

**Ein mehrskaligen- und dualitätsbasiertes
Viskositätsmodell für die inkompressiblen
Navier-Stokes-Gleichungen**

DISSERTATION

zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt von

Arne Rempke

aus Berlin

genehmigt von der Fakultät
für Mathematik/Informatik und Maschinenbau
der Technischen Universität Clausthal

Tag der mündlichen Prüfung:

25. Juni 2015

Vorsitzender der Promotionskommission:

Hauptberichterstatter:

Mitberichterstatter:

Prof. Dr. Michael Kolonko, TU Clausthal

Prof. Dr. Lutz Angermann, TU Clausthal

Prof. Dr. Gert Lube, Universität Göttingen

Abstract

Eine wichtige Anwendung in den Ingenieurwissenschaften und in der Physik ist die Berechnung der Strömung von Flüssigkeiten oder Gasen. Das Verhalten solcher Fluide wird häufig durch die *Navier-Stokes-Gleichungen* beschrieben. Viele dieser Probleme lassen sich nicht analytisch lösen, sondern nur numerisch approximieren. Insbesondere bei turbulenten Strömungen besteht die Schwierigkeit, dass die Diskretisierung nicht alle Skalen auflösen kann. Kleine Verwirbelungen können somit nicht berechnet werden. Üblicherweise werden ihre Effekte jedoch mithilfe von Subgrid-Modellen nachgebildet. Diese Techniken sind auch als *Large Eddy Simulation* oder *variationelle Multiskalenmethode* bekannt. Viele der Subgrid-Modelle gehen als zusätzliche Diffusion in Form einer künstlichen Viskosität in die Differentialgleichungen ein, es wird von *Reynolds-Averaged-Navier-Stokes-Gleichungen* gesprochen.

Moderne Verfahren zur numerischen Lösung von Differentialgleichungen verändern die der Diskretisierung zugrunde liegenden Gitter in Abhängigkeit von der zu bestimmenden Lösung. Als Indikator für eine adaptive Gitterverfeinerung kann ein Funktional der Lösung verwendet werden. Dieses Funktional wird im Vorfeld vom Anwender vorgegeben und beschreibt eine Größe, die ihm bei der Approximation besonders wichtig ist. Mithilfe eines dualen Problems können dann die Gitterelemente identifiziert werden, bei denen eine lokale Änderung der Lösung einen besonders großen Einfluss auf den Wert des Funktionals hat. Bei der verbreiteten Technik der *Dual Weighted Residuals* werden gezielt diese Gitterelemente verfeinert.

In der vorliegenden Arbeit wird auf Grundlage einer a-posteriori-Fehlerabschätzung ein inverses Problem entwickelt, das die künstliche Viskosität nicht als Parameter, sondern als Variable enthält. Statt aus einer vorgegebenen künstlichen Viskosität die resultierende Strömung zu berechnen, wird diejenige künstliche Viskosität gesucht, die eine Strömung liefert, deren Funktionalwert möglichst exakt ist. Es wird ein mehrskaliger Ansatz verwendet, um den Funktionalwert einer feineren Diskretisierung als Referenz nutzen zu können.

Auf Grundlage dieser neuen Idee kann für beliebige (partielle) Differentialgleichungen mitsamt eines vorgegebenen Modellierungsansatzes ein solches inverses Problem formuliert werden. Die Lösung des inversen Problems steuert dabei den Modellierungsansatz in Raum und Zeit. Anschließend wird das resultierende System aus fünf Variationsgleichungen im konkreten Kontext der instationären inkompressiblen Navier-Stokes-Gleichungen formuliert, auf Basis der Finite-Elemente-Methode implementiert und anhand von drei Modellproblemen praktisch gelöst.

Schließlich wird noch der Frage nachgegangen, ob und wie aus der so gewonnenen konkreten künstlichen Viskosität wieder ein (allgemeiner verwendbares) Viskositätsmodell konstruiert werden kann. Dazu werden Approximationsansätze auf Grundlage von reduzierten Basen, nichtlinearer Approximation sowie radialen Basisfunktionen diskutiert, implementiert und auf die vorgestellten Modellprobleme angewendet. Diese neuen Viskositätsmodelle werden mit den bereits gewonnenen Daten verglichen und eingeordnet.

Inhaltsverzeichnis

1	Einleitung und Motivation	1
2	Mathematische Formulierung der zu betrachtenden grundlegenden Probleme	4
2.1	Grundlegende Notation	4
2.2	Differentialgleichungen	5
3	Variationelle Formulierung	8
3.1	Funktionsräume und Abbildungen	8
3.2	Anwendung auf die Navier-Stokes-Gleichung	9
3.3	Symmetrisierung	12
3.4	Physikalische Randbedingungen	13
3.4.1	Undurchlässiger Rand	16
3.4.2	Durchlässiger Rand	17
3.4.3	Mathematische Formulierung	21
3.5	Existenz und Eindeutigkeit	23
4	Diskretisierung	25
4.1	Diskrete Funktionsräume	25
4.2	Gitterbasierte Zerlegung	27
4.3	Diskrete Abbildungen	28
4.4	Zeitdiskretisierung	32
4.5	Diskretisierung des gemischten Problems	34
4.6	Turbulenz und Multiskalen	35
5	A-posteriori-Abschätzung der Navier-Stokes-Gleichung sowie des dualen Problems zur Modellgewinnung	37
5.1	ε -schwache Lösungen	38
5.2	Das duale Problem	38
5.3	Modellierungskorrektur durch a-posteriori-Abschätzung	39
5.4	Skalen- und Modellierungsabschätzung für die Navier-Stokes-Gleichungen	51
5.5	Modellierungskorrektur für die Navier-Stokes-Gleichungen	53
6	Mehrskalige Rechnung zur Parameterbestimmung	55
6.1	Duales instationäres Problem	57
6.2	Modellprobleme	59
6.2.1	Modellproblem A	59
6.2.2	Modellproblem B	61
6.2.3	Modellproblem C	65

7	Modellierung der Stabilisierung als Viskositätsmodell	70
7.1	Aufwandsbetrachtung	72
7.2	Modelltrainingsdaten	73
7.3	Räumlich globale bzw. lokale Regression	73
7.4	Modellierung mittels reduzierter Basisfunktionen	75
7.4.1	Grundprinzip der reduzierten Basis	75
7.4.2	Abbildung mittels reduzierter Basis	78
7.4.3	Anwendung auf die Modelltrainingsdaten	88
7.4.4	Zusammenfassung	94
7.5	Modellierung mittels vorgegebener Ansatzfunktionen	94
7.5.1	Verwendung der Abbildungen als Viskositätsmodell	112
7.6	Modellierung mittels radialer Basisfunktionen	117
7.6.1	Verwendung der Abbildungen als Viskositätsmodell	127
7.7	Zusammenfassung	132
8	Zusammenfassung und Ausblick	137
8.1	Konkrete in der Arbeit aufgetretene Beschränkungen	138
8.2	Mögliche Weiterführungen der grundlegenden Idee	138

