{8} \tau_m^3 \Pi_h u_{ttt}(t^*) \end{aligned} \quad (5.22)$$

Subtrahieren wir (5.22) von (5.21) so erhalten wir:

$$\Pi_h u^m - \Pi_h u^{m-1} = \tau_m \Pi_h u_t(t_{m,0}) + \tau_m^2 \rho_{m,0},$$

wobei:

$$\begin{aligned}\rho_{m,\theta} &= \frac{1}{2} \left(\left(\frac{1-\theta}{2} \right)^2 - \left(\frac{1+\theta}{2} \right)^2 \right) \Pi_h u_{tt}(t_{m,\theta}) \\ &\quad + \frac{1}{6} \left(\left(\frac{1-\theta}{2} \right)^3 + \left(\frac{1+\theta}{2} \right)^3 \right) \tau_m \Pi_h u_{ttt}(t^*).\end{aligned}$$

Analog zu dem Beweis für $\theta \in (0, 1]$ folgt dann:

$$\begin{aligned}\|\rho_{m,\theta}\|_{L^2(\Omega)} &\leq C \|\Pi_h u_{ttt}\|_{L^\infty(I_m; H^1(\Omega))} \\ &\leq C(h+1) \|u_{ttt}\|_{L^\infty(I_m; H^1(\Omega))}\end{aligned}$$

□

Lemma 5.8.2 (Diskrete Version des Lemmas von Gronwall). *Sei $(k_m)_{m=0,\dots,M}$ eine nicht-negative Folge und $\varphi_{m=0,\dots,M}$ eine Folge mit*

$$\begin{aligned}\varphi_0 &\leq g_0 \text{ und} \\ \varphi_m &\leq g_0 + \sum_{s=0}^{m-1} p_s + \sum_{s=0}^{m-1} k_s \varphi_s.\end{aligned}$$

Sind $g_0 \geq 0$ und $p_m \geq 0$, dann gilt für $1 \leq m \leq M$:

$$\varphi_m \leq \left(g_0 + \sum_{s=0}^{m-1} p_s \right) \exp \left(\sum_{s=0}^{m-1} k_s \right).$$

Beweis. Siehe [QSS00, S.159].

□

Wir zeigen noch folgendes Lemma:

Lemma 5.8.3. *Es gilt:*

$$\frac{1}{2\tau_m} \left(\|v^m\|_{L^2(\Omega)}^2 - \|v^{m-1}\|_{L^2(\Omega)}^2 \right) \leq \left(\frac{v^m - v^{m-1}}{\tau_m}, v^{m,\theta} \right)_\Omega$$

Beweis.

$$\begin{aligned}
& \left(\frac{v^m - v^{m-1}}{\tau_m}, v^{m,\theta} \right)_\Omega - \frac{1}{2\tau_m} \left(\|v^m\|_{L^2(\Omega)}^2 - \|v^{m-1}\|_{L^2(\Omega)}^2 \right) \\
&= \frac{1}{2\tau_m} \left((v^m - v^{m-1}, (1+\theta)v^m + (1-\theta)v^{m-1})_\Omega - (v^m, v^m)_\Omega + (v^{m-1}, v^{m-1})_\Omega \right) \\
&= \frac{1}{2\tau_m} \left((v^m - v^{m-1}, (1+\theta)v^m + (1-\theta)v^{m-1})_\Omega - (v^m, v^m)_\Omega + (v^{m-1}, v^{m-1})_\Omega \right) \\
&= \frac{1}{2\tau_m} \left(\theta(v^m, v^m)_\Omega - 2\theta(v^m, v^{m-1})_\Omega + \theta(v^{m-1}, v^{m-1})_\Omega \right) \\
&= \frac{\theta}{2\tau_m} \|v^m - v^{m-1}\|_{L^2(\Omega)}^2 \geq 0.
\end{aligned}$$

□

Die folgenden zwei Sätze für die A priori-Schranken und Konvergenz des volldiskreten Problems finden sich für den Fall von ein und zwei Raumdimensionen in [RW00]. Wir übertragen die Beweise auf die Methode von Feistauer [Fei09], wodurch die Einschränkung auf zwei Raumdimensionen entfällt.

Satz 5.8.4 (A priori Schranken). *Unter den Annahmen (A2) erfüllt die Lösung von (5.17) die Abschätzung:*

$$\begin{aligned}
& \|U\|_{\ell^\infty(L^2(\Omega))}^2 + \kappa \sum_{m=1}^M \tau_m \|U^{m,\theta}\|_{\mathcal{E}}^2 \\
& \leq \tilde{C} \left(\|U^0\|_{L^2(\Omega)}^2 + \sum_{m=1}^M \tau_m \|f(t_{m,\theta})\|_{L^2(\Omega)}^2 + \sum_{\Gamma \in \mathcal{F}_h^N} \sum_{m=1}^M \tau_m \|g_N(t_{m,\theta})\|_{L^2(\Gamma)}^2 \right. \\
& \quad \left. + \hat{C} \sum_{\Gamma \in \mathcal{F}_h^D} \sum_{m=1}^M \tau_m \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}^2 \right),
\end{aligned}$$

mit Konstanten \tilde{C} und \hat{C} unabhängig von h und τ .

Beweis. Wir ersetzen in unserer voll diskreten Variationsgleichung v durch $U^{m,\theta}$ und erhalten nach Anwendung der Koerzitivitätsbeziehung und Lemma 5.8.3:

$$\frac{1}{2\tau_m} (\|U^m\|_{L^2(\Omega)}^2 - \|U^{m-1}\|_{L^2(\Omega)}^2) + \kappa \|U^{m,\theta}\|_{\mathcal{E}}^2 \leq |L(t_{m,\theta}, U^{m,\theta})|.$$

Für die rechte Seite folgt nach Anwendung von Cauchy-Schwarz:

$$\begin{aligned}
|L(t_{m,\theta}, U^{m,\theta})| &\leq \|f(t_{m,\theta})\|_{L^2(\Omega)} \|U^{m,\theta}\|_{L^2(\Omega)} + D \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t_{m,\theta})\|_{L^2(\Gamma)} \|U^{m,\theta}\|_{L^2(\Gamma)} \\
&\quad + D \sum_{\Gamma \in \mathcal{F}_h^D} \left(\|\nabla U^{m,\theta} \cdot n_\Gamma\|_{L^2(\Gamma)} + \|U^{m,\theta}\|_{L^2(\Gamma)} \right) \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}.
\end{aligned}$$

Wir wenden die Spurgleichung an und erhalten für eine Konstante C unabhängig von der Feinheit des Gitters h :

$$\begin{aligned}
&\leq \|f(t_{m,\theta})\|_{L^2(\Omega)} \|U^{m,\theta}\|_{L^2(\Omega)} + C \|U^{m,\theta}\|_{L^2(\Omega)} \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t_{m,\theta})\|_{L^2(\Gamma)} \\
&\quad + C \left(D \|\nabla U^{m,\theta}\|_{L^2(\Omega)}^2 + D \sum_{\Gamma \in \mathcal{F}_h^D} \|U^{m,\theta}\|_{L^2(\Gamma)}^2 \right) \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}^2.
\end{aligned}$$

Einsetzen der Definition für die Energienorm liefert

$$\begin{aligned}
&\leq \|f(t_{m,\theta})\|_{L^2(\Omega)} \|U^{m,\theta}\|_{L^2(\Omega)} + C \|U^{m,\theta}\|_{L^2(\Omega)} \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t_{m,\theta})\|_{L^2(\Gamma)} \\
&\quad + C \|U^{m,\theta}\|_{\mathcal{E}} \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}.
\end{aligned}$$

Anwendung der Young'schen Ungleichung führt dann zu:

$$\begin{aligned}
&\leq \frac{1+C^2}{2} \|U^{m,\theta}\|_{L^2(\Omega)}^2 + \frac{1}{2} \|f(t_{m,\theta})\|_{L^2(\Omega)}^2 + \frac{1}{2} \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t_{m,\theta})\|_{L^2(\Gamma)}^2 \\
&\quad + \frac{\kappa}{2} \|U^{m,\theta}\|_{\mathcal{E}}^2 + \bar{C} \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}^2.
\end{aligned}$$

Wir erhalten also:

$$\begin{aligned}
&\frac{1}{2\tau_m} (\|U^m\|_{L^2(\Omega)}^2 - \|U^{m-1}\|_{L^2(\Omega)}^2) + \kappa \|U^{m,\theta}\|_{\mathcal{E}}^2 \\
&\leq \frac{1+C^2}{2} \|U^{m,\theta}\|_{L^2(\Omega)}^2 + \frac{1}{2} \|f(t_{m,\theta})\|_{L^2(\Omega)}^2 + \frac{1}{2} \sum_{\Gamma \in \mathcal{F}_h^N} \|g_N(t_{m,\theta})\|_{L^2(\Gamma)}^2 \\
&\quad + \frac{\kappa}{2} \|U^{m,\theta}\|_{\mathcal{E}}^2 + \bar{C} \sum_{\Gamma \in \mathcal{F}_h^D} \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}^2.
\end{aligned}$$

Multiplikation mit $2\tau_m$ und Summation über $m = 1, \dots, N$ mit $N < M$ liefert:

$$\begin{aligned}
& \|U^N\|_{L^2(\Omega)}^2 - \|U^0\|_{L^2(\Omega)}^2 + \kappa \sum_{m=1}^N \tau_m \|U^{m,\theta}\|_{\mathcal{E}}^2 \\
& \leq \frac{1+C^2}{2} \sum_{m=1}^N \tau_m \|U^{m,\theta}\|_{L^2(\Omega)}^2 + \sum_{m=1}^N \tau_m \|f(t_{m,\theta})\|_{L^2(\Omega)}^2 + \sum_{\Gamma \in \mathcal{F}_h^N} \sum_{m=1}^N \tau_m \|g_N(t_{m,\theta})\|_{L^2(\Gamma)}^2 \\
& \quad + 2\bar{C} \sum_{\Gamma \in \mathcal{F}_h^D} \sum_{m=1}^M \tau_m \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}^2 \\
& \leq \frac{1+C^2}{2} \sum_{m=1}^M \tau_m \|U^{m,\theta}\|_{L^2(\Omega)}^2 + \sum_{m=1}^M \tau_m \|f(t_{m,\theta})\|_{L^2(\Omega)}^2 + \sum_{\Gamma \in \mathcal{F}_h^N} \sum_{m=1}^M \tau_m \|g_N(t_{m,\theta})\|_{L^2(\Gamma)}^2 \\
& \quad + 2\bar{C} \sum_{\Gamma \in \mathcal{F}_h^D} \sum_{m=1}^M \tau_m \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}^2.
\end{aligned}$$

Anwendung des Lemmas von Gronwall

$$\begin{aligned}
& \|U^N\|_{L^2(\Omega)}^2 + \kappa \sum_{m=1}^N \tau_m \|U^{m,\theta}\|_{\mathcal{E}}^2 \\
& \leq \tilde{C} \left(\|U^0\|_{L^2(\Omega)}^2 + \sum_{m=1}^M \tau_m \|f(t_{m,\theta})\|_{L^2(\Omega)}^2 + \sum_{\Gamma \in \mathcal{F}_h^N} \sum_{m=1}^M \tau_m \|g_N(t_{m,\theta})\|_{L^2(\Gamma)}^2 \right. \\
& \quad \left. + \hat{C} \sum_{\Gamma \in \mathcal{F}_h^D} \sum_{m=1}^M \tau_m \|g_D(t_{m,\theta})\|_{L^2(\Gamma)}^2 \right),
\end{aligned}$$

wobei $\tilde{C} = e^{(1+C^2)T}$. □

Satz 5.8.5. *Unter den Annahmen (A2) erfüllt der Fehler die Abschätzung:*

(i) Für $\theta \in (0, 1]$:

$$\|U - u\|_{\ell^\infty(L^2)}^2 + \kappa \sum_{m=1}^N \tau_m \|U^{m,\theta} - u\|_{\mathcal{E}}^2 \leq \tilde{C} h^{2p} + \hat{C} \tau^2, \text{ für } h \in (0, h_0)$$

(ii) Für $\theta = 0$ (Crank-Nicolson):

$$\|U - u\|_{\ell^\infty(L^2)}^2 + \kappa \sum_{m=1}^N \tau_m \|U^{m,\theta} - u\|_{\mathcal{E}}^2 \leq \tilde{C} h^{2p} + \hat{C} \tau^4, \text{ für } h \in (0, h_0)$$

mit Konstanten \tilde{C} und \hat{C} unabhängig von h und τ .

Beweis. Sind u und U die exakte bzw. approximierte Lösung, so setzen wir

$$\begin{aligned}\eta^m &:= \Pi_h u^m - u^m, \\ \xi^m &:= U^m - \Pi_h u^m \text{ für alle } 0 \leq m \leq M.\end{aligned}$$

Wegen der Konsistenz des Verfahrens gilt:

$$\begin{aligned}& \left(\frac{U^m - U^{m-1}}{\Delta t} - \frac{\partial u^{m,\theta}}{\partial t}, \xi^{m,\theta} \right)_\Omega + a_h^\epsilon(U^{m,\theta} - u^{m,\theta}, \xi^{m,\theta}) + \\ & k(U^{m,\theta} - u^{m,\theta}, \xi^{m,\theta}) + J_h^\sigma(U^{m,\theta} - u^{m,\theta}, \xi^{m,\theta}) = 0.\end{aligned}$$

Wir können $U^{m,\theta} - u^{m,\theta}$ durch $\xi^{m,\theta} + \eta^{m,\theta}$ ersetzen:

$$\begin{aligned}& \left(\frac{\xi_h^m - \xi_h^{m-1}}{\tau_m} + \frac{\Pi_h u^m - \Pi_h u^{m-1}}{\tau_m} - \frac{\partial u^{m,\theta}}{\partial t}, \xi^{m,\theta} \right)_\Omega \\ & + a_h^\theta(\xi^{m,\theta}, \xi^{m,\theta}) + k\|\xi^{m,\theta}\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi^{m,\theta}, \xi^{m,\theta}) \\ & = -a_h^\epsilon(\eta^{m,\theta}, \xi^{m,\theta}) - k(\eta^{m,\theta}, \xi^{m,\theta}) - J_h^\sigma(\eta^{m,\theta}, \xi^{m,\theta}).\end{aligned}$$

Mit der Beziehung $\frac{\Pi_h u^m - \Pi_h u^{m-1}}{\tau_m} = \Pi_h \frac{\partial u^{m,\theta}}{\partial t} + \tau_m \rho^{m,\theta}$ aus Lemma 5.8.1 erhalten wir:

$$\begin{aligned}& \left(\frac{\xi_h^m - \xi_h^{m-1}}{\tau_m}, \xi^{m,\theta} \right)_\Omega + a_h^\theta(\xi^{m,\theta}, \xi^{m,\theta}) + k\|\xi^{m,\theta}\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi^{m,\theta}, \xi^{m,\theta}) \\ & = -\left(\frac{\partial \eta^{m,\theta}}{\partial t}, \xi^{m,\theta} \right)_\Omega - a_h^\epsilon(\eta^{m,\theta}, \xi^{m,\theta}) - k(\eta^{m,\theta}, \xi^{m,\theta}) - J_h^\sigma(\eta^{m,\theta}, \xi^{m,\theta}) - \tau_m(\rho^{m,\theta}, \xi^{m,\theta}).\end{aligned}$$

Es ergibt sich mit Hilfe der Koerzitivitätsungleichung und

$$\frac{1}{2\tau_m} \left(\|\xi^m\|_{L^2(\Omega)}^2 - \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) \leq \left(\frac{\xi_h^m - \xi_h^{m-1}}{\tau_m}, \xi^{m,\theta} \right)_\Omega:$$

$$\begin{aligned}& \frac{1}{2\tau_m} \left(\|\xi^m\|_{L^2(\Omega)}^2 - \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) + \kappa \left(\|\xi^{m,\theta}\|_{H^1(K)}^2 + k\|\xi^{m,\theta}\|_{L^2(\Omega)}^2 + J_h^\sigma(\xi^{m,\theta}, \xi^{m,\theta}) \right) \\ & \leq -\left(\frac{\partial \eta^{m,\theta}}{\partial t}, \xi^{m,\theta} \right)_\Omega - a_h^\epsilon(\eta^{m,\theta}, \xi^{m,\theta}) - k(\eta^{m,\theta}, \xi^{m,\theta}) - J_h^\sigma(\eta^{m,\theta}, \xi^{m,\theta}) - \tau_m(\rho^{m,\theta}, \xi^{m,\theta}).\end{aligned}$$

Die gleichen Argumente, wie im Beweis für das stetige Problem führen zu der Abschätzung:

$$\begin{aligned}& \frac{1}{2\tau_m} \left(\|\xi^m\|_{L^2(\Omega)}^2 - \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) + \kappa \|\xi^{m,\theta}\|_{\mathcal{E}}^2 \\ & \leq C \left\{ \frac{1}{4\kappa} h^{2p+2} \left\| \frac{\partial u^{m,\theta}}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + \kappa \|\xi^{m,\theta}\|_{L^2(\Omega)}^2 + \frac{1}{4\kappa} h^{2p} \|u^{m,\theta}\|_{H^{p+1}(\Omega)}^2 \right\} - \tau_m(\rho^{m,\theta}, \xi^{m,\theta}),\end{aligned}$$

wobei $C > 0$ eine Konstante ist. Anwendung von Cauchy-Schwarz und der Young'schen

Ungleichung führen dann zu:

$$\begin{aligned} & \frac{1}{2\tau_m} \left(\|\xi^m\|_{L^2(\Omega)}^2 - \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) + \kappa \|\xi^{m,\theta}\|_{\mathcal{E}}^2 \\ & \leq C \left\{ \frac{1}{4\kappa} h^{2p+2} \left\| \frac{\partial u^{m,\theta}}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + \kappa \|\xi^{m,\theta}\|_{L^2(\Omega)}^2 + \frac{1}{4\kappa} h^{2p} \|u^{m,\theta}\|_{H^{p+1}(\Omega)}^2 \right\} \\ & \quad + \frac{1}{2} \tau_m^2 \|\rho^{m,\theta}\|_{L^2(\Omega)}^2 + \frac{1}{2} \|\xi^{m,\theta}\|_{L^2(\Omega)}^2. \end{aligned}$$

Anwendung der Dreiecksungleichung und der Ungleichung $(a+b)^2 \leq 2(a^2+b^2)$ führt zu:

$$\begin{aligned} & \frac{1}{2\tau_m} \left(\|\xi^m\|_{L^2(\Omega)}^2 - \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) + \kappa \|\xi^{m,\theta}\|_{\mathcal{E}}^2 \\ & \leq C \left\{ \frac{1}{2\kappa} h^{2p+2} \left(\left\| \frac{\partial u^m}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + \left\| \frac{\partial u^{m-1}}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 \right) + 2\kappa \left(\|\xi^m\|_{L^2(\Omega)}^2 + \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) \right. \\ & \quad \left. + \frac{1}{2\kappa} h^{2p} \left(\|u^m\|_{H^{p+1}(\Omega)}^2 + \|u^{m-1}\|_{H^{p+1}(\Omega)}^2 \right) \right\} + \left(\|\xi^m\|_{L^2(\Omega)}^2 + \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) \\ & \quad + \frac{1}{2} \tau_m^2 \|\rho^{m,\theta}\|_{L^2(\Omega)}^2. \end{aligned}$$

Wir multiplizieren mit $2\tau_m$ und summieren über $m = 1, \dots, N$ mit $N \leq M$:

$$\begin{aligned} & \left(\|\xi^N\|_{L^2(\Omega)}^2 - \|\xi^0\|_{L^2(\Omega)}^2 \right) + \sum_{m=1}^N \tau_m 2\kappa \|\xi^{m,\theta}\|_{\mathcal{E}}^2 \\ & \leq C \sum_{m=1}^N 2\tau_m \left\{ \frac{1}{2\kappa} h^{2p+2} \left(\left\| \frac{\partial u^m}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + \left\| \frac{\partial u^{m-1}}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 \right) + 2\kappa \left(\|\xi^m\|_{L^2(\Omega)}^2 + \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) \right. \\ & \quad \left. + \frac{1}{2\kappa} h^{2p} \left(\|u^m\|_{H^{p+1}(\Omega)}^2 + \|u^{m-1}\|_{H^{p+1}(\Omega)}^2 \right) \right\} + \sum_{m=1}^N 2\tau_m \left(\|\xi^m\|_{L^2(\Omega)}^2 + \|\xi^{m-1}\|_{L^2(\Omega)}^2 \right) \\ & \quad + \sum_{m=1}^N \tau_m^3 \|\rho^{m,\theta}\|_{L^2(\Omega)}^2. \end{aligned}$$

ξ^0 ist per Definition null. Auf der rechten Seite fassen wir die Summen weiter zusammen

und addieren die restlichen Summanden von $N + 1, \dots, M$ und erhalten:

$$\begin{aligned}
& \|\xi\|_{\ell^\infty(L^2(\Omega))}^2 + \sum_{m=1}^M \tau_m 2\kappa \|\xi^{m,\theta}\|_{\mathcal{E}}^2 \\
& \leq \bar{C}\tau \sum_{m=0}^M \left\{ \frac{1}{2\kappa} h^{2p+2} \left\| \frac{\partial u^m}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + 2\kappa \|\xi^m\|_{L^2(\Omega)}^2 + \frac{1}{2\kappa} h^{2p} \|u^m\|_{H^{p+1}(\Omega)}^2 \right\} \\
& \quad + \tau \sum_{m=1}^M 2\|\xi^m\|_{L^2(\Omega)}^2 + \tau^3 \sum_{m=1}^M \|\rho^{m,\theta}\|_{L^2(\Omega)}^2.
\end{aligned}$$

Anwendung der diskreten Version des Lemmas von Gronwall führt zu:

$$\begin{aligned}
& \|\xi\|_{\ell^\infty(L^2(\Omega))}^2 + \sum_{m=1}^M \tau_m 2\kappa \|\xi^{m,\theta}\|_{\mathcal{E}}^2 \\
& \leq \tilde{C} h^{2p} \sum_{m=1}^M 2\tau_m \left\{ \frac{1}{2\kappa} h^2 \left\| \frac{\partial u^m}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + \frac{1}{2\kappa} \|u^m\|_{H^{p+1}(\Omega)}^2 \right\} \exp(2T + \bar{C} 2\kappa T) \\
& \quad + \sum_{m=1}^M \tau_m^3 \|\rho^{m,\theta}\|_{L^2(\Omega)}^2.
\end{aligned}$$

Mit dem Lemma 5.8.1 folgt dann:

(i) für $\theta \in (0, 1]$:

$$\begin{aligned}
& \|\xi\|_{\ell^\infty(L^2(\Omega))}^2 + \gamma \|\xi^{m,\theta}\|_{l^2(\mathcal{E})}^2 \\
& \leq \hat{C} h^{2p} \sum_{m=1}^M \tau_m \left\{ \frac{1}{\gamma} h^2 \left\| \frac{\partial u^m}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + \frac{1}{\gamma} \|u^m\|_{H^{p+1}(\Omega)}^2 \right\} \\
& \quad + \check{C} \tau^2 \sum_{m=1}^M \tau_m \|u_{tt}\|_{L^\infty(I_m, L^2(\Omega))}^2.
\end{aligned}$$

(ii) für $\theta = 0$:

$$\begin{aligned}
& \|\xi\|_{\ell^\infty(L^2(\Omega))}^2 + \gamma \|\xi^{m,\theta}\|_{l^2(\mathcal{E})}^2 \\
& \leq \hat{C} h^{2p} \sum_{m=1}^M \tau_m \left\{ \frac{1}{\gamma} h^2 \left\| \frac{\partial u^m}{\partial t} \right\|_{H^{p+1}(\Omega)}^2 + \frac{1}{\gamma} \|u^m\|_{H^{p+1}(\Omega)}^2 \right\} \\
& \quad + \hat{C} \tau^4 \sum_{m=1}^M \tau_m \|u_{ttt}\|_{L^\infty(I_m, L^2(\Omega))}^2.
\end{aligned}$$

□

Korollar. Insbesondere folgt damit auch die Konvergenz in der $\ell^2(H^1)$ -Norm und $\ell^\infty(L^2)$, das heißt: Unter den Annahmen (A2) erfüllt der Fehler die Abschätzung:

(i) Für $\theta \in (0, 1]$:

$$\begin{aligned}\|U - u\|_{\ell^2(H^1)} &\leq \tilde{C}h^p + \hat{C}\tau, \\ \|U - u\|_{\ell^\infty(L^2)} &\leq \tilde{C}h^p + \hat{C}\tau \text{ für } h \in (0, h_0)\end{aligned}$$

(ii) Für $\theta = 0$ (Crank-Nicolson):

$$\begin{aligned}\|U - u\|_{\ell^2(H^1)} &\leq \tilde{C}h^p + \hat{C}\tau^2, \\ \|U - u\|_{\ell^\infty(L^2)} &\leq \tilde{C}h^p + \hat{C}\tau^2 \text{ für } h \in (0, h_0)\end{aligned}$$

mit Konstanten \tilde{C} und \hat{C} unabhängig von h und τ .

Beweis. Wir schätzen die Terme aus dem Konvergenzsatz (Satz 5.8.5) ab:

$$\begin{aligned}\|U - u\|_{\ell^\infty(L^2)}^2 + \kappa\|U - u\|_{\ell^2(\mathcal{E})}^2 &\geq \frac{1}{M}\|U - u\|_{\ell^2(L^2(\Omega))}^2 + \kappa D\|\nabla(U - u)\|_{\ell^2(L^2(\Omega))}^2 \\ &\geq \min\left\{\frac{1}{M}, \kappa D\right\}\|U - u\|_{\ell^2(H^1(\Omega))}^2.\end{aligned}$$

Aus dem vorhergehenden (Satz 5.8.5) folgt dann nach Wurzelziehen das gewünschte Resultat. \square

6 Implementierung

6.1 Das Programm: Dune

Mit Hinblick auf die Anwendung auf unser biologisches Problem benötigen wir eine Software, welche Werkzeuge zur Lösung von partiellen Differentialgleichungen bereitstellt und flexibel an unser Problem angepasst werden kann. Die Software muss in der Lage sein unstrukturierte Gitter zu verwalten, um mit komplexen Geometrien umgehen zu können. Dazu gehört auch, dass es möglich ist Gitter einzulesen, welche wiederum von gängigen CAD-Programmen erzeugt werden können. Das Programm muss zudem in der Lage sein, Ansatz und Testfunktionen für dieses Gitter zu verwalten. Wir brauchen Quadraturregeln und Transformationen auf die entsprechenden Referenzelemente, um die Integrale in der vorgestellten Variationsformulierung (5.17) ausrechnen zu können. Wir wollen in der Lage sein, diese Methoden für zwei- und dreidimensionale Gebiete nutzen zu können. Wir brauchen lineare und nichtlineare Löser, welche parallel auf einem Cluster große Gleichungssysteme lösen können. Die Verwendung eines gängigen Dateiformats für die Visualisierung ist ebenfalls unabdingbar. Die Entwicklung einer solchen Software nimmt für gewöhnlich viele Arbeitsjahre in Anspruch. Deshalb werden wir in dieser Diplomarbeit auf die bereits existierende Software DUNE zurückgreifen. Sie wird unter anderem von Arbeitsgruppen an den Universitäten in Heidelberg, Freiburg, Stuttgart und an der FU Berlin entwickelt. Es handelt sich dabei um eine Open-Source-Software, welche eine effiziente generische Schnittstelle für unterschiedliche Gittermanager für strukturierte und unstrukturierte Gitter enthält und verschieden Methoden für ein weites Spektrum von Problemen bereitstellt [BDE⁺05]. Eine ausführliche Darstellung der generischen Gitterschnittstelle und der Verwendung von parallelen Gittern findet sich in [BBD⁺08] und [BBD⁺07]. Wir verwenden hier insbesondere das Diskretisierungs-Modul DUNE-PDELAB, welches von der Arbeitsgruppe von Prof. Peter Bastian entwickelt wird. Viele gängige Methoden wurden dort bereits implementiert, darunter die klassischen Finite-Volumen-Verfahren, Finite-Elemente-Verfahren und das hier vorgestellte DG-Verfahren. Als Gittermanager verwenden wir hier das Gitter **UGGrid**, welches unstrukturierte Gitter in 2D und 3D verwalten und parallel verwendet werden kann. Einen Überblick über den Leistungsumfang gibt [BJL⁺02]. Wir können in DUNE mit dem **UGGrid** Gitter im MSH Format einlesen, welche in *Gmsh* erzeugt wurden. *Gmsh* wiederum akzeptiert eine Reihe von gängigen CAD-Formaten wie GEO, STEP, IGES, BREP und ACIS.

6.2 Parallelisierung

In DUNE wird unter anderem der überlappende additive Schwarz-Algorithmus für die Parallelisierung verwendet. Diesen Algorithmus wollen wir hier kurz beschreiben. Durch die Parallelisierung erwarten wir erhebliche Geschwindigkeitsvorteile bei der numerischen Berechnung der Lösung von partiellen Differentialgleichungen. Numerische Experimente, in denen die erzielten Geschwindigkeitsvorteile nachgewiesen werden, finden sich in 7. Eine allgemeine ausführliche Übersicht über die aktuellen Verfahren zur Parallelisierung von partiellen Differentialgleichungen findet sich in dem Buch von Toselli und Widlund [TW04]. Der klassische additive Schwarz-Algorithmus wird dort ebenfalls dargestellt.

6.2.1 Der klassische additive überlappende Schwarz-Algorithmus für zwei Gebiete

Wir betrachten die Poisson-Gleichung

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega, \\ u &= 0 \text{ auf } \partial\Omega. \end{aligned} \tag{6.1}$$

Dabei ist $\Omega \subset \mathbb{R}^2$ ein beschränktes Gebiet und f eine gegebene Funktion in $L^2(\Omega)$. Sei

$$\Omega = \Omega_1 \cup \Omega_2, \quad \Omega_1 \cap \Omega_2 = \emptyset \tag{6.2}$$

eine Zerlegung von Ω in zwei disjunkte Teilmengen. Wir wollen eine Überlappung beider Gebiete definieren, das heißt wir erweitern Ω_1 um Punkte, die in Ω_2 liegen und umgekehrt. Wir erweitern also beide Gebiete um die Punkte mit Abstand $\beta > 0$:

$$\hat{\Omega}_i = \{x \in \Omega \mid \text{dist}(x, \Omega_i) < \beta\}, \quad i = 1, 2. \tag{6.3}$$

Wir bezeichnen mit

u^k die Lösung in Ω im k -ten Iterationsschritt,
und u_i^k die Lösung in $\hat{\Omega}_i$ im k -ten Iterationsschritt.

Initialisiere für $k = 0$ mit der Eingangsvermutung $u_i^0 \equiv 0$ für $i \in \{1, 2\}$.

Wiederhole:

Löse für $i = 1, 2$

$$\begin{aligned} -\Delta u_i^k &= f \text{ in } \hat{\Omega}_i, \\ u_1^{k+1} &= u_{2|\partial\hat{\Omega}_1\cap\Omega_2}^k, \\ u_2^{k+1} &= u_{1|\partial\hat{\Omega}_2\cap\Omega_1}^k. \end{aligned}$$

Definiere u^{k+1} als

$$u^{k+1}(x, y) = \begin{cases} u_1^k(x, y) & \text{falls } (x, y) \in \hat{\Omega}_1, \\ u_2^k(x, y) & \text{sonst.} \end{cases}$$

Bis: $k < \text{MaxIterationen}$

Algorithmus 1: Klassischer additiver überlappender Schwarz-Algorithmus für zwei Gebiete.

Die vorgestellte Idee kann auf mehrere Teilgebiete erweitert werden.