Abbildungsverzeichnis

3.1	Ausfluss des Fluids aus dem Simulationsgebiet (a) mit bzw. (b) ohne dem Randintegral $\int_{\partial\Omega} n \cdot \nabla u \cdot v$ der Diffusion	14
3.2	Ausfluss des Fluids aus dem Simulationsgebiet (a) ohne bzw. (b) mit dem Randintegral $\frac{1}{2}d(u, u \cdot v)$ der Konvektion	15
3.3	Exemplarisches Simulationsgebiet: Rohrströmung mit Hindernis in physikalische Ränder unterteilt	16
3.4	Ausfluss des Fluids aus dem Simulationsgebiet an einem Neumann-Rand: (a) do-nothing bzw. (b) vorgegebene Rohrspannung	19
3.5	Ausfluss des Fluids aus einem verwirbelten Simulationsgebiet: Randbedingung (a) do-nothing bzw. (b) künstlich	20
3.6	Randmodellierung für das exemplarische Simulationsgebiet	21
6.1	Geometrie und Gitter des Modellproblems A	59
6.2	Entwicklung (a) des Funktionals sowie (b) der Parameter des inversen Problems bei Modellproblem A	60
6.3	Geometrie und Gitter des Modellproblems B	61
6.4	Entwicklung (a) des Funktionals sowie (b) der Parameter des inversen Problems bei Modellproblem B	62
6.5	Geschwindigkeitsquerschnitte bei Modellproblem B über den Zeitraum (a) $t \in (2.25, 2.75)$ sowie (b) $t \in (2.75, 5)$	64
6.6	Vergleichsströmungsprofile aus [DAG ⁺ 11]	65
6.7	Geometrie und Gitter des Modellproblems C	65
6.8	Entwicklung (a) des Funktionals sowie (b) der Parameter des inversen Problems bei Modellproblem C	67
6.9	Geschwindigkeitsquerschnitte bei Modellproblem C über den Zeitraum (a) $t \in (3, 6)$ sowie (b) $t \in (8, 10)$	68
7.1	Schematische Darstellung der Grundidee zur Gewinnung eines Viskositätsmodells	71
7.2	Approximationsfehler mit klassisch reduzierter Basis bei synthetischen Daten	78
7.3	Approximationsfehler mit reduzierter Basis als Abbildung bei synthetischen Daten	79
7.4	Approximationsfehler mit orthonormalisierten Funktionalen bei synthetischen Daten	83
7.5	Approximationsfehler mit näherungsweise Singulärwertzerlegung bei synthetischen Daten	84
7.6	Approximationsfehler mit reduzierter Singulärwertzerlegung bei synthetischen Daten	87
7.7	Vergleich der Approximationsfehler mit den verschiedenen Verfahren bei synthetischen Daten	87

7.8	Vergleich der Approximationsfehler bei (a) globaler bzw. (b) lokaler Regression der Modelltrainingsdaten A mit den verschiedenen Verfahren	89
7.9	Vergleich der Approximationsfehler bei (a) globaler bzw. (b) lokaler Regression der Modelltrainingsdaten B mit den verschiedenen Verfahren	91
7.10	Vergleich der Approximationsfehler bei (a) globaler bzw. (b) lokaler Regression der Modelltrainingsdaten C mit den verschiedenen Verfahren	93
7.11	Approximation 1 der Modelltrainingsdaten A	96
7.12	Approximation 1 der Modelltrainingsdaten B	97
7.13	Approximation 1 der Modelltrainingsdaten C	98
7.14	Approximation 2 der Modelltrainingsdaten A	100
7.15	Approximation 2 der Modelltrainingsdaten B	101
7.16	Approximation 2 der Modelltrainingsdaten C	102
7.17	Approximation 3 der Modelltrainingsdaten A	103
7.18	Approximation 3 der Modelltrainingsdaten B	104
7.19	Approximation 3 der Modelltrainingsdaten C	105
7.20	Approximation 4 (Smagorinsky) der Modelltrainingsdaten A	106
7.21	Approximation 4 (Smagorinsky) der Modelltrainingsdaten B	107
7.22	Approximation 4 (Smagorinsky) der Modelltrainingsdaten C	108
7.23	Vergleich der Approximationsfehler bei den verschiedenen Approximationen .	111
7.24	Entwicklung des Funktionals über die Zeit bei Modellproblem A und lokaler Regression nach Approximationen 2, 3 sowie Smagorinsky-Modell	113
7.25	Entwicklung des Funktionals über die Zeit bei Modellproblem B und lokaler Regression nach Approximationen 2, 3 sowie Smagorinsky-Modell	113
7.26	Entwicklung des Funktionals über die Zeit bei Modellproblem C und lokaler Regression nach Approximationen 2, 3 sowie Smagorinsky-Modell	114
7.27	Geschwindigkeitsquerschnitte bei Modellproblem B mit den lokalen Approximationen über den Zeitraum (a) $t \in (2.25, 2.75)$ sowie (b) $t \in (2.75, 5)$	115
7.28	Geschwindigkeitsquerschnitte bei Modellproblem C mit den lokalen Approximationen über den Zeitraum (a) $t \in (3, 6)$ sowie (b) $t \in (8, 10)$	116
7.29	Eindimensionale Beispieldaten (mit Rauschen) zur Approximation mit radialen Basisfunktionen mit $M \in \{1, 4, 10\}$	118
7.30	Approximation durch radiale Basisfunktionen der Modelltrainingsdaten A: (a) Koeffizienten der Ansatzfunktionen und (b) Wahl der Ansatzfunktionen . . .	123
7.31	Approximation durch radiale Basisfunktionen der Modelltrainingsdaten B: (a) Koeffizienten der Ansatzfunktionen und (b) Wahl der Ansatzfunktionen . . .	124
7.32	Approximation durch radiale Basisfunktionen der Modelltrainingsdaten C: (a) Koeffizienten der Ansatzfunktionen und (b) Wahl der Ansatzfunktionen . . .	125
7.33	Fehler bei Approximation der Modelltrainingsdaten A, B und C mit radialen Basisfunktionen	126
7.34	Entwicklung des Funktionals über die Zeit bei Modellproblem A und lokaler Regression mit 2, 10 bzw. 50 radialen Basisfunktionen	128
7.35	Entwicklung des Funktionals über die Zeit bei Modellproblem B und lokaler Regression mit 2, 10 bzw. 50 radialen Basisfunktionen	128
7.36	Entwicklung des Funktionals über die Zeit bei Modellproblem C und lokaler Regression mit 2, 10 bzw. 50 radialen Basisfunktionen	129
7.37	Geschwindigkeitsquerschnitte bei Modellproblem B mit den lokalen RBF-Ansätzen über den Zeitraum (a) $t \in (2.5, 3)$ sowie (b) $t \in (3.25, 3.5)$	130