6.2.2 Der abstrakte additive überlappende Schwarz-Algorithmus

In wir verwenden in DUNE den abstrakten überlappenden Schwarz-Algorithmus. Eine ausführliche Darstellung der Methode und Aspekte der Implementierung in DUNE finden sich in [BB08]. Eine gute Einführung findet sich auch in [Bas09]. Hier wird dieses Konzept auf die verwendete DG-Methode übertragen und ergänzt. Wir nehmen an, dass wir die Berechnung unserer partiellen Differentialgleichung auf p Prozessoren aufteilen. Für unseren Algorithmus teilen wir das Gebiet Ω in p Teilgebiete auf, das heißt

$$\Omega = \bigcup_{i=1, \dots, p} \Omega_i \text{ mit } \Omega_i \cap \Omega_j = \emptyset \text{ für } i \neq j. \quad (6.4)$$

Wie im vorangegangenen Abschnitt erweitern wir jedes Gebiet Ω_i um alle Punkte mit Abstand $\beta > 0$, um eine Überlappung zu erzeugen. Wir definieren

$$\hat{\Omega}_i = \{x \in \Omega \mid \text{dist}(x, \Omega_i) < \beta\}. \quad (6.5)$$

Durch die Diskretisierung mit dem DG-Verfahren entsteht ein Gleichungssystem (vergleiche (5.18)), dass wir hier vereinfacht schreiben als

$$Ax = b. \quad (6.6)$$

Wir verwenden für das DG-Verfahren eine Monombasis. Bezüglich der Freiheitsgrade definieren wir uns eine Indexmenge

$$I = \text{Indexmenge aller Freiheitsgrade bezüglich } \mathcal{T}_h. \quad (6.7)$$

Der Koeffizientenvektor x hat genau $|I|$ Einträge und $(x)_i, i \in I$ ist der i -te Freiheitsgrad. Da für das DG-Verfahren mit der verwendeten Monombasis die Freiheitsgrade bezüglich der Elemente der Triangulierung \mathcal{T}_h gewählt werden, definieren wir für die überlappende Zerlegung $(\hat{\Omega}_i)_{i=1,\dots,p}$ folgende Indexmengen:

$$\hat{I}_i = \{j \in I \mid j \text{ ist Freiheitsgrad bezüglich eines Elementes } K \in \mathcal{T}_h \text{ mit } K \in \hat{\Omega}_i\}. \quad (6.8)$$

Es werden nun Gleichungen lokal in den Gebieten $\hat{\Omega}_i$ aufgestellt, die dann von dem jeweiligen Prozess gelöst werden. Dazu definieren wir uns eine Restriktion $R_i : \mathbb{R}^I \rightarrow \mathbb{R}^{\hat{I}_i}$ mit

$$(R_i x)_j = (x)_j \text{ für alle } j \in \hat{I}_i.$$

Entsprechend sei $R_i^T : \mathbb{R}^{\hat{I}_i} \rightarrow \mathbb{R}^I$ definiert durch:

$$(R_i x_i)_j = \begin{cases} (x_i)_j, & \text{falls } j \in \hat{I}_i, \\ 0 & \text{sonst} \end{cases}$$

und

$$A_i := R_i A R_i^T = A_{\hat{I}_i, \hat{I}_i}.$$

Der abstrakte additive überlappende Schwarz-Algorithmus ist ein iterativer Algorithmus, wir wollen deshalb erklären was in jedem einzelnen Iterationsschritt passiert. Sei $x^{alt} \in \mathbb{R}^I$ eine beliebige Iterierte. Wir können eine Defektgleichung formulieren

$$Ae = b - Ax^{alt}, \quad (6.9)$$

wobei $e = x - x^{alt}$ der betrachtete Defekt ist. Wir wollen den Defekt lokal für die Indizes aus \hat{I}_i berechnen, das heißt also

$$e = R_i^T e_i \text{ mit } e_i \in \mathbb{R}^{\hat{I}_i}.$$

Einsetzen in (6.9) und Multiplikation mit R_i liefert

$$R_i A R_i^T e_i = R_i (b - Ax^{alt}). \quad (6.10)$$

Eine verbesserte Lösung ergibt sich aus:

$$x^{neu} = x^{alt} + R_i^T e_i = x^{alt} + R_i^T A_i^{-1} R_i (b - Ax^k).$$

Das Berechnen des Vektors x^{neu} entspricht dem einmaligen Lösen unseres Problems im Teilgebiet $\hat{\Omega}_i$ unter Fixieren der aktuellen Randwerte in der Schwarz-Iteration. Summieren wir die Defekte für alle Teilprozessoren auf und multiplizieren sie mit einem Dämpfungsfaktor $\omega > 0$, so erhalten wir den abstrakten additiven Schwarz-Algorithmus.

Initialisiere für $k = 0$ mit der Eingangsvermutung x^0 .

Wiederhole:

Löse

$$x^{k+1} = x^k + \omega \sum_{i=1,\dots,p} R_i^T A_i^{-1} R_i (b - Ax^k).$$

Bis: $k < \text{MaxIterationen}$

Algorithmus 2: Abstrakter additiver überlappender Schwarz-Algorithmus verteilt auf p -Prozesse.

6.2.3 Anwendung des Algorithmus auf das DG-Verfahren

Wir verwenden das unstrukturierte Gitter `UGGrid`. Dieses Gitter kann in DUNE in paralleler Form eingesetzt werden. Bei der parallelen Unterteilung in die Teilgebiete erzeugt `UGGrid` eine Überlappung von einem Element. Das heißt, zu jeder äußeren Kante $\Gamma \in \partial\Omega_i \cap \Omega$ eines Teilgebietes Ω_i gibt es ein benachbartes Element $K_\Gamma \in \mathcal{T}_h \in \Omega \setminus \Omega_i$, welches zur Überlappung gehört. Kanten und Knoten, die außerhalb von Ω_i liegen, gehören nicht zum Überlappungsbereich.

Für diese Art von Überlappung ist das DG-Verfahren besonders geeignet. Wir verwenden für das DG-Verfahren nämlich eine Monombasis, deren Funktionen elementweise definiert werden. Die Gleichungen in der Diskretisierung werden beim DG-Verfahren lokal für jedes Element aufgestellt. Die Freiheitsgrade beziehen sich damit wie oben vorgestellt auf die entsprechenden Elemente, wobei zusätzlich Kopplungsterme bezüglich der Kanten entstehen. Wir verwenden den `ISTLBackend.OVLP_BCGS_SSORk`-Löser. Dies ist ein stabilisierter bikonjugierter Gradientenlöser, welcher mit k Schritten des SSOR-Verfahrens vorkonditioniert wird. Das Verhalten des Verfahrens für unterschiedliche Anzahlen von Teilgebieten werden wir in Kapitel 7 untersuchen.

6.3 Systeme von Reaktions-Diffusions-Gleichungen

Wir haben bereits in 4.1 einen Existenzsatz für die Lösung von Systemen formuliert und auf das Brüsselator-System übertragen. Wir können die Existenz der Lösungen von Gleichungen

chungen der Form

$$u_t - D \cdot \Delta u = f(u) \text{ in } \Omega \times [0, T] \quad (6.11)$$

$$u = 0 \text{ auf } \partial\Omega \times [0, T] \quad (6.12)$$

$$u = u^0 \text{ auf } \Omega \times \{t = 0\} \quad (6.13)$$

beweisen, wenn f Lipschitz-stetig ist. Bei der Implementierung verwenden wir jedoch kein Fixpunktverfahren, sondern koppeln das System durch die Variationsformulierung. Wir betrachten ein gekoppeltes System mit L Komponenten. Der Raum der Ansatz- und Testfunktionen hat die Gestalt eines Produktraumes:

$$S_h^{p,L} = \underbrace{S_h^p(K) \times \cdots \times S_h^p(K)}_{L\text{-mal}}.$$

Wir erhalten damit als semidiskrete Variationsformulierung

$$\begin{aligned} (a) \quad & \vec{u}_h \in \mathcal{C}^1([0, T]; S_h^{p,L}), \\ (b) \quad & \sum_{i=1}^L \left(\frac{\partial u_h^i}{\partial t}, v_h^i \right)_\Omega + \sum_{i=1}^L a_h^\epsilon(u_h^i, v_h) + \sum_{i=1}^L J_h^\sigma(u_h^i, v_h) + \sum_{i=1}^L (f^i(\vec{u}_h), v_h^i)_\Omega = \sum_{i=1}^L L_h(v_h^i), \\ & \forall \vec{v}_h \in S_h^{p,L} \text{ und für fast alle } t \in (0, T), \\ (c) \quad & \vec{u}_h(0) = \vec{u}_h^0. \end{aligned} \quad (6.14)$$

Da wir die Testfunktionen v^i unabhängig voneinander variieren, erhalten wir folgendes System von Gleichungen:

$$\begin{aligned} (a) \quad & \vec{u}_h \in \mathcal{C}^1([0, T]; S_h^{p,L}), \\ (b) \quad & \left(\frac{\partial u_h^i}{\partial t}, v_h^i \right)_\Omega + a_h^\epsilon(u_h^i, v_h) + J_h^\sigma(u_h^i, v_h) + (f^i(\vec{u}_h), v_h^i)_\Omega = L_h(v_h^i), \\ & \forall v_h^i \in S_h^{p,h} \text{ für } i = 1, \dots, L \text{ und für fast alle } t \in (0, T), \\ (c) \quad & \vec{u}_h(0) = \vec{u}_h^0. \end{aligned} \quad (6.15)$$

Wir legen also für jede Komponente das gleiche Gitter zugrunde. Eine Beschreibung der Implementierung dieser Methode findet sich in [BDT10, S. 71-85].

7 Numerische Experimente

In diesem Kapitel sollen verschiedene numerische Tests durchgeführt werden, um die Funktionsfähigkeit des Verfahrens, wie auch der Implementierung nachzuweisen. Im ersten Teil dieses Kapitels lösen wir die instationäre Wärmeleitungsgleichung in einem einfachen Einheitsquadrat und in einem Einheitsquadrat mit Löchern, um die in Kapitel 5 bewiesene Konvergenzordnung in einem konvexen und einem nicht konvexen Gebiet nachzuweisen. Im zweiten Teil dieses Kapitels wollen wir den in Kapitel 6 vorgestellten überlappenden Schwarzalgorithmus auf einem Linux-Cluster testen. Dazu lösen wir die Helmholtz-Gleichung auf der Einheitssphäre in drei Raumdimensionen.

7.1 Konvergenztests für eine parabolische Gleichung

Wir lösen die Wärmeleitungsgleichung auf dem jeweiligen Gebiet Ω mit vorgegebener Dirichlet-Randbedingung:

$$\partial_t u - \Delta u = f \text{ in } \Omega, \quad (7.1)$$

$$u = g \text{ auf } \partial\Omega_D. \quad (7.2)$$

Wir geben eine analytische Funktion vor, mit der wir die approximierte Lösung vergleichen wollen. Als analytische Lösung wählen wir die Funktion

$$u(x, t) = (1 - \exp^{-t}) \exp^{-\|x\|} \text{ für alle } (x, t) \in \Omega \times (0, 1). \quad (7.3)$$

Der Fehler zum Zeitpunkt t_m auf dem Gitter \mathcal{T}_h berechnet sich aus der Differenz der analytischen und approximierten Lösung

$$e_{\tau,h}^m = u(t_m) - U_{\tau,h}^m \text{ für } m = 1, \dots, M. \quad (7.4)$$

Die Konvergenzrate β wird dann durch die Formel

$$\beta = \frac{1}{\ln \frac{h_1}{h_2}} \ln \left(\frac{\|e_{\tau_1, h_1}\|}{\|e_{\tau_2, h_2}\|} \right) \quad (7.5)$$

berechnet, wobei mit $\|\cdot\|$ die entsprechende Norm bezeichnet wird. Wir verwenden für die numerischen Tests das NIPG-Verfahren für den Ort und das implizite Euler-Verfahren für den Zeitschritt. Abhängig vom Polynomgrad p können wir dann in der $\ell^\infty(L^2(\Omega))$ - bzw. $\ell^2(H^1(\Omega))$ -Norm eine Konvergenzordnung von $\mathcal{O}(h^p) + \mathcal{O}(\tau)$ erwarten. Die Schrittweiten τ_1 und τ_2 werden dabei jeweils an die entsprechende Konvergenzordnung angepasst. Das

heißt die Abschätzung für den Fehler im Ort, also der Term $\mathcal{O}(h^p)$ und die Abschätzung für den Fehler in der Zeit, also $\mathcal{O}(\tau)$ sollen bei jedem Verfeinerungsschritt um den gleichen Faktor schrumpfen. Für einen Verfeinerungsschritt von h_1 auf h_2 wählen wir also

$$\tau_2 = \left(\frac{h_2}{h_1}\right)^{\frac{1}{p}} \cdot \tau_1, \quad (7.6)$$

damit das Verhältnis $\left(\frac{h_1}{h_2}\right)^p = \frac{\tau_1}{\tau_2}$ erfüllt wird. Für das unstetige Galerkin-Verfahren müssen wir nach Annahmen (A1) und (A2) einen entsprechenden Strafparameter wählen, um Unstetigkeiten zu bestrafen. Der Strafparameter σ ergibt sich also nach Gl. 5.8 für ein konformes Gitter aus

$$\sigma = \frac{C_W}{h_{K_\Gamma}}, \Gamma \in \mathcal{F}_h. \quad (7.7)$$

Im Fall des untersuchten NIPG-Verfahrens sind wir in der komfortablen Situation, dass wir den Parameter C_W nach den Annahmen (A1) und (A2) einfach nur größer als null wählen müssen um die angestrebte Konvergenzordnung einzuhalten. In diesem Fall wählen wir

$$C_W = 4.0.$$

Wir beginnen den Konvergenztest mit der Lösung der Wärmeleitungsgleichung in einem Einheitsquadrat. Die Lösung sehen wir in Abbildung 7.3. Die Konvergenzentwicklung in der $\ell^\infty(L^2(\Omega))$ - bzw. $\ell^2(H^1(\Omega))$ -Norm für das NIPG-Verfahren im Ort verbunden mit dem impliziten Euler-Verfahren für den Zeitschritt ist in den Tabellen 7.1 und 7.3 ersichtlich. Als nicht konvexes Gebiet wählen wir ein Einheitsquadrat mit Löchern. Die Lösung der Wärmeleitungsgleichung in diesem Gebiet erkennt man in Abbildung 7.3. Die Konvergenzentwicklung hierfür wird in den Tabellen 7.2 und 7.4 dargestellt. Eine entsprechende graphische Darstellung der Ergebnisse für die $\ell^\infty(L^2(\Omega))$ - bzw. $\ell^2(H^1(\Omega))$ ist in den Abbildungen 7.2 bzw. 7.4 zu sehen. Für die numerischen Tests wurden für das Einheitsquadrat wie für das Einheitsquadrat mit Löchern jeweils zwei Gitter erstellt. In **UGGrid** führt nämlich ein Verfeinerungsschritt zu einer Vervierfachung der Anzahl der Gitterelemente. Um Werte für den Fehler dazwischen zu erhalten, wurden für jedes Gebiet zwei Ausgangsgitter erzeugt, die dann schrittweise verfeinert wurden. Außerdem hat man bei einem Gebiet mit Löchern das Problem, dass man eine gewisse Elementzahl benötigt, um die Geometrie auflösen zu können. Es kommt deshalb zu kleinen Fluktuationen in der Konvergenzordnung, sie wird jedoch im Mittel wieder ausgeglichen und erfüllt die theoretische Konvergenzordnung aus Korollar 5.8.

| dofs | τ | $\ e_{\tau,h}\ _{\ell^\infty(0,T,L^2(\Omega))}$ | $\beta_{\ell^\infty(0,T,L^2(\Omega))}$ | $\ e_{\tau,h}\ _{\ell^2(0,T,H^1(\Omega))}$ | $\beta_{\ell^2(0,T,H^1(\Omega))}$ |
|-------|----------|---|--|--|-----------------------------------|
| 168 | 0.1 | 0.000951104 | - | 0.0203287 | - |
| 366 | 0.071 | 0.000641189 | 1.13770686 | 0.0119428 | 1.53475368 |
| 672 | 0.05 | 0.000458032 | 0.970602562 | 0.00723379 | 1.44663499 |
| 1464 | 0.0355 | 0.000298774 | 1.23278758 | 0.00420105 | 1.56800352 |
| 2688 | 0.025 | 0.000228044 | 0.779484789 | 0.00256065 | 1.42847977 |
| 5856 | 0.01775 | 0.000148859 | 1.23073153 | 0.00148345 | 1.57510754 |
| 10752 | 0.0125 | 0.000114636 | 0.753772548 | 0.00090579 | 1.42341557 |
| 23424 | 0.008875 | 7.47912e-05 | 1.23223948 | 0.000524264 | 1.57776626 |
| 43008 | 0.00625 | 5.75344e-05 | 0.756887451 | 0.000320424 | 1.42062254 |

Tabelle 7.1: Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Verfahren für P1-Polynome auf dem Einheitsquadrat.

| dofs | τ | $\ e_{\tau,h}\ _{\ell^\infty(0,T,L^2(\Omega))}$ | $\beta_{\ell^\infty(0,T,L^2(\Omega))}$ | $\ e_{\tau,h}\ _{\ell^2(0,T,H^1(\Omega))}$ | $\beta_{\ell^2(0,T,H^1(\Omega))}$ |
|-------|--------|---|--|--|-----------------------------------|
| 528 | 0.1 | 0.000445305 | - | 0.0154889 | - |
| 1770 | 0.05 | 0.00010629 | 2.06678794 | 0.00537594 | 1.52664575 |
| 3144 | 0.035 | 6.36961e-05 | 1.47745785 | 0.00394308 | 0.894388111 |
| 7080 | 0.025 | 2.78656e-05 | 2.38543965 | 0.00189764 | 2.11023331 |
| 12576 | 0.0176 | 2.05027e-05 | 0.885342614 | 0.0014337 | 0.808926304 |
| 28320 | 0.0125 | 1.45152e-05 | 0.99649892 | 0.000674588 | 1.85743302 |