7.38	Geschwindigkeitsquerschnitte bei Modellproblem C mit den lokalen RBF-Ansätzen über den Zeitraum (a) $t \in (3, 6)$ sowie (b) $t \in (8, 10)$	131
7.39	Zeitliche Mittelung des Funktionals aus den verschiedenen Approximationen bei Modellproblem A (a) über $I = (0, 20)$ und (b) über $I = (2, 20)$	133
7.40	Zeitliche Mittelung des Funktionals aus den verschiedenen Approximationen bei Modellproblem B (a) über $I = (0, 5)$ und (b) über $I = (1.5, 5)$	134
7.41	Zeitliche Mittelung des Funktionals aus den verschiedenen Approximationen bei Modellproblem C (a) über $I = (0, 10)$ und (b) über $I = (8, 10)$	135

Tabellenverzeichnis

6.1	Parameter der Modellprobleme	69
7.1	Koeffizienten sowie deren Genauigkeit bei Approximation 1 der Modelltrainingsdaten A und B	109
7.2	Koeffizienten sowie deren Genauigkeit bei Approximation 1 der Modelltrainingsdaten C	110
7.3	Koeffizienten sowie deren Genauigkeit bei Approximation 2	110
7.4	Koeffizienten sowie deren Genauigkeit bei Approximation 3	110
7.5	Koeffizienten sowie deren Genauigkeit bei Approximation 4 (Smagorinsky) . .	110
7.6	Vergleich der Approximationsfehler bei den verschiedenen Approximationen .	111
7.7	Daten der radialen Basisfunktionen zur Approximation der Modelltrainingsdaten A	120
7.8	Daten der radialen Basisfunktionen zur Approximation der Modelltrainingsdaten B	121
7.9	Daten der radialen Basisfunktionen zur Approximation der Modelltrainingsdaten C	122
7.10	Fehler bei Approximation der Modelltrainingsdaten A, B und C mit radialen Basisfunktionen	126

Algorithmenverzeichnis

4.1	Ein Schritt einer allgemeinen s -stufigen Einschrittmethode	33
6.1	Mehrskalige Parameterbestimmung	55
6.2	Regularisierung	58
7.1	Klassische Reduktion der Basis von R	77
7.2	Darstellung von r in Abhängigkeit von d mit reduzierter Basis	80
7.3	Darstellung von r in Abhängigkeit von d mit reduzierter Basis durch Orthonormalisierung des kanonischen Duals	82
7.4	Darstellung von r in Abhängigkeit von d mit reduzierter Basis durch näherungsweise Singulärwertbestimmung	85
7.5	Approximation mittels radialer Basisfunktionen	119

Danksagung

Ein Werk wie eine Dissertation entsteht nicht über Nacht und auch nicht von alleine. Daher soll dieser Platz zunächst den Unterstützern im Hintergrund gewidmet sein.

Die Promotion ist als Teil eines Ausbildungsprozesses nicht selbstverständlich. Als stetige Begleiter dieses langjährigen Weges, der schon viel eher als das Verfassen dieser Arbeit begonnen hat, möchte ich mich daher zunächst bei meiner engeren Familie bedanken:

- Meiner Mutter Karin für die langjährige Unterstützung und den Rückhalt, den ich jederzeit erhalten konnte.
- Meinem Vater Thomas für unzählige spannende Diskussionen und so manch kritische Frage.
- Meiner Tante Bärbel für das Näherbringen der Mathematik sowie die Verbindung nach Clausthal.

Die vorliegende Arbeit ist während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Mathematik der Technischen Universität Clausthal entstanden. Im Rückblick war dies eine sehr schöne Zeit, die ich nicht missen möchte. Grund dafür war zum einen das angenehme Arbeitsklima sowie die kollegiale Zusammenarbeit am Institut, zum anderen ein breitgefächertes soziales Netzwerk in ganz Clausthal.

An erster Stelle danke ich natürlich meinem Betreuer Prof. Lutz Angermann für die Anleitung, den regen Austausch, das entgegengebrachte Vertrauen und die sehr gute Zusammenarbeit. Ich danke Prof. Gert Lube, der, als ich ihm mein Thema vorstellte, nicht nur unmittelbar sehr angetan war, sondern auch in schwierigen Zeiten das Zweitgutachten übernommen hat. Dank geht insbesondere an Prof. Ulrich Mertins, Prof. Hans-Heinrich Kairies, Prof. Walter Klotz, Bernd Mulansky und Henning Behnke, in deren Vorlesungen und Übungen so manche Grundlage gelegt wurde. Auch allen anderen Mitarbeitern im Institut danke ich für die konstruktive Zusammenarbeit. Ich danke meinem Freund Sven für etliche spannende Koch- und Diskussionsabende. Der Mensa-Runde danke ich für regen mittäglichen Austausch – fachlich wie interdisziplinär – oder auch mal Ablenkung: Alexander, Anja, Annika, Christian, Daniel, Dennis, Dorothee, Fabian, Hendrik, Nora, Sven, Wiebke und Wiebke. Vielen Dank auch an die Mitglieder der Gruppierungen MundHarmoniker, Sinfonieorchester, Highflyers sowie Tanzbären für viel Spaß, eine angenehme Atmosphäre und die Möglichkeit zur abwechslungsreichen Freizeitgestaltung.

Schließlich möchte ich der Person danken, die vermutlich das Meiste in der stressigen Schlussphase aushalten musste, auf die ich mich aber dennoch jederzeit verlassen konnte: Meiner Freundin Nicole.

1 Einleitung und Motivation

Ein zentrales – wenn nicht sogar das zentrale – Anliegen der Naturwissenschaften ist es, die Welt, in der wir leben, zu verstehen und zu beschreiben. Ein wichtiges Werkzeug aller Naturwissenschaften ist die Mathematik. Sie dient unter anderem dazu, naturwissenschaftliches, konkret etwa physikalisches Verhalten, präzise zu beschreiben. Anhand dieser Beschreibungen können theoretische Betrachtungen, Analysen und Berechnungen durchgeführt werden, die zu Vorhersagen führen. Häufig lässt sich auf diesem Weg der Ausgang eines Experiments vorhersagen, das in der Regel auch in der realen Welt durchführbar ist. Idealerweise stimmt die Vorhersage mit dem Experiment überein und wir haben eine *Simulation*, die das Experiment gut beschreibt.

Es gibt eine Reihe von Vorteilen bei Simulationen:

- Statt einen Versuch immer und immer wieder mit leicht veränderten Bedingungen durchzuführen, kann dies bei Simulationen durch Anpassung einer einzigen oder weniger Variablen geschehen. Dies führt möglicherweise zu einer deutlichen Reduktion des Aufwands (z. B. bei der Entwicklung und Optimierung einer Komponente).
- Manche Versuche sind nicht umkehrbar, z.B. werden bei manchen Experimenten die verwendeten Stoffe zerstört. Die Wiederherstellung kann teuer sein, in bestimmten Fällen ist sie sogar unmöglich. Dann ist eine Durchführung des Experiments nicht empfehlenswert (z. B. Festigkeitsversuch bis zum Zerreißen).
- In den meisten Simulationen werden mehr Informationen als bei den entsprechenden Versuchen gewonnen. Dies kann etwa ein Blick ins Innere eines Körpers sein oder eine detailliertere Beschreibung bei zeitlich kritischen oder räumlich filigranen Prozessen. Spezielle Informationen aus Experimenten zu ziehen kann zwar gelingen, aber dies kann auch den Versuch beeinflussen, etwa die Platzierung von Sensoren (z. B. Wärmeentwicklung im Inneren eines Werkstücks, Verhalten auf subatomarer Größenordnung).
- Manche Versuche sind unmöglich durchzuführen, da es sich eher um Gedankenexperimente oder sehr weitreichende Fragestellungen dreht, (z. B. Klimasimulationen, astronomische Fragen, militärische Simulationen oder volkswirtschaftliche Auswirkungen).

Auch wenn die Simulation in vielen Fällen ein hilfreiches Werkzeug darstellt, muss sichergestellt sein, dass sie richtig angewendet wird. Schon anhand des eingangs skizzierten Ablaufes einer Simulation ist es ersichtlich, dass viele Einzelschritt nötig sind, in denen jeweils Fehler entstehen können. Auch muss festgehalten werden, dass bei weitem nicht jede Simulation möglich ist. In den vergangenen Jahrzehnten wurden zwar bereits fundamentale Fortschritte in Bezug auf die behandelbare Problemgröße gemacht, dennoch sind manche Fragestellungen noch völlig unbeantwortet. Einen Wasserfall mit all seinen Verwirbelungen, dem Aufschäumen, einer Bläschen- und Tröpfchenbildung sowie der Berücksichtigung der mitgerissenen Luft zu simulieren ist mit derzeitigen Mitteln nicht ansatzweise simulierbar.

Die Simulation eines physikalischen Phänomens ist interdisziplinär: Angefangen bei Experimen-

ten (Naturwissenschaft, Technik) die zur Beschreibung und Verifikation der Gesetzmäßigkeiten benötigt werden, über die Formulierung des Problems (Mathematik), bis hin zum Berechnen (Informatik), sind bereits viele große Disziplinen eingebunden. Auf allen Bereichen gibt es regelmäßig Fortschritte und Verbesserungen zu verzeichnen: Z. B. steigende Rechenleistungen der Computer oder ein verbesserter Algorithmus. Die Tendenz, dass sich hier auch in Zukunft noch einiges entwickelt und die Grenzen des Machbaren weitergeschoben werden, ist deutlich.

Auch die vorliegende Arbeit ist ein Beitrag, diese Grenze zu erweitern. Statt jedoch die Ergebnisse von Simulationen quantitativ zu verbessern, also mehr Daten zu erzeugen, etwa durch den Einsatz von Hochleistungsrechnern, liegt der Fokus auf einer qualitativen Verbesserung: Die zur Verfügung stehenden Ressourcen zur Simulation sollten besser dafür verwendet werden, die Teile zu berechnen, die bei dieser Simulation interessant sind.

Oben wurde bereits angeführt, dass bei typischen Simulationen mehr Informationen als bei einem Experiment anfallen. Der Grundgedanke der vorliegenden Arbeit ist es nun, sich bei der Simulation nur auf bestimmte, im Vorhinein festzulegende Eigenschaften zu konzentrieren. Es soll also bereits vor Beginn der Simulation festgelegt werden, welche Ergebnisse der Simulation *wichtig* sind. Andere Größen, die zwar ebenfalls anfallen, aber eben nicht so wichtig sind, können wir entsprechend als *unwichtig* bezeichnen. Um zu erreichen, dass die wichtigen Ergebnisse eine möglichst hohe Präzision erreichen, kann mehr Rechenleistung auf ihre Berechnung vereint werden. Im Gegenzug kann zur Bestimmung der unwichtigen Ergebnisse Rechenkapazität der Simulation eingespart werden. Ziel ist es, auf diesem Wege mit der zur Verfügung stehenden Rechenkapazität eine höhere Güte der wichtigen Simulationsergebnisse zu erzielen als auf herkömmlichen Wege.

Eine im Vorhinein durch den Anwender vorzunehmende Klassifizierung der Ergebnisse als wichtig oder unwichtig stellt in der Regel kein Problem dar. Zu einem Abgleich mit experimentellen Daten bietet es sich natürlich an, die in den Experimenten gemessenen Größen als wichtig zu deklarieren. Bei der Entwicklung und Optimierung einer Komponente sind in der Regel auch bestimmte Kenngrößen von besonderer Relevanz der Simulation, z. B. die Festigkeit, Verformung oder Temperatur.

Wesentlich schwieriger gestaltet sich die konkrete Umsetzung der Idee. Die unwichtigen Ergebnisse können nicht einfach weggelassen werden, denn in der Regel hängen sie (zum Teil sehr stark) mit den wichtigen Ergebnissen zusammen. Insbesondere bei nichtlinearen Problemen kann dieser Effekt sehr stark ausgeprägt sein. Stattdessen kann versucht werden, die unwichtigen Ergebnisse weniger genau zu bestimmen, indem Approximationen und Modellierungen verwendet werden.

Die oben skizzierte Idee wird in der vorliegenden Arbeit umgesetzt und untersucht. Als konkrete Simulationsanwendung dient dazu eine inkompressible Strömung. Ihre mathematische Beschreibung wird in Kapitel 2 dargelegt. In Kapitel 3 wird dieses Problem in einer Weise umgeformt, die gewissermaßen universell für eine Reihe von ähnlichen Problemen und gleichzeitig Grundlage der Lösungsstrategie ist. Das Kapitel 4 stellt dann vor, wie solche Probleme auf herkömmlichem Wege numerisch gelöst werden können.

Kapitel 5 stellt anschließend die zentrale Idee der Arbeit vor: Die Klasse der in Kapitel 3 vorgestellten Probleme wird mithilfe eines dynamischen Modells so approximiert, dass der Fehler bezüglich einer zuvor festgelegten Eigenschaft klein ist. Zentrale Fragestellung ist dabei die Entwicklung des dynamischen Modells bzw. der Gewinnung seiner Parameter.

In Kapitel 6 wird ein Verfahren zur Lösung der in Kapitel 5 entwickelten Probleme konkret auf die inkompressible Strömung angewendet. Es werden drei Modellprobleme vorgestellt und die numerischen Ergebnisse dargelegt.

Da die Referenz des in Kapitel 5 beschriebenen Verfahrens aus einer aufwendigeren Simulation bestand, und daher der Aufwand des bislang beschriebenen Verfahrens eher größer geworden ist, werden in Kapitel 7 verschiedene Ansätze besprochen, wie hier approximiert werden kann. Auch diese Ansätze werden jeweils auf die zuvor eingeführten Modellprobleme angewendet und numerisch ausgewertet.

Kapitel 8 schließlich fasst die Ergebnisse zusammen und stellt Ideen für Folgeuntersuchungen oder weitergehende Anwendungen vor.