Tabelle 7.2: Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Zeitschritt für P1-Polynome auf einem Einheitsquadrat mit Löchern.

| dofs | τ | $\ e_{\tau,h}\ _{\ell^\infty(0,T,L^2(\Omega))}$ | $\beta_{\ell^\infty(0,T,L^2(\Omega))}$ | $\ e_{\tau,h}\ _{\ell^2(0,T,H^1(\Omega))}$ | $\beta_{\ell^2(0,T,H^1(\Omega))}$ |
|-------|------------|---|--|--|-----------------------------------|
| 336 | 0.1 | 0.000882622 | - | 0.00158853 | - |
| 732 | 0.05 | 0.000462948 | 1.86189111 | 0.00057515 | 2.93161519 |
| 1344 | 0.025 | 0.000236401 | 1.93922803 | 0.000202097 | 3.01752986 |
| 2928 | 0.0125 | 0.000120344 | 1.94814385 | 7.3306e-05 | 2.92608943 |
| 5376 | 0.00625 | 6.04191e-05 | 1.98817523 | 2.56954e-05 | 3.02484235 |
| 11712 | 0.003125 | 3.04433e-05 | 1.97775988 | 9.27813e-06 | 2.93920829 |
| 21504 | 0.0015625 | 1.51994e-05 | 2.00422076 | 3.24675e-06 | 3.02967552 |
| 46848 | 0.00078125 | 7.63499e-06 | 1.9866324 | 1.1681e-06 | 2.94966503 |

Tabelle 7.3: Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Verfahren für P2-Polynome auf dem Einheitsquadrat.

| dofs | τ | $\ e_{\tau,h}\ _{\ell^\infty(0,T,L^2(\Omega))}$ | $\beta_{\ell^\infty(0,T,L^2(\Omega))}$ | $\ e_{\tau,h}\ _{\ell^2(0,T,H^1(\Omega))}$ | $\beta_{\ell^2(0,T,H^1(\Omega))}$ |
|-------|----------|---|--|--|-----------------------------------|
| 1056 | 0.1 | 0.000135633 | - | 0.000919086 | - |
| 3540 | 0.05 | 6.10499e-05 | 1.31983418 | 0.000236003 | 2.24783708 |
| 14160 | 0.0125 | 1.53009e-05 | 1.99637242 | 2.83392e-05 | 3.05793426 |
| 56640 | 0.003125 | 3.90717e-06 | 1.96942058 | 3.53605e-06 | 3.00258845 |

Tabelle 7.4: Konvergenz des NIPG-Verfahrens in Kombination mit dem impliziten Euler-Verfahren für P2-Polynome auf einem Einheitsquadrat mit Löchern.



Abbildung 7.1: Lösung der Wärmeleitungsgleichung auf dem Einheitsquadrat.

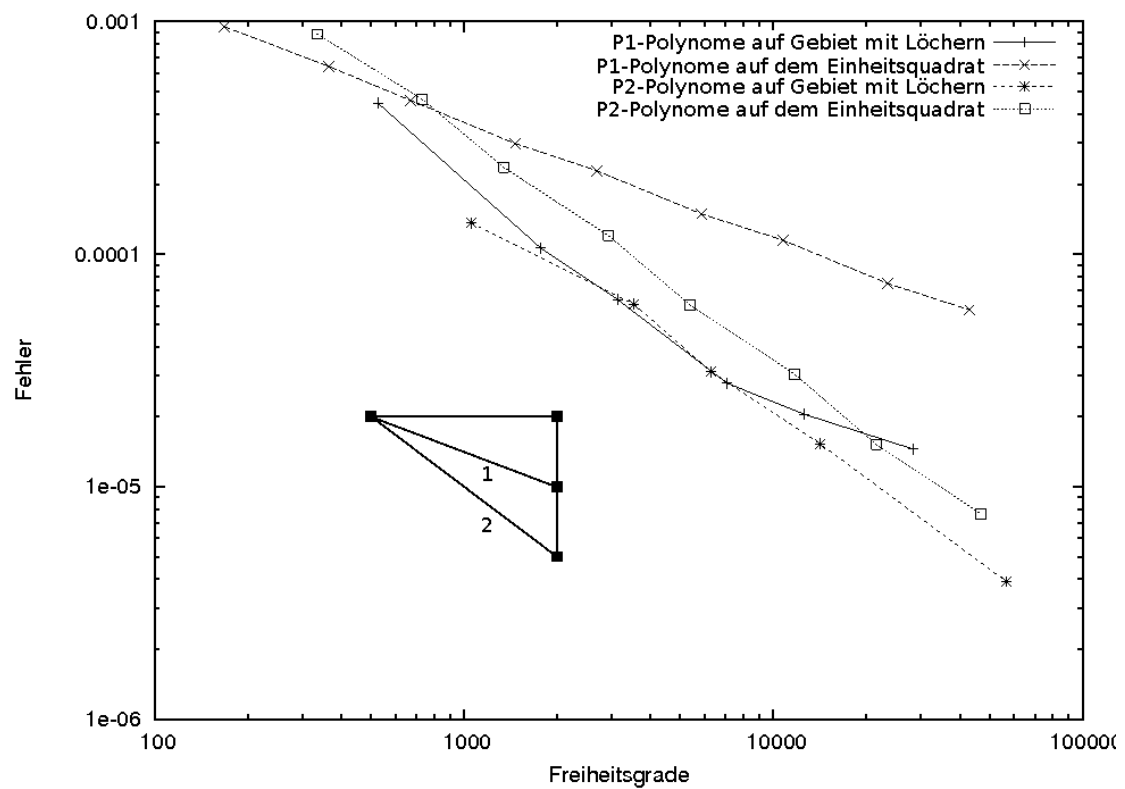


Abbildung 7.2: Konvergenzgraph für die $\ell^\infty(L^2)$ -Norm.

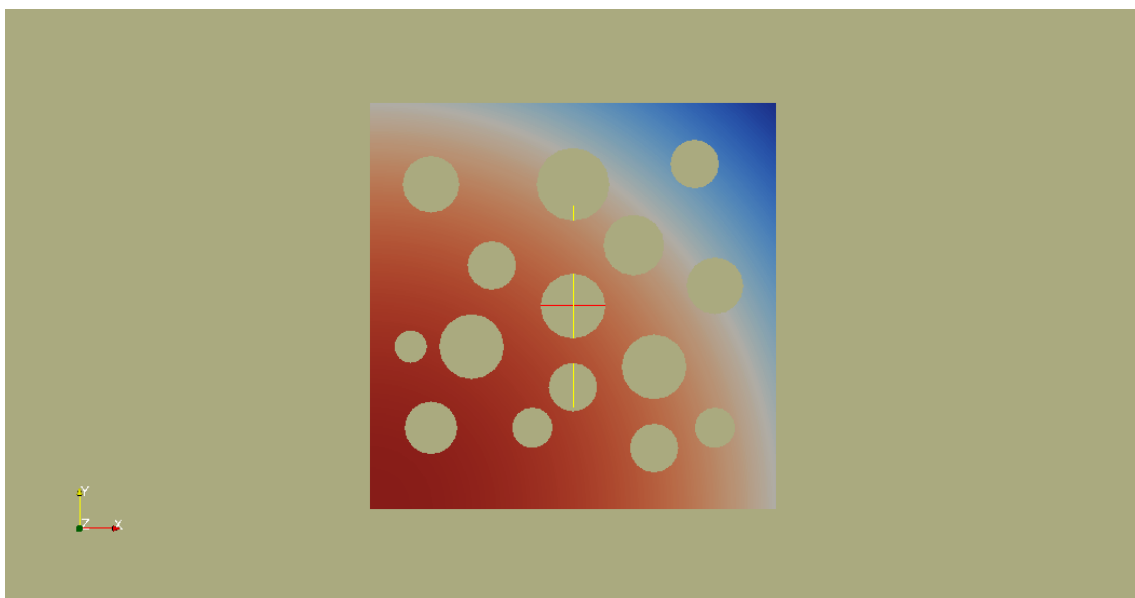
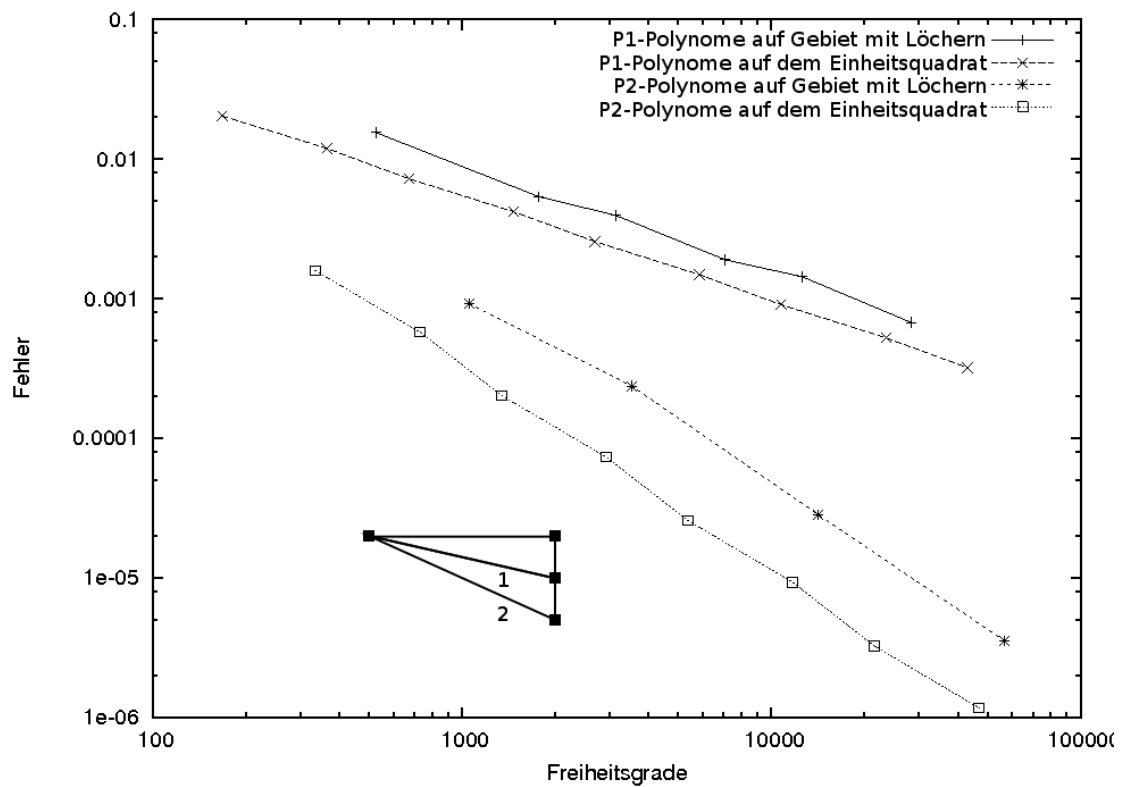


Abbildung 7.3: Lösung der Wärmeleitungsgleichung auf einem Gebiet mit Löchern.

Abbildung 7.4: Konvergenzgraph für die $\ell^2(H^1)$ -Norm.

7.2 Tests auf einem Linux-Cluster mit 256 Kernen

Wir wollen das in Kapitel 6 vorgestellte abstrakte überlappende Schwarz-Verfahren auf einem Linux-Cluster testen. Dazu lösen wir die Helmholtz-Gleichung auf der dreidimensionalen Einheitskugel $\Omega = \{x \in \mathbb{R}^3 \mid \|x - 1\| < 1\}$:

$$-\Delta u + ku = f \text{ in } \Omega, \quad (7.8)$$

$$u = g_D \text{ auf } \partial\Omega. \quad (7.9)$$

Hier wählen wir $k \equiv 1$. Die Lösung der Helmholtz-Gleichung kommt mit entsprechender rechter Seite und Koeffizienten k dem Lösen eines Zeitschrittes mit Hilfe des impliziten Euler-Verfahrens oder der Crank-Nicolson-Methode gleich. Die folgende Tabelle dient also als Richtwert für das Lösen eines Zeitschrittes für ein instationäres Problem.

Als analytische Lösung geben wir die Exponentialfunktion

$$u = \exp(-\|x\|), x \in \bar{\Omega} \quad (7.10)$$

vor und passen entsprechend die Dirichlet-Randwerte an. Die Lösung wird in Abbildung 7.5 und 7.6 veranschaulicht.

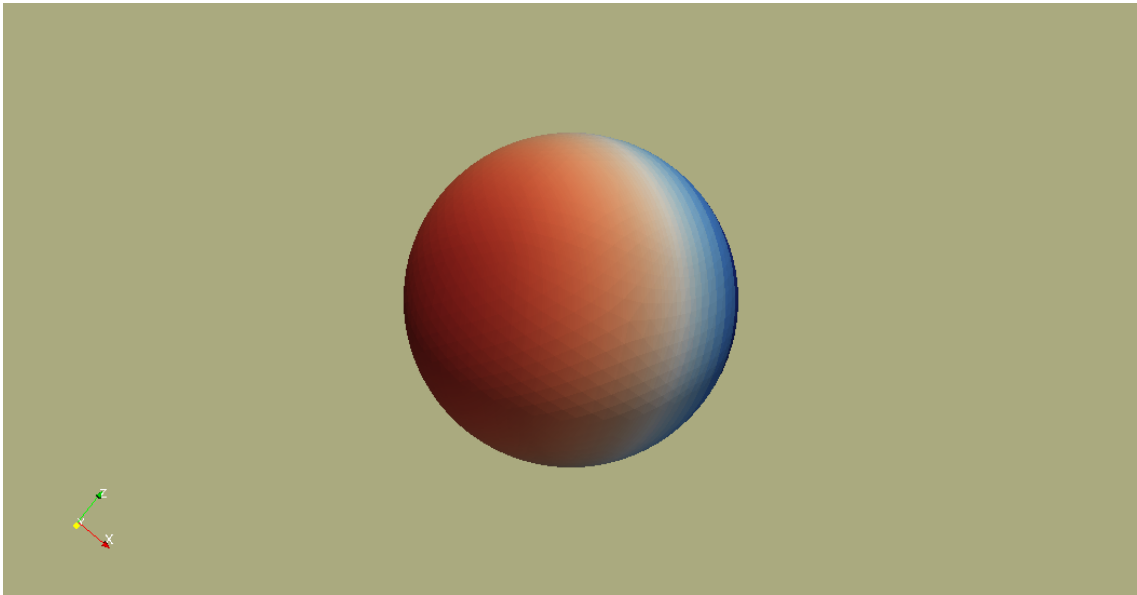


Abbildung 7.5: Lösung der Helmholtz-Gleichung auf der Einheitskugel.

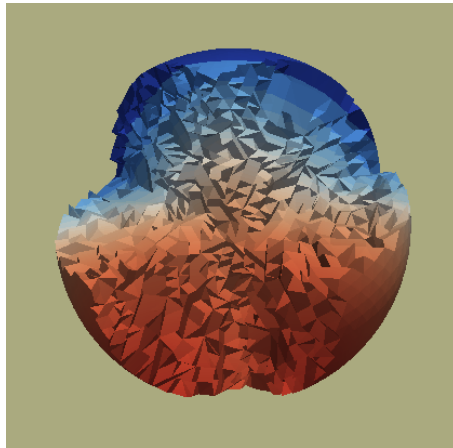


Abbildung 7.6: Ergebnis einer parallelen Rechnung auf 10 Prozessoren. 4 von 10 Teilgebiete wurden entfernt. Man erkennt auch die einzelnen Tetraederelemente.

Auf dem verwendeten Cluster stehen 256 Kerne zur Verfügung. Die Kommunikation zwischen den Prozessoren wird durch das *Message Passing Interface* (MPI) gesteuert. Wir verwenden zwei unterschiedliche Versionen: OpenMPI und MVAPICH. Aus technischen Gründen können bislang mit OpenMPI bis zu 248 und mit MVAPICH bis zu 30 Prozesse gleichzeitig angesteuert werden. In Abbildung 7.7 wird die parallele Gebietsverteilung auf die Prozessoren mit unterschiedlichen Farben veranschaulicht.