2 Mathematische Formulierung der zu betrachtenden grundlegenden Probleme

In diesem Kapitel werden die mathematischen Grundlagen zur Beschreibung einer inkompressiblen Strömung dargestellt. Insbesondere wird zunächst die in dieser Arbeit verwendete Notation vorgestellt. Im zweiten Teil werden verschiedene Möglichkeiten zur Beschreibung von inkompressiblen Strömungen mithilfe von Differentialgleichungen aufgeführt.

2.1 Grundlegende Notation

Es sei $d \in \{2, 3\}$ die räumliche Dimension und $\Omega \subset \mathbb{R}^d$ ein beschränktes Gebiet, auf dem das Strömungsverhalten eines Fluids bestimmt werden soll. Die gesuchte Strömung wird im stationären Fall dargestellt durch

- das Vektorfeld der Geschwindigkeit

$$u : \Omega \rightarrow \mathbb{R}^d,$$

- sowie den Druck

$$p : \Omega \rightarrow \mathbb{R}.$$

Bekannt seien

- das Vektorfeld des Einflusses äußerer Kräfte

$$f : \Omega \rightarrow \mathbb{R}^d,$$

- sowie die (kinematische) Viskosität

$$\nu : \Omega \rightarrow \mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\}.$$

Häufig ist die Viskosität eine vom Fluid abhängige Konstante, d. h. mit $\nu_{\text{const}} > 0$ ist

$$\nu(x) = \nu_{\text{const}} \quad \forall x \in \Omega.$$

Das physikalische Verhalten kann durch Differentialgleichungen beschrieben werden, die wir mit den folgenden Differentialoperatoren formulieren:

- Nabla-Operator ∇ :

$$\nabla p := \begin{pmatrix} \partial_1 p \\ \vdots \\ \partial_d p \end{pmatrix} := \begin{pmatrix} \frac{\partial}{\partial x_1} p \\ \vdots \\ \frac{\partial}{\partial x_d} p \end{pmatrix}$$

- Divergenz-Operator $\nabla \cdot$:

$$\nabla \cdot u := \sum_{k=1}^d \partial_k u_k := \sum_{k=1}^d \frac{\partial}{\partial x_k} u_k$$

- Laplace-Operator Δ :

$$\Delta u := \sum_{k=1}^d \partial_k^2 u_k := \sum_{k=1}^d \frac{\partial^2}{\partial x_k^2} u_k$$

2.2 Differentialgleichungen

Sofern die Daten der Aufgabe hinreichend glatt und gutartig gewählt sind, stellt sich eine stationäre Strömung ein, die durch eines der folgenden Probleme beschrieben werden kann:

1. Stationäres inkompressibles Stokes-Problem

$$\begin{aligned} -\Delta u + \nabla p &= f \\ \nabla \cdot u &= 0. \end{aligned} \tag{2.1}$$

2. Stationäres inkompressibles Navier-Stokes-Problem

$$\begin{aligned} -\nu \Delta u + \nabla p + u \cdot \nabla u &= f \\ \nabla \cdot u &= 0. \end{aligned} \tag{2.2}$$

Während das Navier-Stokes-Problem die Strömung physikalisch korrekter beschreibt, ist das Stokes-Problem mathematisch einfacher zu behandeln. Der wesentliche Unterschied ist die nichtlineare Konvektion $u \cdot \nabla u$. In speziellen Anwendungsfällen mag sie vernachlässigbar sein, in der Praxis ist jedoch (2.2) meist von größerer Relevanz.

Ist eine Beschreibung als stationäres Problem nicht möglich, dann verändert sich die Strömung mit der Zeit. Die Beschreibungsgrößen müssen also zeitabhängig erfasst werden. Konkret betrachten wir das Zeitintervall $I := (0, T)$ mit $T > 0$. Dann ist

$$\begin{aligned} u : I \times \Omega &\rightarrow \mathbb{R}^d, & f : I \times \Omega &\rightarrow \mathbb{R}^d, \\ p : I \times \Omega &\rightarrow \mathbb{R}, & \nu : I \times \Omega &\rightarrow \mathbb{R}^+. \end{aligned}$$

Die Nabla-, Divergenz- und Laplace-Operatoren wirken weiterhin – wie oben definiert – nur auf den räumlichen Abhängigkeiten. Zur Beschreibung der zeitlichen Veränderung verwenden

wir ∂_t als Notation der partiellen Ableitung nach t :

$$\partial_t u := \frac{\partial}{\partial t} u := \begin{pmatrix} \frac{\partial}{\partial t} u_1 \\ \vdots \\ \frac{\partial}{\partial t} u_d \end{pmatrix}.$$

Nun können wir die zeitabhängigen Probleme formulieren:

3. Instationäres inkompressibles Stokes-Problem

$$\begin{aligned} \partial_t u - \Delta u + \nabla p &= f \\ \nabla \cdot u &= 0. \end{aligned} \tag{2.3}$$

4. Instationäres inkompressibles Navier-Stokes-Problem

$$\begin{aligned} \partial_t u - \nu \Delta u + \nabla p + u \cdot \nabla u &= f \\ \nabla \cdot u &= 0. \end{aligned} \tag{2.4}$$

Damit es sich um sinnvoll gestellte Probleme handelt, müssen hier jeweils noch Anfangs- bzw. Randbedingungen ergänzt werden. Im stationären Fall genügen Randbedingungen, im instationären Fall sind zusätzlich auch Anfangsbedingungen nötig.

Wir wollen mit $\partial\Omega$ den Rand von Ω , sowie mit

$$\bar{\Omega} := \Omega \cup \partial\Omega$$

den Abschluss von Ω bezeichnen.

Die Funktionen u bzw. p , ggf. auch ∇u , sollen nun stetig auf den Rand $\partial\Omega$ fortsetzbar sein:

$$u \in (C(\bar{\Omega}))^d \quad \text{und} \quad p \in C(\bar{\Omega})$$

bei den stationären bzw.

$$u \in (C(\bar{I} \times \bar{\Omega}))^d \quad \text{und} \quad p \in C(\bar{I} \times \bar{\Omega})$$

bei den instationären Problemen. Typische Anfangs- bzw. Randbedingungen sind etwa folgende (hier nur für die instationären Probleme formuliert):

- Anfangsbedingung durch eine vorgegebene Anfangsströmung $u_0 : \Omega \rightarrow \mathbb{R}^d$:

$$u(0, x) = u_0(x) \quad \forall x \in \Omega.$$

- Dirichlet-Randbedingungen auf $\Gamma_D \subseteq \partial\Omega$: Es sei $g : [0, T] \times \Gamma_D \rightarrow \mathbb{R}^d$, dann gelte

$$u(t, x) = g(t, x) \quad \forall (t, x) \in I \times \Gamma_D.$$

- Neumann-Randbedingungen auf $\Gamma_N \subseteq \partial\Omega$: Es sei $j : [0, T] \times \Gamma_N \rightarrow \mathbb{R}^d$, dann gelte

$$(\nu(t, x) \nabla u(t, x) - p(t, x) \mathbb{I}) n(x) = j(t, x) \quad \forall (t, x) \in I \times \Gamma_N.$$

Dabei bezeichne $n(x)$ den äußeren Normalen-Vektor von Ω in den Punkt $x \in \partial\Omega$ und \mathbb{I} die Einheitsmatrix.