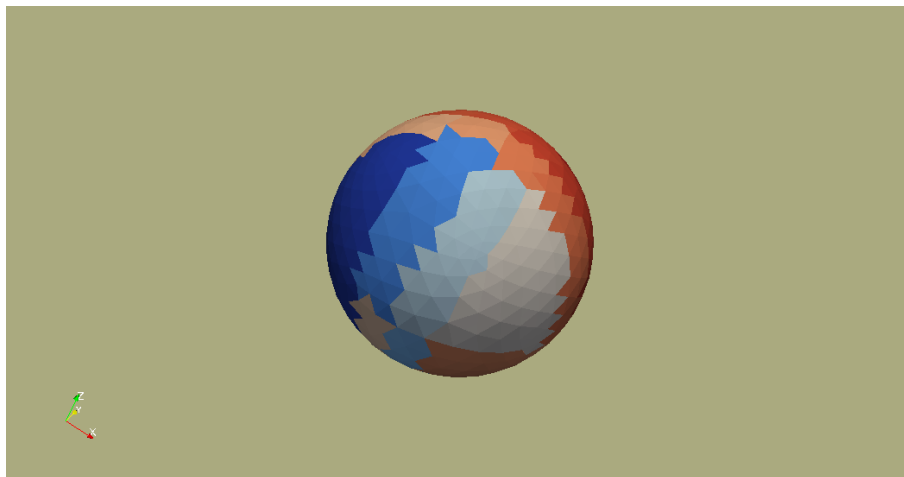


Abbildung 7.7: Eine 3-dimensionale Kugel wird in 24 Teilgebiete auf dem Cluster zerlegt. Die Teilgebiete werden durch unterschiedliche Farben dargestellt.

Das angewendete Schwarz-Verfahren ergibt erhebliche Geschwindigkeitszuwächse bei der parallelen Ausführung auf dem Cluster. In den Tests wurde zwischen der reinen Ausführungszeit für das Lösen des Problems und der Zeit, die für den gesamten Prozess unterschieden. Die Zeit für den gesamten Prozess, beinhaltet neben dem Lösen auch das Einlesen des Gitters und das Versenden der Graphikdateien. Wir haben die Tests für die zwei genannten MPI-Versionen durchgeführt.

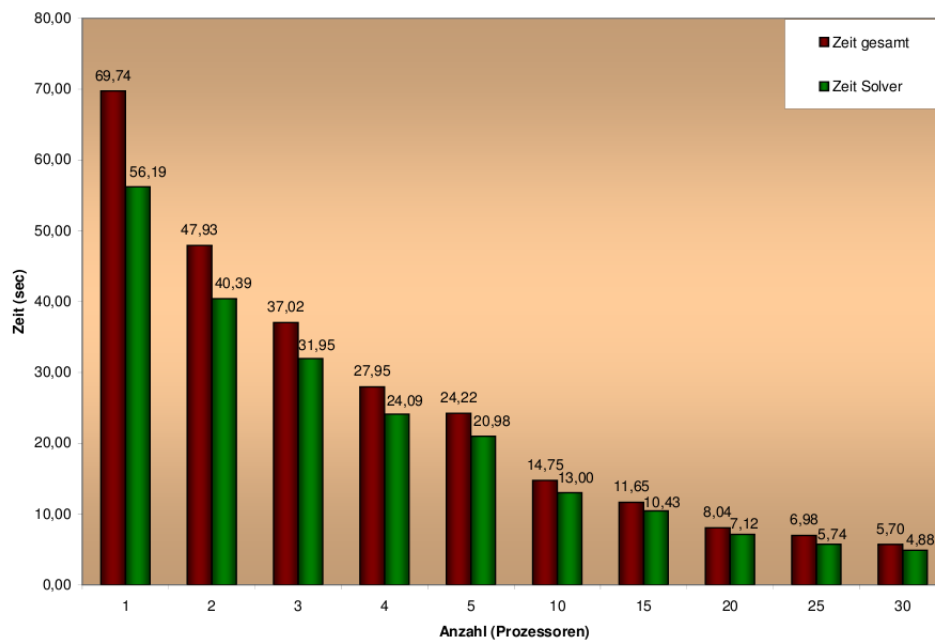


Abbildung 7.8: Abbildung der Rechenzeiten für die Lösung der Helmholtz-Gleichung auf einer Kugel mit 66560 Elementen unter Verwendung von MVAPICH. Die grünen Balken stellen die reine Zeit zum parallelen Lösen des Gleichungssystems dar. Die roten Balken beinhalten den gesamten Prozess vom Verteilen des Gitters, der Assemblierung der Matrix bis zum Versenden der vtu-Bilddateien auf den Head-Knoten.

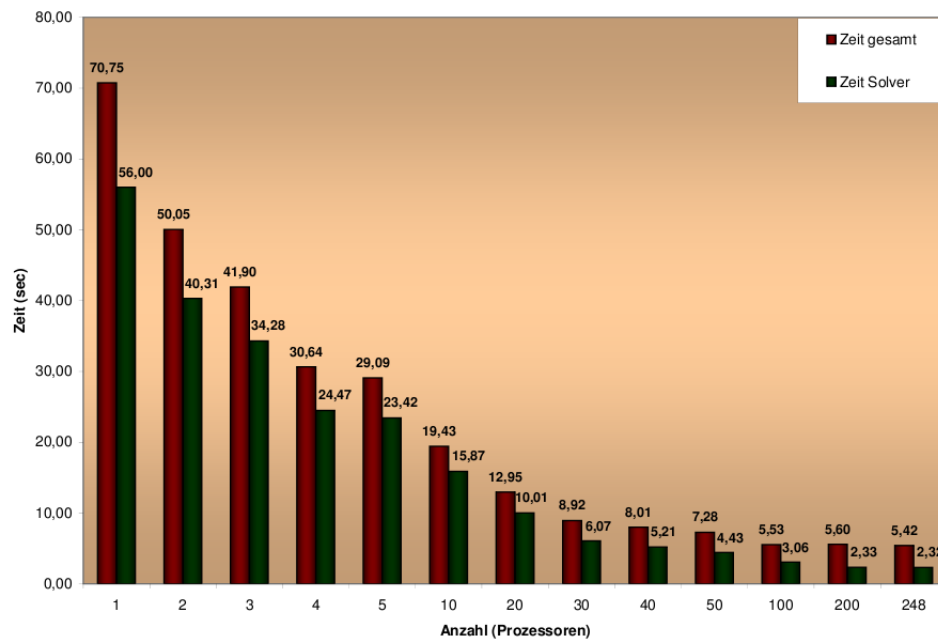


Abbildung 7.9: Abbildung der Rechenzeiten für die Lösung der Helmholtz-Gleichung auf einer Kugel mit 66560 Elementen unter Verwendung von OpenMPI. Die grünen Balken stellen die reine Zeit zum parallelen Lösen des Gleichungssystems dar. Die roten Balken beinhalten den gesamten Prozess vom Verteilen des Gitters, der Assemblierung der Matrix bis zum Versenden der vtU-Bilddateien auf den Head-Knoten.

In den Diagrammen sehen wir, dass MVAPICH bei 30 Prozessoren das Problem zwölfmal schneller löst als bei sequentieller Ausführung. Für OpenMPI verdoppelt sich dieser Geschwindigkeitsvorteil bei der Verwendung von 200 Prozessoren für das Lösen sogar auf den Faktor 24. In Abbildung 7.9 können wir erkennen, dass der Anteil für die Kommunikation der Graphikdateien und das Aufsetzen des Problems gegenüber dem reinen Lösen dominiert.

Bis ca. 100 Prozessoren ergeben sich für den Gesamtprozess, wie auch für das Lösen, erhebliche Geschwindigkeitsvorteile. Bei der Verwendung von mehr als 100 Prozessoren skaliert jedoch unser Problem nicht mehr so gut. So nimmt die Gesamtzeit von 100 auf 200 Prozessoren sogar zu, die Zeit für den Löser fällt gerade noch um ein Viertel. Von 200 auf 248 Prozessoren stagniert die Gesamtzeit, wie auch die Zeit für den Löser. Für die Gesamtzeit scheint die Kommunikation zwischen den Prozessoren ein Engpass zu sein. Doch auch die Lösungsmethode sollte für die Verwendung auf mehreren hundert Prozessoren weiter angepasst werden. Da wir eine Überlappung von einem Element an den Außenkanten der Teilgebiete haben, nimmt die prozentuale Überlappung ab, wenn wir unser Gitter verfeinern. Dies wirkt sich negativ auf die Konvergenz des Verfahrens aus (s. [TW04, S.29ff]), so dass wir mehr Iterationsschritte und damit mehr Zeit benötigen. Im verwendeten überlappenden additiven Schwarz-Verfahren kann man aufgrund von technischen Problemen bislang die Überlappung der einzelnen Teilgebiete nicht vergrößern.

8 Biologische Anwendung in der interzellulären Kommunikation bei Hefezellen

8.1 Biologischer Hintergrund

Die Hefezellen der Bäckerhefe *Saccharomyces cerevisiae* treten in der haploiden (einfacher Chromosomensatz) oder diploiden (doppeltem Chromosomensatz) Form auf. In der haploiden Form gibt es zwei Paarungstypen (mating types) *MATa* und *MAT α* , die zu einer diploiden *MATa/ α* -Zelle verschmelzen können. Das Problem besteht darin, dass sich jeweils zwei Zellen mit entgegengesetztem Paarungstyp finden müssen. Die *MAT α* -Zellen können dazu ein spezifisches Pheromon α -Faktor ausstoßen, welches die *MATa*-Zellen erkennen können. Die *MATa*-Zellen können sich in Richtung dieses Signals verformen, um dann mit einer *MAT α* -Zelle zu verschmelzen. Dies ist jedoch ein hochkomplexer Vorgang, der vom Lokalisieren des Signals bis hin zur Verformung der *MATa*-Zelle auf einer ganzen Reihe von Stoffwechselprozessen und Signalkaskaden beruht. Eine Darstellung des intrazellulären Pheromonsignalweges findet sich in [Bar05]. Die *MATa*-Zelle kann durch Rezeptor-Proteine den α -Faktor an der Zelloberfläche wahrnehmen. Dabei ist die Verteilung des α -Faktors von entscheidender Bedeutung. Die Hefezellen haben keine Möglichkeit sich aktiv zu bewegen, sie müssen also das Signal durch den Gradienten des α -Faktors an der Zelloberfläche wahrnehmen. Durch polarisiertes Wachstum können sich die *MATa*-Zellen in Richtung dieses Gradienten verformen um mit einem Paarungspartner zu verschmelzen [MS98]. Das zielgerichtete Wachstum der *MATa*-Zellen bedingt die inhomogene Verteilung von Signalmolekülen im Inneren der Zellen.

Die Erkennung wird auf mehreren Ebenen unter anderem durch eine Signalkaskade realisiert. Um die Prozesse zu verstehen sind räumliche Modelle von entscheidender Bedeutung. In [CSQC08] wurde die Entstehung, Verstärkung und Aufrechterhaltung eines Gradienten im Inneren der Zelle mit Hilfe eines Systems von RD-Gleichungen beschrieben. Wir wollen in dieser Arbeit daran anknüpfen, wollen dabei jedoch nicht das Innere der Zellen, sondern die Entstehung von Signal-Gradienten außerhalb der Zellen modellieren. Über die Länge einer Zelle gerechnet mögen die Konzentrationsunterschiede an der Zelloberfläche sehr klein sein. Es ist deshalb wichtig zu verstehen, ob und wie der Signalgradient im extrazellulären Bereich verstärkt werden kann. Es wird vermutet, dass die *MATa*-Zellen auf die Erkennung von α -Faktor mit dem Aussenden der Protease Bar1p reagieren. Hinweise darauf finden sich in [Man83]. Die Bar1p kann α -Faktor degradieren. Eine Darstellung der extrazellulären Vorgänge findet sich in Abbildung 8.1. Wir wollen die Verteilung von α -Faktor simulieren und im Modell untersuchen ob und wie Bar1p zur Verstärkung und

räumlichen Lokalisierung der Signal-Gradienten beiträgt.

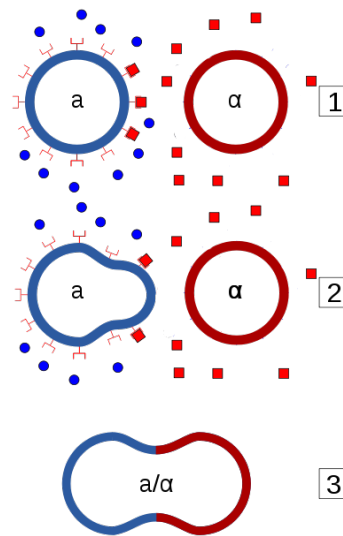


Abbildung 8.1: Vereinigungsprozess der *Saccharomyces cerevisiae*. Die roten Punkte stellen α -Faktor dar, die blauen Punkte stehen für Bar1p. Abbildung übernommen und verändert aus http://en.wikipedia.org/wiki/File:Yeast_mating_scheme.svg.

8.2 Grundlagen der Modellierung

8.2.1 Physikalische Grundlagen von Reaktions-Diffusions-Gleichungen

Reaktions-Diffusions-Gleichungen können auf makroskopischer Ebene Prozesse beschreiben, bei denen Diffusion auftritt und lokale Effekte wie Degradation, Wachstum oder andere Wechselwirkungen stattfinden. Die RD-Gleichungen sind partielle Differentialgleichungen. Diese sind kontinuierlich bzgl. des Raumes. Dennoch können sie auf diskrete physikalische Probleme angewendet werden. Murray [Mur04] beispielsweise verwendet RD-Gleichungen zur Beschreibung der Verteilung von Morphogenen, die notwendig sind um Muster auf Säugetierfellen zu erklären [Mur04, S. 142] und sogar für die Verteilung von Tieren oder Pflanzen, z.B. von Wolfsterritorien [Mur04, S. 722]. In "The chemical basis of morphogenesis" [Tur52] von Alan Turing werden RD-Gleichungen für die Verteilung von verschiedenen Morphogenen verwendet, um Struktur- und Musterbildung zu erklären, welche beim Wachstum von Pflanzen oder der Entwicklung von Tieren auftreten. Diffusion beruht auf der Brown'schen Molekularbewegung der betrachteten Moleküle in einem Medium. Sind die Moleküle ungleich verteilt, so bewegen sich statistisch mehr Moleküle in den Bereich niedriger Konzentration als in den Bereich hoher Konzentration. Makroskopisch betrachtet resultiert ein Nettofluss von Molekülen entlang des Konzentrationsgefälles.

Dieser Sachverhalt kann durch das erste Fick'sche Gesetz beschrieben werden [Cra56, S.2].

$$J = -D\nabla c. \quad (8.1)$$

Dabei ist $[c] = \frac{\text{mol}}{\text{m}^3}$ die Molekülkonzentration, $[J] = \frac{\text{mol}}{\text{m}^2 \cdot \text{s}}$ die Teilchenstromdichte und $[D] = \frac{\text{m}^2}{\text{s}}$ ist der Diffusionskoeffizient. Zur eigentlichen Diffusionsgleichung gelangt man durch Anwendung der Kontinuitätsgleichung

$$\frac{\partial c}{\partial t} + \text{div} J = r(c, x, t). \quad (8.2)$$

Hier steht r für lokale Reaktions- bzw. Quellterme. Einsetzen des ersten Fick'schen Gesetzes in die Kontinuitätsgleichung liefert dann eine RD-Gleichung:

$$\frac{\partial c}{\partial t} - \text{div}(D\nabla c) = r(c, x, t). \quad (8.3)$$

Im Fall von $r \equiv 0$ ist Gl. 8.3 das zweite Fick'sche Gesetz. Für $D \equiv \text{const}$ und $r(c, x, t) = f(x, t) - k(x, t)c$ ergibt sich unsere Reaktionsgleichung aus Gl. 1.1

$$\frac{\partial c}{\partial t} - D\Delta c + kc = f. \quad (8.4)$$

Diffusion ist ein spontan ablaufender Prozess, der ohne Energiezufuhr stattfindet. Man spricht deshalb bei der Diffusion auch von passivem Stofftransport. In einem abgeschlossenen System bewirkt Diffusion den Abbau von Konzentrationsunterschieden bis hin zur vollständigen Homogenität.

In unserem Modell werden wir die Diffusionsprozesse in einem 2-dimensionalen Gebiet untersuchen. Die ablaufenden Prozesse können phänomenologisch weitestgehend auf ein 3-dimensionales Gebiet übertragen werden, es gelten also die gleichen Gesetzmäßigkeiten. Konstanten und Längenmaße können jedoch nicht eins zu eins übertragen werden.

8.2.2 Randbedingungen

Werden die RD-Gleichungen in beschränkten Gebieten gelöst, müssen Randbedingungen an den Gebietsgrenzen vorgeben werden. In unserem Modell für die interzelluläre Kommunikation wird dies besonders wichtig sein, um die Sekretion und Absorption der einzelnen Zellen auf der Zellwand zu beschreiben. Wir verwenden Neumann- und Dirichlet-Randwerte. Sei $\Omega \subset \mathbb{R}^d$ ein Gebiet und $\partial\Omega$ sein Rand. Mit Γ_D bezeichnen wir den Teil des Randes, auf dem die Dirichlet-Randdaten gelten und mit Γ_N den Teil, auf dem die Neumann-Randdaten gelten.

Dirichlet-Randwerte

Wir fixieren die Konzentrationswerte auf dem Rand durch eine gegebene Funktion, deren Werte auf dem Rand in Ort und Zeit variieren können. Wir nennen diese Funktion g_D :

$(0, T) \times \Gamma_D \rightarrow \mathbb{R}$, dabei ist $(0, T)$ ein Zeitintervall. Wir setzen also im Kontext der Gl. 8.4

$$c = g_D \text{ auf } (0, T) \times \Gamma_D. \quad (8.5)$$

Im Besonderen hat die Bedingung

$$c = 0 \text{ auf } (0, T) \times \Gamma_D \quad (8.6)$$

zum Beispiel die physikalische Bedeutung, dass alle Moleküle, die durch die Brown'sche Molekularbewegung den Rand von Ω überschreiten, nicht wieder nach Ω zurückgelangen können [Per10]. Die Gleichung 8.4 wird also zu:

$$\frac{d}{dt} \int_{\Omega} c dx = \int_{\partial\Omega} \frac{\partial c}{\partial n} dS + \int_{\Omega} (f - kc) dx.$$

Setzen wir den Quellterm $f \equiv 0$, dann liefert die Integration über $(0, t)$:

$$\int_{\Omega} c(t) dx = \int_{\Omega} c(0) dx + \int_0^t \int_{\partial\Omega} \frac{\partial c}{\partial n} dS - \int_0^t \int_{\Omega} k c dx.$$

Wegen $c > 0$ ist $\frac{\partial c}{\partial n}$ über den gesamten Zeitraum negativ. Das heißt, die Gesamtkonzentration in Ω nimmt permanent ab.

Neumann-Randwerte

Anstelle von festen Werten auf dem Rand tritt eine Flux-Bedingung. Dies erfolgt in Form einer orts- und zeitabhängigen Funktion $g_N : (0, T) \times \Gamma_N \rightarrow \mathbb{R}$. Wir bestimmen also die Menge an Molekülen, die netto über den Rand hinein- beziehungsweise hinausströmt. Im Kontext von Gl. 8.4 hat die Neumann-Randbedingung folgende Form:

$$\vec{n} \cdot \nabla c = g_N \text{ auf } (0, T) \times \Gamma_N. \quad (8.7)$$

Dabei bezeichnet \vec{n} die ins Gebietsäußere zeigende Einheitsnormale. Beispielsweise ist

$$\vec{n} \cdot \nabla c = 0 \text{ auf } (0, T) \times \partial\Omega \quad (8.8)$$

eine Null-Flux Bedingung. Das heißt, netto verlässt kein Molekül das Gebiet Ω über den Rand und es strömt auch keines hinein. Die Gleichung 8.4 wird also zu:

$$\frac{d}{dt} \int_{\Omega} c dx = \int_{\Omega} (f - kc) dx.$$

Integration über $(0, t)$ liefert

$$\int_{\Omega} c(t) dx = \int_{\Omega} c(0) dx + \int_0^t \int_{\Omega} (f - kc) dx.$$

Die Gesamtkonzentration in Ω wird also nur durch die Quell- und Degradationsterme im Gebietsinneren von Ω beeinflusst.

8.3 Mathematisches Modell

Das mathematische Modell soll die Ausbreitung des α -Faktors sowie den Ausstoß von Bar1p räumlich darstellen. Wir verwenden dazu ein System von RD-Gleichungen.

Die Gleichungen beschreiben die Ausbreitung und Interaktion der beiden genannten Stoffe in dem Gebiet außerhalb der Zellen. Im Modell wird dieser Bereich durch ein beschränktes Gebiet $\Omega \subset \mathbb{R}^2$ dargestellt. Der zeitliche Ablauf beschränkt sich auf das Beobachtungsintervall $[0, T]$. Für die Konzentration des α -Faktors bzw. Bar1p-Pheromons führen wir zwei Funktionen $u : \Omega \times [0, T] \rightarrow \mathbb{R}$ bzw. $v : \Omega \times [0, T] \rightarrow \mathbb{R}$ ein. Das Modell besteht aus zwei gekoppelten partiellen Differentialgleichungen

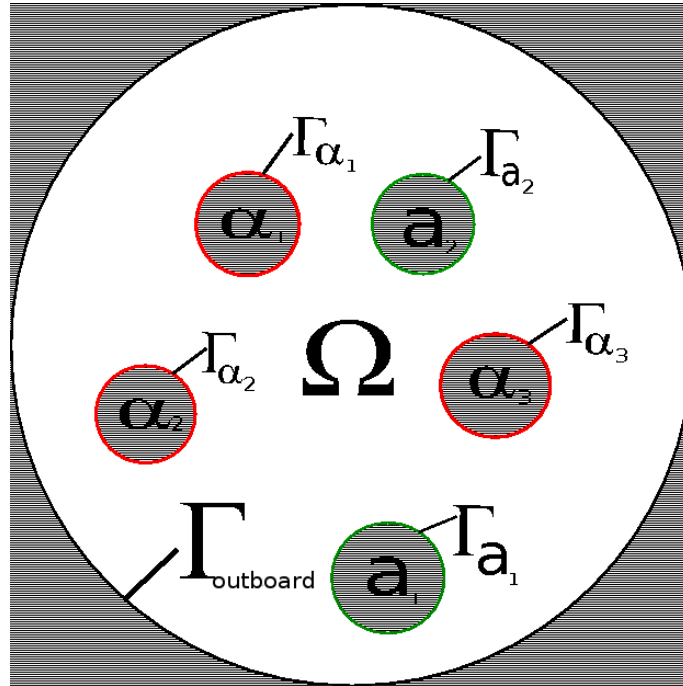
$$\frac{\partial u}{\partial t} - D_u \Delta u = -kuv \text{ in } (0, T) \times \Omega, \quad (8.9)$$

$$\frac{\partial v}{\partial t} - D_v \Delta v = 0 \text{ in } (0, T) \times \Omega \quad (8.10)$$

mit den Variablen und Parametern

$$\begin{aligned} u &= [\text{Konzentration von } \alpha\text{-Faktor}], \\ v &= [\text{Konzentration von Bar1p}], \\ D_u &= [\text{Diffusionsrate für } \alpha\text{-Faktor}], \\ D_v &= [\text{Diffusionsrate für Bar1p}], \\ k &= [\text{Degradationsrate für } \alpha\text{-Faktor}]. \end{aligned}$$

In dem durch Ω beschriebenen Bereich kann das Bar1p frei durch das die Zellen umgebende Medium diffundieren. Dieser Prozess wird in Gleichung (8.10) durch eine Diffusionsgleichung beschrieben, wobei wir für diesen Prozess eine konstante Diffusionsrate D_v annehmen. Das α -Faktor kann ebenfalls frei durch das Medium mit der Diffusionsrate D_u diffundieren. Es wird mit der Abbaurate k abgebaut, wenn es mit Bar1p zusammentrifft. Dies wird in Gleichung (8.9) durch den Reaktionsterm $(-kuv)$ beschrieben. Wir haben

Abbildung 8.2: Schematische Darstellung des Gebietes Ω .

gemischte Dirichlet- und Neumann-Randbedingungen

$$\begin{aligned} u &= g_D^u \text{ auf } (0, T) \times \Gamma_D^u, & v &= g_D^v \text{ auf } (0, T) \times \Gamma_D^v, \\ \frac{\partial u}{\partial n} &= g_N^u \text{ auf } (0, T) \times \Gamma_N^u, & \frac{\partial v}{\partial n} &= g_N^v \text{ auf } (0, T) \times \Gamma_N^v \end{aligned}$$

und die Startbedingungen

$$u, v = 0 \text{ auf } \{0\} \times \Omega.$$

Die Randbedingungen werden im Folgenden formuliert und näher erläutert.

Randbedingungen für das α -Faktor

Um den Ausstoß von α -Faktor durch die MAT α -Zellen zu beschreiben verwenden wir eine Dirichlet-Randbedingung. Dabei setzen wir mit g_α die Werte von u auf dem Rand fest. Die Absorption des α -Faktors am Zellrand der MAT α -Zelle kann durch eine Neumann-Bedingung ausgedrückt werden. Die Absorption ist dabei linear abhängig von der Differenz des inneren α -Faktorniveaus der jeweiligen MAT α -Zelle zum Zeitpunkt $(t - \tau)$ und des Niveaus von α -Faktor am Zellrand; σ ist dabei die maximale Absorption. Auf dem äußeren

Gebietsrand nehmen wir an, dass das α -Faktor abprallt, also kein α -Faktor das Gebiet Ω durch den Außenrand verlassen kann.

(a) **Randbedingungen auf den MAT α -Zellen**

Die Abgabe von α -Faktor durch MAT α -Zellen wird durch eine Dirichlet-Bedingung ausgedrückt:

$$u = C_{\alpha_i} \text{ auf } (0, T) \times \Gamma_{\alpha_i}^u \text{ für alle } i \in \mathbb{N}.$$

Dabei bestimmt C_{α_i} die Sekretionsrate von α -Factor auf dem Rand der i -ten α -Zelle.

(b) **Randbedingungen auf dem äußeren Gebietsrand**

Auf dem äußeren Rand des Gebietes nehmen wir eine Null-Flux-Bedingung an:

$$\frac{\partial u}{\partial n} = 0 \text{ auf } (0, T) \times \Gamma_{\text{Außen}}. \quad (8.11)$$

(c) **Randbedingungen auf den MAT α -Zellen**

Der Flux von α -Faktor auf dem Zellrand der i -th MAT α -Zelle hängt von der Konzentration von α -Faktor auf dem Zellrand und der Menge von bereits gebundenem α -Faktor auf der Zellwand ab:

$$A(t) = \int_{t-\tau_{\text{rez}}}^t \int_{\Gamma_{\alpha_i}} \frac{\partial u}{\partial n} dS,$$

$$\frac{\partial u}{\partial n}(t) = -\sigma \left(1 - \frac{A(t)}{A_{\text{max}}}\right) u \text{ on } (0, T) \times \Gamma_{\alpha_i} \text{ für alle } i \in \mathbb{N}.$$

Dabei verwenden wir die folgenden Variablen und Parameter:

$\tau_{\text{rec}} = [\text{Erholungszeit der Rezeptoren auf der Zellwand, um nach Aktivierung erneut } \alpha\text{-Faktor binden zu können }],$

$\sigma = [\text{Maximaler Flux für die Absorption von } \alpha\text{-Faktor}],$

$A_{\text{max}} = [\text{Maximale Gesamtabsorption von } \alpha\text{-Faktor}].$

Randbedingungen für Bar1p

(a) **Auf den MAT α -Zellen**

Die Sekretion von Bar1p auf der i -ten MAT α -Zelle hängt von der Gesamtabsorption von α -Faktor auf der Zellwand ab. Wir nehmen eine verzögerte Antwort an, dass heißt die Sekretion von Bar1p findet erst nach der Zeitspanne τ_{Bar1p} nach der Aktivierung

statt:

$$v(t) = C_{\text{Bar1p}} \frac{A(t)}{A(t) + K_M} \text{ on } (0, T) \times \Gamma_{a_i} \text{ für alle } i \in \mathbb{N},$$

$$\text{wobei } A(t) = \int_0^{t-\tau_{\text{Bar1p}}} \int_{\Gamma_{a_i}} \frac{\partial u}{\partial n} dS.$$

(b) **Auf den MAT α -Zellen und dem äußeren Rand des Gebietes**

Auf dem Rand der MAT α -Zellen und auf dem äußeren Gebietsrand nehmen wir eine Null-Flux-Bedingung an:

$$\frac{\partial v}{\partial n} = 0.$$

Die auftretenden Variablen und Parameter sind

$$\begin{aligned} \tau_{\text{Bar1p}} &= [\text{Zeitverzögerung für die Bar1p-Antwort}], \\ C_{\text{Bar1p}} &= [\text{Maximaler Sekretionsrate von Bar1p}], \\ K_M &= [\text{Michaelis-Menten-Konstante}]. \end{aligned}$$

8.4 Numerische Resultate und Interpretation des Modells

Die Simulation der Prozesse basiert auf dem Programm DUNE und DUNE-PDELAB. Wir haben dabei das NIPG-Verfahren mit Polynomgrad $p = 2$ im Ort und das implizite Crank-Nicolson als Zeitschrittverfahren verwendet. Die Gitter wurden mit dem Gittergenerierungsprogramm *Gmsh* erzeugt. Die Funktionsweise von *Gmsh* in Bezug auf Algorithmen und Implementierung wird in [GR08] dargestellt.

Verstärkung der entfernungsabhängigen Absorption von α -Faktor durch Bar1p

Für dieses numerische Experiment verwenden wir ein künstlich generiertes Gitter mit drei Zellen. Das Gitter wurde mit *Gmsh* erstellt. Alle drei Zellen haben einen Radius von $r = 2.5\mu\text{m}$. Die Verteilung der Zellen ist in Abbildung 8.3 zu sehen. Die Mittelpunkte der MAT α -Zellen haben einen Abstand von $8\mu\text{m}$ bzw. $16\mu\text{m}$ zur in der Mitte gelegenen MAT α -Zelle. Das kreisförmige Gebiet hat einen Radius von $40\mu\text{m}$. Eine Übersicht der

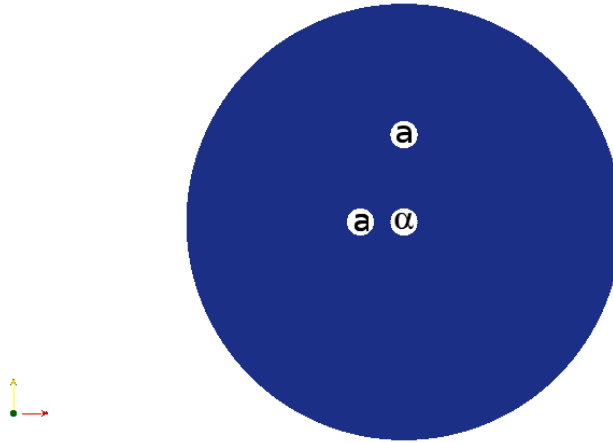


Abbildung 8.3: Verteilung der MATa/α-Zellen.

verwendeten Parameter werden nachfolgend aufgeführt:

$$\begin{aligned}
 D_{\alpha} &= 324.03 \frac{\mu m^2}{s}, \\
 D_{\text{Bar1p}} &= 93.38 \frac{\mu m^2}{s}, \\
 k &= 28, \\
 C_{\alpha_i} &= 11.64 \frac{\text{molec}}{\mu m} \text{ für alle MAT}\alpha\text{-Zellen,} \\
 C_{\text{Bar1p}} &= 2.76 \frac{\text{molec}}{\mu m}, \\
 \tau_{\text{Bar1p}} &= 2.0s, \\
 \sigma &= 0.1 \frac{1}{\mu m}, \\
 A_{\text{max}} &= 10.0 \text{molec}, \\
 \tau_{\text{rec}} &= 1.0s, \\
 K_M &= 1.0 \text{molec}.
 \end{aligned}$$

Die Parameter wurden von Christian Diener aus der Arbeitsgruppe von Prof. Edda Klipp zur Verfügung gestellt. Wir wollen unser Augenmerk auf die qualitativ ablaufenden Prozesse legen. Das Beobachtungsintervall erstreckt sich über einen Zeitraum von 20s. In der Abbildung 8.4 sieht man die Verteilung von α-Faktor. Die Konzentration wird durch eine logarithmische Farbenskala wiedergegeben, wie rechts im Bild zu erkennen ist. In Bild 1 sieht man wie die MATα-Zelle α-Faktor abgibt. Durch die hohe Diffusionsrate breitet

sich der α -Faktor relativ schnell im Medium aus. Dies ist in der Abbildung 8.4 durch die rote Farbe zu erkennen. Die beiden MAT α -Zellen absorbieren den α -Faktor über die Zelloberfläche. Nach einer kurzen Zeitverzögerung reagieren beide Zellen mit dem Ausstoß von Bar1p. Die Zelle, welche näher an der MAT α -Zelle liegt, absorbiert mehr als die weiter entfernt liegende und reagiert deshalb heftiger. Der α -Faktor wird durch das ausgestoßene Bar1p degradiert. In dem Bereich hinter den beiden MAT α -Zellen geht die *alpha*-Faktor-Konzentration besonders stark zurück. Wie man an der blauen Einfärbung in der Simulation sehen kann, geht die Konzentration von *alpha*-Faktor dort fast auf Null zurück. In Abbildung 8.5 erkennt man auch, wie die Bar1p-Sekretion der weiter entfernt liegenden Zelle zurückgeht und sich gleichzeitig Bar1p weiter im Medium verteilt. Dieses Bild entsteht nach 15s. Nach 20s erreicht die Simulation einen stationären Zustand. Die α -Konzentration lokalisiert sich in der Region um die MAT α -Zelle. In der Abbildung 8.4 sieht man, dass der Gradient nur für die nahe gelegene MAT α -Zelle wahrnehmbar ist.