Die stationären Probleme entsprechen dem gemischten Problem

$$\begin{aligned}\mathcal{L}u + \nabla p &= f \\ \nabla \cdot u &= 0,\end{aligned}\tag{2.5}$$

wobei \mathcal{L} ein (nicht notwendig linearer) Differentialoperator zweiter Ordnung ist. Es handelt sich um ein *gemischtes Problem*, da die Lösungsgrößen u und p durch die Gleichungen gekoppelt sind. Bei der Diskretisierung und Lösung ergeben sich folglich gewisse Abhängigkeiten der Lösungskomponenten, die etwa in [Gal94] besprochen werden.

Mit geeigneter Zeitdiskretisierung können die instationären Probleme auf Folgen stationärer Probleme zurückgeführt werden. Daher soll im nächsten Abschnitt zunächst auf stationäre Probleme eingegangen werden.

Ein Überblick über gängige Ansätze zur Behandlung der inkompressiblen Navier-Stokes-Gleichungen ist in [Lay08] zu finden.

3 Variationelle Formulierung

Variationsgleichungen sind die Grundlage vieler numerischer Verfahren. Um auch die inkompressiblen Navier-Stokes-Gleichungen numerisch lösen zu können, wird in diesem Kapitel zunächst eine allgemeine Form von Variationsgleichungen zur Beschreibung des gemischten Problems vorgestellt. Die variationelle Formulierung stellt eine Schwächung der Anforderungen an die Lösung dar, so dass wir hierbei von *schwachen Lösungen* sprechen wollen. Im Unterschied dazu stellen klassische Lösungen der Differentialgleichung *starke Lösungen* dar. Die neue Formulierung wird anschließend zur Beschreibung der stationären inkompressiblen Navier-Stokes-Gleichung verwendet. Schließlich werden noch einige Besonderheiten der Formulierung diskutiert, insbesondere verschiedene Typen von Randbedingungen besprochen.

3.1 Funktionsräume und Abbildungen

Im Folgenden seien

- V, Q Hilberträume über \mathbb{R} mit den aus ihren Skalarprodukten erzeugten Normen:

$$\begin{aligned}\|v\| &:= \|v\|_V := \sqrt{(v, v)} \quad \forall v \in V, \\ \|q\| &:= \|q\|_Q := \sqrt{(q, q)} \quad \forall q \in Q.\end{aligned}$$

- $a : V \times V \rightarrow \mathbb{R}$ eine Bilinearform mit den Eigenschaften

– *Koerziv*: Es gibt ein $c_k > 0$, so dass

$$c_k \|v\|_V^2 \leq a(v, v) \quad \forall v \in V.$$

– *Beschränkt*: Es gibt ein $c_{b1} > 0$, so dass

$$a(v, w) \leq c_{b1} \|v\|_V \|w\|_V \quad \forall v, w \in V.$$

- $b : V \times Q \rightarrow \mathbb{R}$ eine Bilinearform mit der Eigenschaft

– *Beschränkt*: Es gibt ein $c_{b2} > 0$, so dass

$$b(v, q) \leq c_{b2} \|v\|_V \|q\|_Q \quad \forall (v, q) \in V \times Q.$$

- $c : V \times V \times V \rightarrow \mathbb{R}$ eine Trilinearform mit der Eigenschaft

– *Beschränkt*: Es gibt ein $c_{b3} > 0$, so dass

$$c(u, v, w) \leq c_{b3} \|u\|_V \|v\|_V \|w\|_V \quad \forall u, v, w \in V.$$

Gegeben seien außerdem stetige lineare Funktionale $f \in V^*$ und $g \in Q^*$.

Betrachtet wird nun das gemischte Problem mit der variationellen Formulierung: Finde $(u, p) \in V \times Q$ mit

$$\begin{aligned} a(u, v) + b(v, p) + c(u, u, v) &= \langle f, v \rangle & \forall v \in V \\ b(u, q) &= \langle g, q \rangle & \forall q \in Q. \end{aligned} \quad (3.1)$$

Wir definieren die folgenden den Bi- bzw. Trilinearformen zugeordneten Operatoren:

- $A : V \rightarrow V^*$, $w \mapsto Aw \Leftrightarrow \langle Aw, v \rangle = a(w, v) \quad \forall v \in V$,
- $B : V \rightarrow Q^*$, $w \mapsto Bw \Leftrightarrow \langle Bw, q \rangle = b(w, q) \quad \forall q \in Q$,
- $B^* : Q \rightarrow V^*$, $r \mapsto B^*r \Leftrightarrow \langle B^*r, v \rangle = b(v, r) \quad \forall v \in V$,
- $C : V \rightarrow V^*$, $w \mapsto Cw \Leftrightarrow \langle Cw, v \rangle = c(w, w, v) \quad \forall v \in V$.

A, B und B^* sind linear und stetig.

Das betrachtete Problem (3.1) kann nun gleichwertig in Operatorformulierung formuliert werden: Finde $(u, p) \in V \times Q$ mit

$$\begin{aligned} Au + B^*p + Cu &= f \\ Bu &= g. \end{aligned}$$

Mit der Quotientenraumnorm

$$\|u\|_{U/W} := \inf_{w \in U} \|u + w\|_U$$

kann die kontinuierliche Inf-Sup-Bedingung formuliert werden:

$$\exists C > 0 : \inf_{q \in Q / \ker B^*} \sup_{v \in V \setminus \{0\}} \frac{b(v, q)}{\|v\|_V \|q\|_{Q / \ker B^*}} \geq C \quad (3.2)$$

oder (äquivalent)

$$\exists C > 0 : \inf_{v \in V / \ker B} \sup_{q \in Q \setminus \{0\}} \frac{b(v, q)}{\|v\|_{V / \ker B} \|q\|_Q} \geq C$$

Sie ist äquivalent dazu, dass $\text{im}(B)$ abgeschlossen in Q^* ist. Vertiefende Betrachtungen sowie Beweise für diesen Zusammenhang sind in [OD96, Gal94] zu finden.