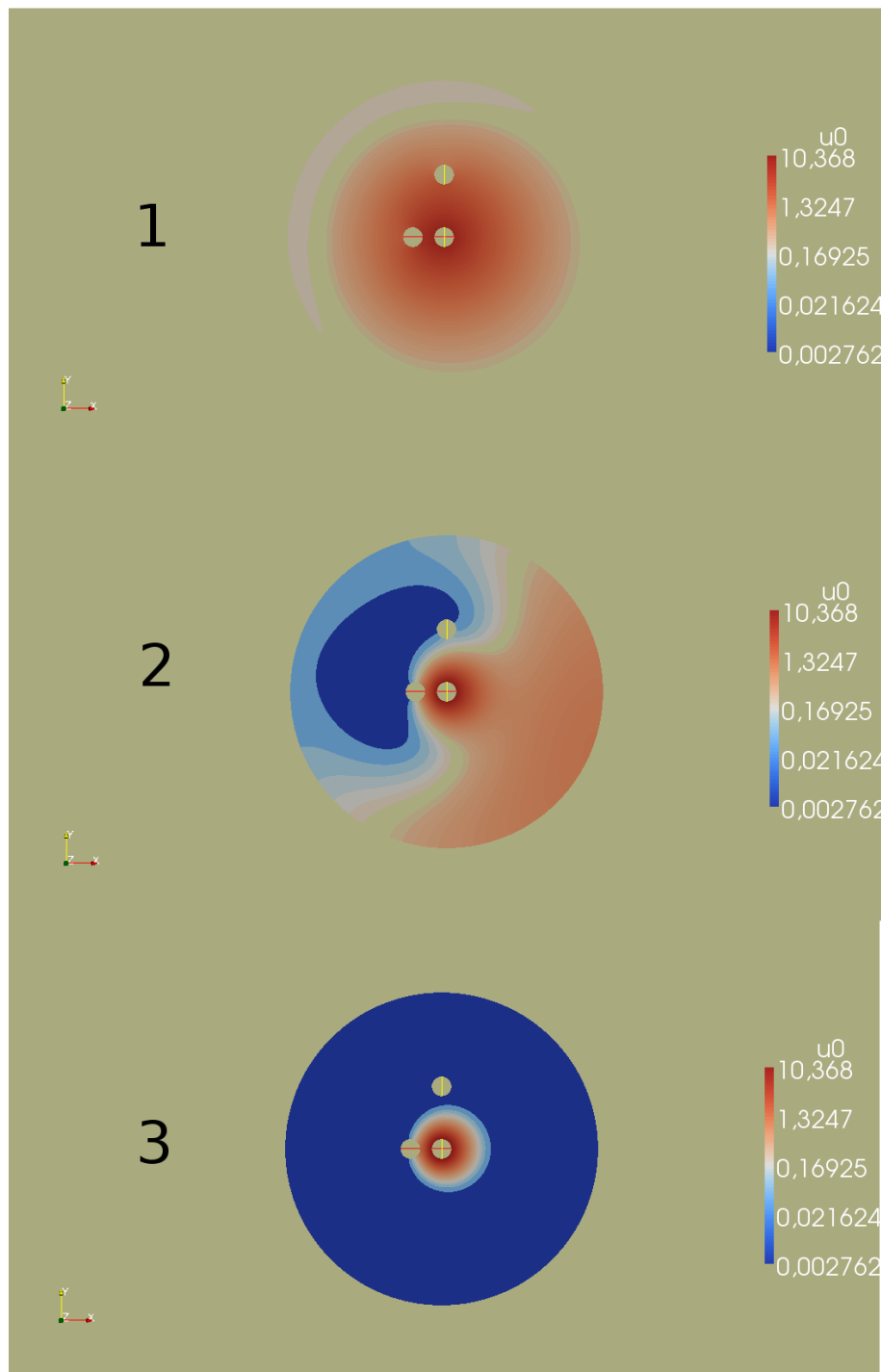


Abbildung 8.4: Verteilung von Alpha-Faktor nach 2s, 10s und 20s.

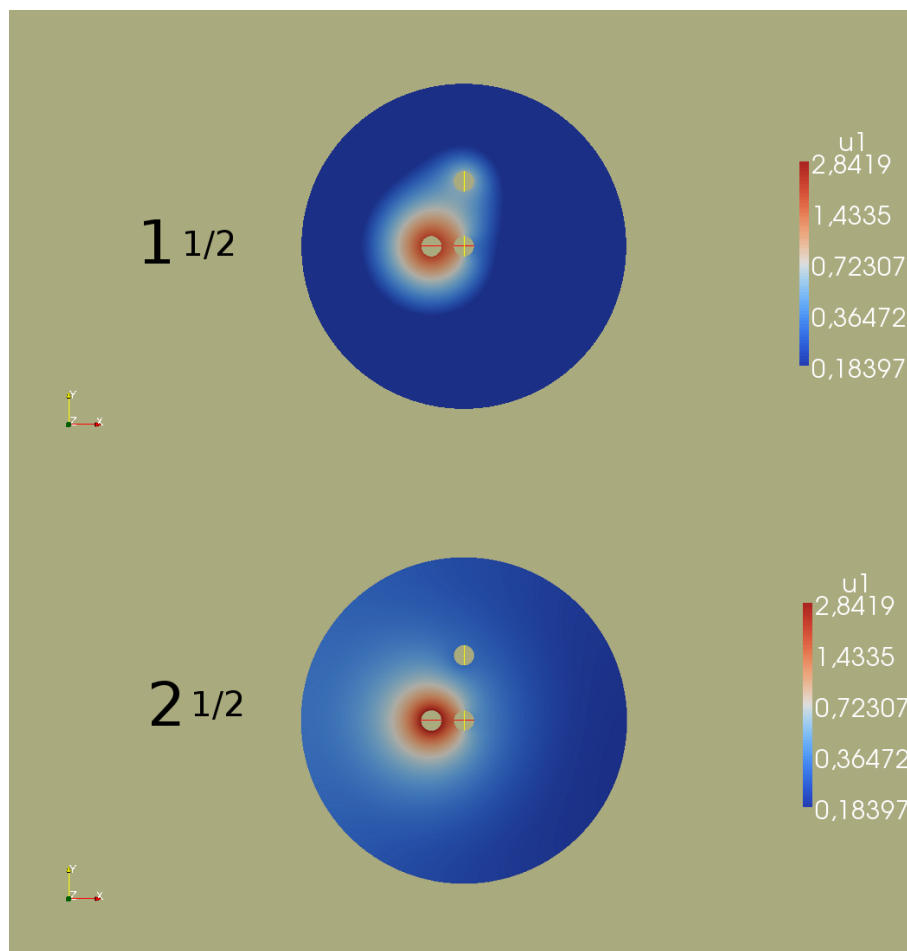


Abbildung 8.5: Verteilung von Bar1p nach 5 und 15s.

Wirkung von Bar1p in einer größeren Population

Diesmal wollen wir in einer Simulation eine größere Population von Zellen beobachten. Zellpositionen werden aus einem Mikroskop-Bild (s. Abbildung 8.6) herausgelesen und daraus wird ein Gitter (s. Abbildung 8.7) erzeugt. Die Zellen wurden in dieser Simulation durch Kreise approximiert. Der Radius variiert je nach Zellvolumen von $2.0\mu m$ bis $3.0\mu m$. Das betrachtete Kompartiment für die Simulation hat einen Radius von $100\mu m$. Wir betrachten die Interaktion von 174 Zellen, wovon die Hälfte $MAT\alpha$ -Zellen und die andere Hälfte $MATa$ -Zellen sind.

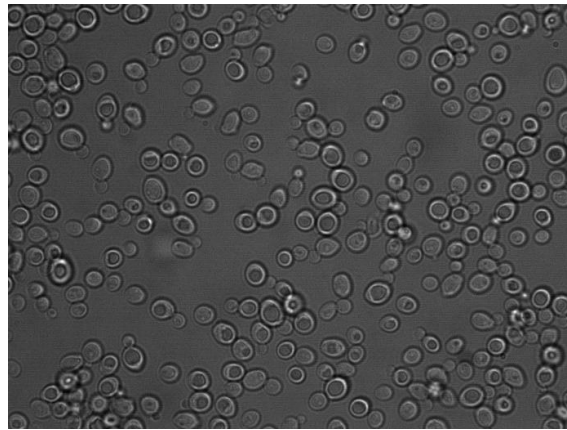


Abbildung 8.6: Mikroskopbild. Das Bild wurde von Dr. Gabriele Schreiber (AG Prof. Andreas Herrmann) und Christian Diener bereitgestellt.

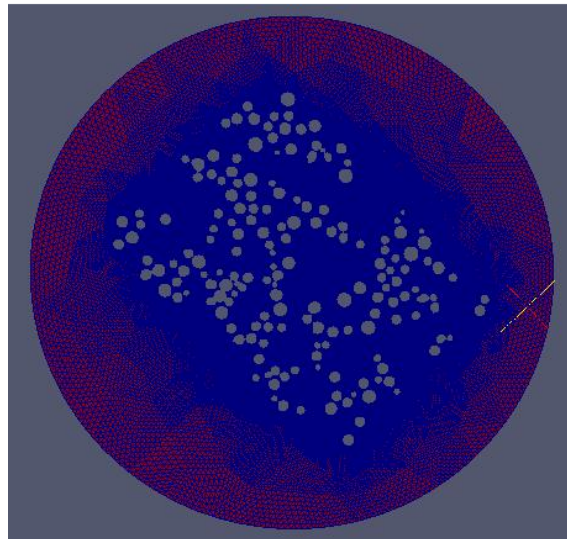


Abbildung 8.7: Ein Gitter mit 45.489 Elementen. Die Gitter-Erzeugung basiert auf einem Skript, welches in Zusammenarbeit mit Christian Diener erstellt wurde.

In den nachfolgenden Abschnitten haben wir folgende Modell-Parameter verwendet:

$$\begin{aligned}
 D_{\alpha} &= 324.03 \frac{\mu m^2}{s}, \\
 D_{\text{Bar1p}} &= 93.38 \frac{\mu m^2}{s}, \\
 k &= 28, \\
 C_{\alpha_i} &= 2.91 \frac{\text{molec}}{\mu m} \text{ für alle Mat}\alpha\text{-Zellen,} \\
 C_{\text{Bar1p}} &= 0.68 \frac{\text{molec}}{\mu m}, \\
 \tau_{\text{Bar1p}} &= 2.0s, \\
 \sigma &= 0.1 \frac{1}{\mu m}, \\
 A_{\text{max}} &= 10.0 \text{ molec}, \\
 \tau_{\text{rec}} &= 1.0s, \\
 K_M &= 1.0 \text{ molec}.
 \end{aligned}$$

Die Parameter wurden von Christian Diener zur Verfügung gestellt. Diesmal sind jedoch die Sekretionsraten von Bar1p und α -Faktor kleiner gewählt. In der ersten Simulation setzen wir die maximale Sekretionsrate von Bar1p auf Null. Das heißt, die MAT α -Zellen antworten nicht auf das Signal, was durch die MAT α Zellen in Form von α -Faktor ausgesendet wird. Wir sehen das Resultat nach ca. 20s in Abbildung 8.8. In dem roten Bereich ist die α -Faktor-Konzentration besonders stark und wir sehen, dass α -Faktor überall in dem Bereich der Zellen verteilt ist. Es werden in diesem Szenario alle MAT α -Zellen stimuliert, unabhängig davon, ob sich in unmittelbarer Nähe eine MAT α -Zelle befindet oder nicht. Desweiteren gibt es keine starken Gradienten, da sich die Konzentration von α -Faktor im Gebiet durch den Diffusionsprozess immer weiter angleicht. Es ist zu erwarten, dass die Konzentration von α -Faktor mit zunehmender Zeit homogen wird. Im zweiten Szenario setzen wir die maximale Sekretionsrate von Bar1p, also C_{Bar1p} , auf $0.68 \frac{\text{molec}}{\mu m}$. Das Ergebnis ist in Abbildung 8.9 zu sehen. Die MAT α -Zellen reagieren mit Sekretion von Bar1p. Dadurch wird α -Faktor degradiert. Das Signal konzentriert sich also wirklich nur auf die direkte Umgebung der MAT α -Zellen. Gradienten können auch nur dort wahrgenommen werden. Die Stimulation der MAT α -Zellen erfolgt also nicht mehr willkürlich.

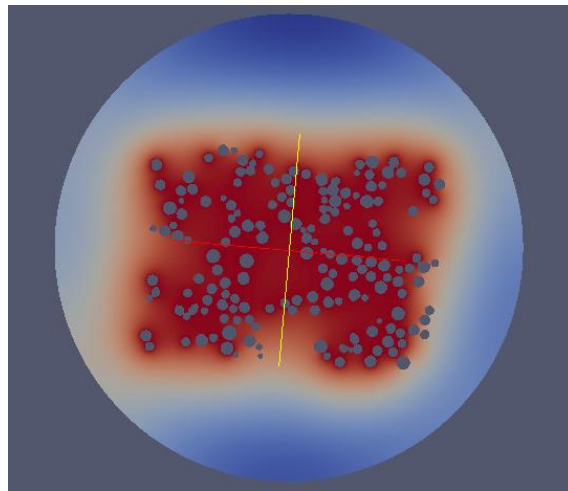


Abbildung 8.8: Verteilung von α -Faktor ohne Bar1p-Sekretion.

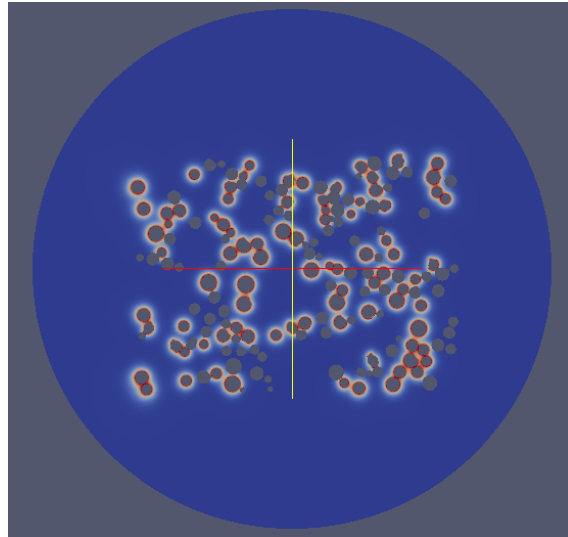


Abbildung 8.9: Verteilung von α -Faktor mit Einfluss von Bar1p.

9 Schlussbetrachtung und Ausblick

Diese Diplomarbeit ist interdisziplinär angelegt und verbindet die Fachbereiche Mathematik und Biologie. Die Arbeit erstreckt sich über funktionalanalytische Existenz- und Eindeutigkeitsaussagen, über ein numerisches Verfahren bis hin zu physikalischen Grundlagen und deren Anwendung auf ein biologisches Problem.

Die Ausarbeitungen in den Kapiteln 2 - 4 klären vor allem die theoretischen Grundlagen. Wir erfahren, welche Anforderungen an die Modelldaten gestellt werden müssen, damit ein numerisches Verfahren überhaupt konvergieren kann.

Es wurde das unstetige Galerkin-Verfahren untersucht. Dieses Verfahren ist durch seine Komplexität aufwändiger in der Implementierung als vergleichbare Finite-Elemente- oder Finite-Volumen-Methoden. Die Implementierung in DUNE hat sich als funktionsfähig und robust erwiesen. Die numerischen Tests sowie die Umsetzung in dem biologischen Modell haben gezeigt, dass dieses Verfahren vielseitig einsetzbar ist. Wir haben die Möglichkeit das Verfahren auf 2-dimensionalen, wie auf 3-dimensionalen unstrukturierten Gittern mit komplexen Geometrien zu nutzen. Die Möglichkeit der Parallelisierung beschleunigt das Verfahren erheblich. Eine Implementierung wurde auf dem Cluster erfolgreich getestet.