3.2 Anwendung auf die Navier-Stokes-Gleichung

Das inkompressible stationäre Navier-Stokes-Problem (2.2) lässt sich mit den Bezeichnungen

$$\begin{aligned} V &:= (W_2^1(\Omega))^d \\ Q &:= L_2(\Omega) \end{aligned} \quad (3.3)$$

sowie

$$\begin{aligned}
\nu(t, x) &:= \nu_{\text{const}} > 0 & \forall (t, x) \in I \times \Omega, \\
a(u, v) &:= (\nu \nabla u, \nabla v) & \forall u, v \in V, \\
b(u, p) &:= -(\nabla \cdot u, p) & \forall (u, p) \in V \times Q, \\
c(u, v, w) &:= (u \cdot \nabla v, w) & \forall u, v, w \in V, \\
\langle g, q \rangle &:= 0 & \forall q \in Q
\end{aligned} \tag{3.4}$$

als variationelle Formulierung eines gemischtes Problem interpretieren. $W_r^s(\Omega)$ bezeichnet dabei den Sobolev-Raum derjenigen reellwertigen Funktionen $v \in L_r(\Omega)$, deren gemischte partielle schwache Ableitungen bis zur Ordnung s im Lebesgue-Raum $L_r(\Omega)$ liegen, und (\cdot, \cdot) bezeichnet das $L_2(\Omega)$ -Skalarprodukt:

$$(u, v) := \int_{\Omega} u(x)v(x)dx.$$

Wir definieren nun die folgenden Randintegrale:

$$\begin{aligned}
d(u, p) &:= \int_{\partial\Omega} p(u \cdot n) & \forall (u, p) \in V \times Q, \\
e_1(u, v) &:= \int_{\partial\Omega} ((v \cdot \nabla)u) \cdot n & \forall u, v \in V, \\
e_2(u, v) &:= \int_{\partial\Omega} ((n \cdot \nabla)u) \cdot v & \forall u, v \in V.
\end{aligned} \tag{3.5}$$

Bemerkung 3.2.1. Der hier verwendete äußere Normaleneinheitsvektor ist nur für hinreichend glatte Randstücke eindeutig definiert. Problematisch sind also hier die Kanten und Ecken zwischen den stückweise zusammengesetzten Rändern. Für die obigen Definitionen ist dies unerheblich, da ihr Maß 0 ist und wir die Integrale über $\partial\Omega$ einfach durch Integrale über die einzelnen glatten Randstücke $\Gamma_1, \Gamma_2, \dots, \Gamma_n$ ersetzen können.

Satz 3.2.1. Sei $\Omega \in \mathbb{R}^d$ ein beschränktes Gebiet mit C^1 -Rand. Dann existiert ein beschränkter, linearer Operator

$$T : (W_r^1(\Omega))^d \rightarrow (L_r(\partial\Omega))^d,$$

so dass

$$Tu = u|_{\partial\Omega} \quad \text{falls} \quad u \in (W_r^1(\Omega))^d \cap (C(\bar{\Omega}))^d$$

und

$$\|Tu\|_{(L_r(\partial\Omega))^d} \leq C \|u\|_V$$

für alle $u \in V$ gilt. Die Konstante C hängt nur von d, r und Ω ab.

Beweis. Siehe z. B. [AF03]. □

Definition 3.2.1. Wir bezeichnen T aus Satz 3.2.1 als *Spuroperator*.

Im Folgenden setzen wir voraus, dass Ω ein beschränktes Gebiet mit stückweisem C^1 -Rand ist und zusätzlich gilt

$$u, v \in (L_4(\Omega))^d.$$

Dann ist zum einen

$$\int_{\Omega} (u \cdot u)^2 = \|u \cdot u\|_{\Omega}^2 \leq C, \quad \text{d. h.} \quad u \cdot v \in L_2(\Omega).$$

Zum anderen ergibt sich dann mithilfe des Spuroperators aus 3.2.1, dass

$$u \in (L_4(\partial\Omega))^d,$$

und schließlich

$$\int_{\partial\Omega} (u \cdot u)^2 = \|u \cdot u\|_{\partial\Omega}^2 \leq C, \quad \text{d. h.} \quad Tu \cdot Tv \in L_2(\partial\Omega).$$

Lemma 3.2.1. Für das Navier-Stokes-Problem (3.1) mit (3.3) und (3.4) gilt für $u \in V$, $v, w \in V \cap (L_4(\partial\Omega))^d$

$$c(u, v, w) + c(u, w, v) = b(u, v \cdot w) + d(u, v \cdot w),$$

und für $u, v \in V \cap (C^2(\bar{\Omega}))^d$

$$((\nabla u)^\top, \nabla v) = (\nabla \cdot u, \nabla \cdot v) + e_1(u, v) - \int_{\partial\Omega} (\nabla \cdot u)(n \cdot v).$$

Beweis. Produktregel und partielle Integration ergeben

$$\begin{aligned} c(u, v, w) + c(u, w, v) &= \int_{\Omega} \sum_{i,j} (u_i \partial_i v_j w_j + u_i \partial_i w_j v_j) \\ &= \int_{\Omega} \sum_{i,j} u_i \partial_i (v_j w_j) \\ &= - \int_{\Omega} \sum_{i,j} \partial_i u_i v_j w_j + \int_{\partial\Omega} \sum_{i,j} n_i u_i v_j w_j \\ &= b(u, v \cdot w) + \int_{\partial\Omega} (n \cdot u)(v \cdot w). \end{aligned}$$

Durch partielle Integration und den Satz von Schwarz erhalten wir

$$\begin{aligned} ((\nabla u)^\top, \nabla v) &= \int_{\Omega} \sum_{i,j} \partial_i u_j \partial_j v_i \\ &= - \int_{\Omega} \sum_{i,j} \partial_j \partial_i u_j v_i + \int_{\partial\Omega} \sum_{i,j} \partial_i u_j n_j v_i \\ &= \int_{\Omega} \sum_{i,j} \partial_j u_j \partial_i v_i + \int_{\partial\Omega} \sum_{i,j} (\partial_i u_j n_j v_i - \partial_j u_j n_i v_i) \\ &= (\nabla \cdot u, \nabla \cdot v) + \int_{\partial\Omega} (v \cdot \nabla u - v \nabla \cdot u) \cdot n. \end{aligned}$$

□

3.3 Symmetrisierung

Statt der Formulierung (3.1) werden häufig symmetrisierte Terme verwendet:

$$\begin{aligned} a_s(u, v) &:= (\nu(\nabla u + (\nabla u)^\top), \nabla v) & \forall u, v \in V, \\ c_s(u, v, w) &:= \frac{1}{2}(c(u, v, w) - c(u, w, v)) & \forall u, v, w \in V. \end{aligned}$$

Eine Lösung von (2.2) erfüllt auch die veränderte Gleichung mit a_s sowie c_s statt a bzw. c , allerdings entstehen beim Übergang Randintegrale, die ggf. in den Randbedingungen berücksichtigt werden müssen. Diese Randintegrale wollen wir im Folgenden genauer untersuchen.