Das DG-Verfahren wurde in Bezug auf Konvergenz untersucht. Es wurden Konvergenzraten für die $\ell^2(H^1)$ - und $\ell^\infty(L^2)$ -Norm hergeleitet. Die Konvergenzraten in der $\ell^2(H^1)$ -Norm sind optimal [RW00]. Dabei wurden Resultate für das NIPG/SIPG/IIPG-Verfahren im Ort für zwei und drei Raumdimensionen und dem impliziten Euler- bzw. Crank-Nicolson-Verfahren in der Zeit hergeleitet. Es wurde der Ansatz von Feistauer [Fei09] als Ausgangspunkt gewählt, wobei hier der Konvektionsterm weggelassen wurde, da er in unserem Problem keinen Beitrag leistet. Dieser Ansatz wurde um die Diskretisierung in der Zeit für das implizite Euler- und das Crank-Nicolson-Verfahren ergänzt. Hierzu wurden Resultate aus [RW00] verwendet. Aufgrund der technischen Komplexität wurde in der vorgenannten Arbeit [RW00] der Fall von drei Raumdimensionen ausgespart, in dieser Arbeit jedoch ausgeführt.

Es wurde ein Modell für die Kommunikation von Hefezellen mit Hilfe von partiellen Differentialgleichungen erstellt. Mit numerischen Experimenten konnte der erwartete Einfluß von Bar1p dargestellt und diskutiert werden. Dies gelang mit dem Modell auch auf der Populationsebene mit mehr als hundert Zellen.

Der Quellcode für das biologische Modell, wie auch für die numerischen Experimente kann auf Nachfrage eingesehen werden.

Erweiterung des Modells und Ausblicke

In den vorangehenden Simulationen haben wir die Zellen durch kreisrunde Scheiben approximiert. Der Modellfehler kann jedoch weiter reduziert werden, indem wir natürliche

Geometrien für die Zellen verwenden. Durch die natürlichen Zellformen erhalten wir genauere Daten über Absorption bzw. Sekretion, wenn wir das Modell mit dem Experiment vergleichen. Desweiteren ist so eine bildhafte Identifikation und Wiedererkennung des Mikroskop-Bildes zum Aufzeigen verschiedener Effekte möglich. Außerdem können die Dirichlet-Randdaten durch Neumann-Bedingungen ersetzt werden, um experimentelle Daten wie explizite Sekretionsraten direkt umzusetzen.

In Abbildung 9.2 sehen wir ein Gitter mit natürlichen Geometrien. Das zugrunde liegende Mikroskopbild ist in Abbildung 9.1 zu sehen. Vorläufige numerische Experimente werden in den Abbildungen 9.3 und 9.4 gezeigt. Abgebildet sind 8 *MAT α* -Zellen und 55 *MATa*-Zellen.

Verschiedene Erweiterungen des Modells sind denkbar. Mit 3-dimensionalen Bilddaten kann man auch 3-dimensionale Gitter erzeugen, um dort die Modellgleichungen numerisch zu lösen. Im weiteren Verlauf könnte das PDE-System mit einem Modell der Prozesse im Zellinneren gekoppelt werden. Dies könnte dann durch ein System gewöhnlicher Differentialgleichungen modelliert werden oder auch durch Kopplung mit einzelnen Gittern für die Zellen. In zukünftigen Entwicklungen von DUNE ist z.B auch die Kopplung von verschiedenen Gittern vorgesehen, auf denen unterschiedliche Gleichungen betrachtet werden können.

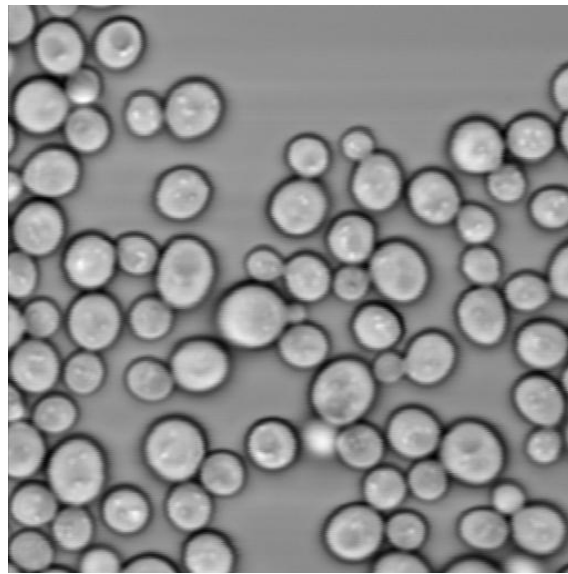


Abbildung 9.1: Mikroskopbild als Grundlage der Gittergenerierung. (Bild von Dr. Gabriele Schreiber)

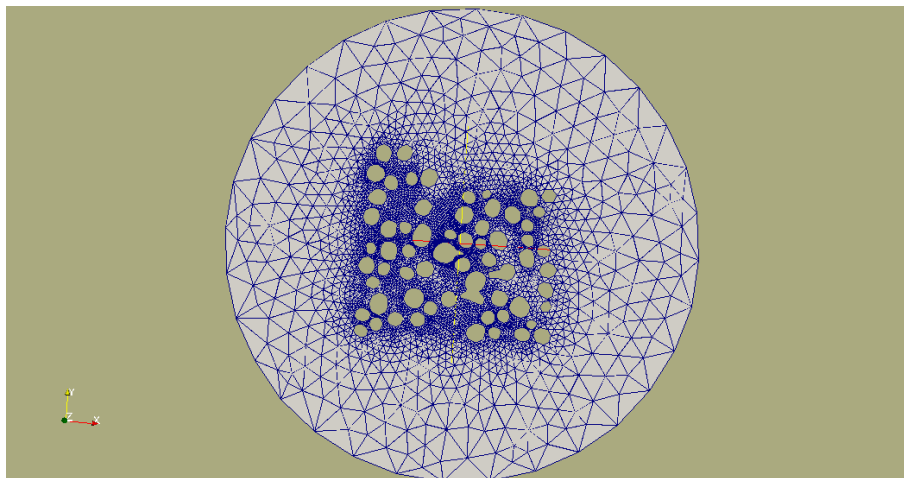


Abbildung 9.2: Gitter mit natürlichen Zellformen. Die Gittergenerierung auf Basis von Mikroskopdaten basiert auf auf einem Skript von Christian Diener.

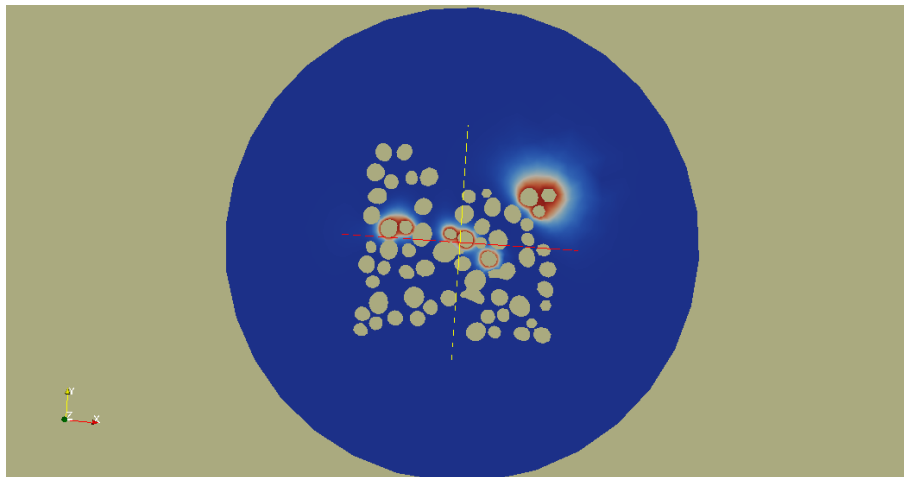


Abbildung 9.3: Verteilung von α -Faktor mit Einfluss von Bar1.

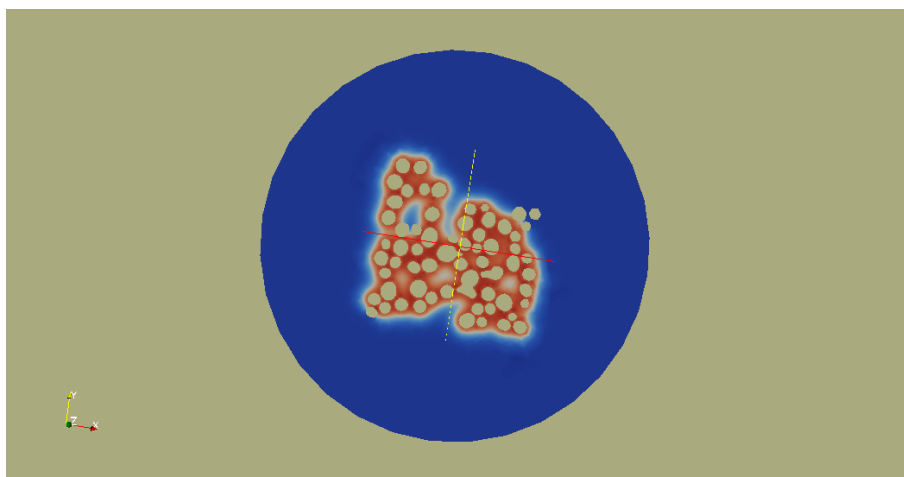


Abbildung 9.4: Verteilung von Bar1p.

Literaturverzeichnis

- [ABCM01] D. Arnold, F. Brezzi, B. Cockburn, and D. Marini. Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems. *SIAM Journal on Numerical Analysis*, 2001.
- [Alt06] Hans Wilhelm Alt. *Lineare Funktionalanalysis*. Springer-Verlag, 2006.
- [Ama95] Herbert Amann. *Gewöhnliche Differentialgleichungen*. de Gruyter Lehrbücher. Gruyter, 2. auflage edition, 1995.
- [Arn82] Douglas N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM Journal on Numerical Analysis*, 19(4):742–760, 1982.
- [Bar05] Bardwell. Mating. *Jounrnal of Computational Biology*, 2005.
- [Bas09] Peter Bastian. Parallele Löung großer linearer Gleichungssysteme, 2009. Vorlesungsskript.
- [BB08] Markus Blatt and Peter Bastian. On the generic parallelisation of iterative solvers for the finite element method. *International Journal of Computational Science and Engineering*, pages 56–69, 2008.
- [BBD⁺07] Peter Bastian, Markus Blatt, Andreas Dedner, Christian Engwer, Robert Klöfkorn, Ralf Kornhuber, and S Mario Ohlberger Oliver. A generic grid interface for parallel and adaptive scientific computing. part ii: Implementation and tests in dune, 2007.
- [BBD⁺08] P. Bastian, M. Blatt, A. Dedner, C. Engwer, R. Klöfkorn, M. Ohlberger, and O. Sander. A generic grid interface for parallel and adaptive scientific computing. part i: abstract framework. *Computing*, 82:103–119, July 2008.
- [BDE⁺05] Peter Bastian, Marc Droske, Christian Engwer, Robert Klöfkorn, Thimo Neubauer, Mario Ohlberger, and Martin Rumpf. Towards a unified framework for scientific computing. In *In Proc. of the 15th International Conference on Domain Decomposition Methods*, 2005.
- [BDT10] Peter Bastian and DUNE-Team. dune-pdelab howto. www.dune-project.org, October 2010.
- [BG99] Mary F. Wheeler Béatrice Rivière and Vivette Girault. Improved energy estimates for interior penalty, constrained and discontinuous galerkin methods for elliptical problems, 1999.

- [BJL⁺02] Peter Bastian, Klaus Johannsen, Stefan Lang, Christian Wieners, Volker Reichenberger, and Gabriel Wittum. Parallel numerical simulation of time-dependent problems on adaptive unstructured grids. 2002.
- [BMMP00] F. Brezzi, G. Manzini, D. Marini, and A. Pietra. Discontinuous galerkin approximations for elliptic problems. *Numerical Methods for Partial Differential Equations*, 16:365–378, 2000.
- [BO99] Carlos Erik Baumann and J. Tinsley Oden. A discontinuous hp finite element method for convection–diffusion problems. *Computer Methods in Applied Mechanics and Engineering*, 175(3-4):311–341, 1999.
- [BR97] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible navier-stokes equations. *Journal of Computational Physics*, 131(2):267–279, 1997.
- [Bra07] D. Braess. *Finite elemente*. Springer, 2007.
- [Cra56] J. Crank. *The Mathematics of Diffusion*. Oxford University Press, 1956.
- [CS98] Bernardo Cockburn and Chi-Wang Shu. The local discontinuous galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463, 1998.
- [CSQC08] Ching-Shan, Nie Qing, and Yi Tau-Mu Chou. Modeling robustness tradeoffs in yeast cell polarization induced by spatial gradients. *PLoS ONE*, 3(9):e3103, 09 2008.
- [DFKS08] V. Dolešji, M. Feistauer, V. Kučera, and V. Sobotíková. An optimal $L^\infty(L^2)$ -error estimate of the discontinuous galerkin method for a nonlinear nonstationary convection-diffusion problems. *IMA J. Numer. Anal.*, 3(28):496–521, 2008.
- [DSW04] C. Dawson, S. Sun, and M.F. Wheeler. Compatible algorithms for coupled flow and transport* 1. *Computer Methods in Applied Mechanics and Engineering*, 193(23-26):2565–2580, 2004.
- [Eva98] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.
- [Fei09] M. Feistauer. Workshop: DGFEM for a model scalar nonlinear nonstationary convection-diffusion problem. Presented at Humboldt-Universität zu Berlin, July 2009.
- [Gaj74] Groeger Zacharias Gajewski. *Nichtlineare Operatorgleichungen und Operator-differentialgleichungen*. Akademie Verlag, 1974.

- [GR08] Christophe Geuzaine and Jean-François Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 2008.
- [JD76] JR. J. Douglas and T. Dupont. Interior penalty procedures for elliptic and parabolic galerkin methods. *Lecture Notes in Phys.* 76, 1976.
- [KEP⁺00] D. Keyes, A. Ecer, J. Periaux, N. Satofuka, and P. Fox. Parallel implementation of the discontinuous galerkin method. In *Parallel computational fluid dynamics: towards teraflops, optimization, and novel formulations: proceedings of the Parallel CFD'99 Conference*, page 115. North-Holland, 2000.
- [Man83] T. R. Manney. Expression of the *bar1* gene in *saccharomyces cerevisiae*: induction by the alpha mating pheromone of an activity associated with a secreted protein. *J Bacteriol*, 155(1):291–301, Jul 1983.
- [MS98] K. Madden and M. Snyder. Cell polarity and morphogenesis in budding yeast. *Annu Rev Microbiol*, 52:687–744, 1998.
- [Mur02] J.D. Murray. *Mathematical biology I: An Introduction*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer, 3rd edition, 2002.
- [Mur04] J.D. Murray. *Mathematical biology II: Spatial Models and Biomedical Applications*, volume 18 of *Interdisciplinary Applied Mathematics*. Springer, 3rd edition, 2004.
- [Per10] Benoit Perthame. Growth, reaction, movement and diffusion from biology. www.ann.jussieu.fr/~perthame/cours_M2.pdf, November 2010.
- [QSS00] A. Quarteroni, R. Sacco, and F. Saleri. *Numerische Mathematik 2*. Springer-Verlag Berlin, 2000.
- [RH73] W. H. Reed and T. R. Hill. Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory, Los Alamos, MN, 1973.
- [Riv08] Béatrice Rivière. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations - Theory and Implementation*. SIAM, Frontiers in Applied Mathematics, 2008.
- [RW00] Béatrice Rivière and Mary F. Wheeler. A discontinuous galerkin method applied to nonlinear parabolic equations, 2000.
- [Tur52] A.M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37, 1952.

- [TW04] Andrea Toselli and Olof Widlund. *Domain Decomposition Methods - Algorithms and Theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer, 2004.
- [Wer09] D. Werner. *Funktionalanalysis*. Springer, 6. auflage edition, 2009.
- [Yan96] D. Yang. A parallel iterative nonoverlapping domain decomposition procedure for elliptic problems. *IMA Journal of Numerical Analysis*, 16(1):75, 1996.
- [You07] Yuncheng You. Global dynamics of the brusselator equations. *Dynamics of PDE*, 4(2):167–196, 2007.
- [Zei90] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications*, volume II/A: Linear Monotone Operators. Springer-Verlag, 1990.

10 Erklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Wolfgang Giese
Berlin, den 15. Dezember 2010

Einverständniserklärung

Hiermit erkläre ich mich einverstanden, dass ein Exemplar meiner Diplomarbeit in der Bibliothek des Institutes für Mathematik verbleibt.

Wolfgang Giese
Berlin, den 15. Dezember 2010