Für den Zusammenhang zwischen (2.2) und (3.1) gilt durch partielle Integration, für hinreichend glatte (d. h. hinreichend oft differenzierbare) u, v, p, q :

$$\begin{aligned} &(-\nu\Delta u + \nabla p + u \cdot \nabla u, v) - d(v, p) + e_2(u, \nu v) \\ &= a(u, v) + b(v, p) + c(u, u, v). \end{aligned}$$

Nun ist

$$a_s(u, v) - a(u, v) = (\nu\nabla \cdot u, \nabla \cdot v) + e_1(u, \nu v) - \int_{\partial\Omega} \nu(\nabla \cdot u)(n \cdot v)$$

und

$$c_s(u, v, w) - c(u, v, w) = -\frac{1}{2}(c(u, v, w) + c(u, w, v)) = -\frac{1}{2}(b(u, v \cdot w) + d(u, v \cdot w)).$$

Ist u divergenzfrei, also $\nabla \cdot u = 0$, fallen folgende Terme weg:

$$(\nu\nabla \cdot u, \nabla \cdot v) = b(u, u \cdot v) = d(u, u \cdot v) = \int_{\partial\Omega} \nu(\nabla \cdot u)(n \cdot v) = 0.$$

Folglich ist

$$\begin{aligned} &a(u, v) + c(u, u, v) \\ &= a_s(u, v) + c_s(u, u, v) + \frac{1}{2}d(u, u \cdot v) - e_1(u, \nu v). \end{aligned}$$

Unter dieser Voraussetzung bekommen wir schließlich den Zusammenhang

$$\begin{aligned} &(-\nu\Delta u + \nabla p + u \cdot \nabla u, v) - d(v, p) + e_2(u, \nu v) \\ &= a(u, v) + b(v, p) + c(u, u, v) \\ &= a_s(u, v) + b(v, p) + c_s(u, u, v) - e_1(u, \nu v) + \frac{1}{2}d(u, u \cdot v). \end{aligned} \tag{3.6}$$

Bemerkung 3.3.1. Die Randterme gliedern sich wie folgt:

- $e_2(u, \nu v) - d(v, p) = \int_{\partial\Omega} v \cdot (\nu\nabla u - p\mathbb{I})n$ aus dem Übergang von der starken zur schwachen Formulierung.

Zur Veranschaulichung der Auswirkungen betrachten wir Abbildung 3.1: Die Begrenzungen am oberen und unteren Rand sind mit homogenen Dirichlet-Randbedingungen vorgegeben, der Einfluss am linken Rand ist ebenfalls durch Dirichlet-Randbedingungen fixiert. Am rechten Ausflussrand wird mit einer Neumann-Randbedingung gerechnet,

wobei sich die Randintegrale in dem obigen Term unterscheiden. Die Viskosität ist $\nu = 1$. Die Stromlinien sind gemäß der Geschwindigkeit $\|u\|$ eingefärbt, die Hintergrundfarben im Simulationsgebiet stellen die Druckverteilung dar. In (a) tendiert die Strömung dazu, orthogonal zum Rand abzufließen. Insbesondere fließt hier auch Fluid vom „Abflussrand“ in das Simulationsgebiet hinein. In (b) hingegen verhält sich die Strömung gleichmäßiger und ungestörter.

- $-e_1(u, \nu v) = -\int_{\partial\Omega} (\nu(v \cdot \nabla)u) \cdot n$ aus der Symmetrisierung der Diffusion (Symmetrischer Spannungstensor, Übergang von a zu a_s). Dieser Term ist für turbulente Strömungen ($\nu \ll 1$) klein.
- $\frac{1}{2}d(u, u \cdot v) = \frac{1}{2} \int_{\partial\Omega} (u \cdot v)(u \cdot n)$ aus der Symmetrisierung der Konvektion (Übergang von c zu c_s). Dieser Ausdruck ist für turbulente Strömungen nicht zu vernachlässigen.

Siehe dazu etwa Abbildung 3.2. Abermals sind oberer und unterer Rand mit homogenen Dirichlet-Randbedingungen vorgegeben, der Einfluss am linken Rand ist ebenfalls durch Dirichlet-Randbedingungen fixiert. Am rechten Ausflussrand wird mit einer Neumann-Randbedingung gerechnet, wobei sich die Randintegrale in dem obigen Term unterscheiden. Die Viskosität ist $\nu = 10^{-4}$. Die Stromlinien sind gemäß der Geschwindigkeit $\|u\|$ eingefärbt, die Hintergrundfarben im Simulationsgebiet stellen die Druckverteilung dar. Offenbar wirkt die Randbedingung in (a) sich deutlich auf das Strömungsverhalten in Ω aus, bei (b) wird die Strömung erwartungsgemäß nicht wesentlich durch den Rand beeinflusst, sondern schlicht abgeschnitten.

3.4 Physikalische Randbedingungen

Typische zu modellierende Probleme haben physikalische Vorgaben, nach denen der Rand eingeteilt werden kann:

- $\Gamma_{\text{fluid}} \subseteq \partial\Omega$ bezeichne Ränder, durch die Fluid fließen kann. Hier ist also ein Zu- oder Abfluss möglich, aber nicht notwendig.
- $\Gamma_{\text{solid}} \subseteq \partial\Omega$ bezeichne Ränder, durch die kein Fluss des Fluids möglich sein soll, etwa begrenzende Wände.

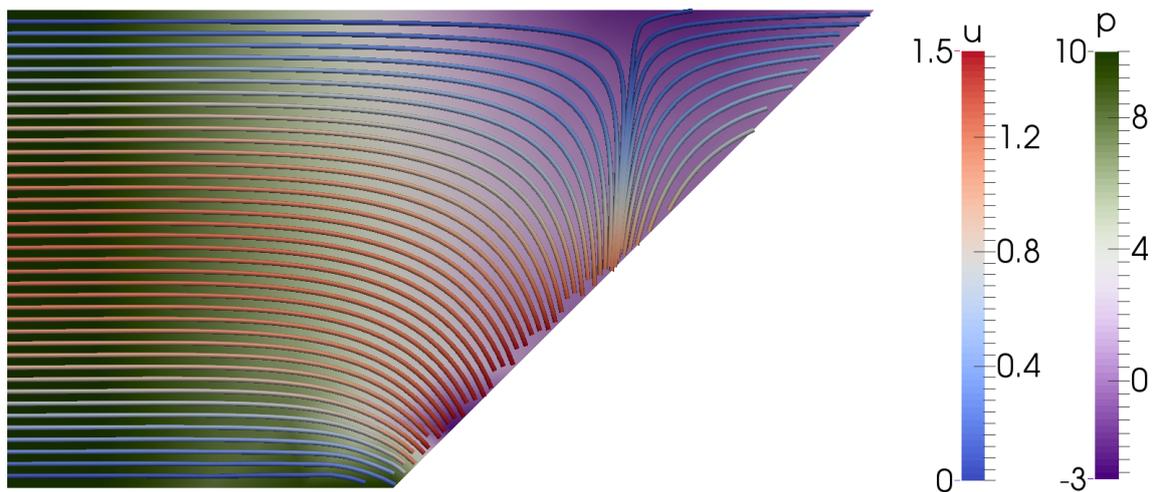
Auch von diesen Kategorien erfasst sind sogenannte poröse Medien. Es handelt sich dabei um Festkörper, die (partiell) durchlässig für das Fluid sind. Gemäß der obigen Definition wollen wir sie zu Γ_{fluid} ordnen.

Wir wollen mit Γ_{fluid} und Γ_{solid} jeweils solche Mengen bezeichnen, die bezüglich $\partial\Omega$ offen sind. Den Abschluss von Γ bezeichnen wir mit $\bar{\Gamma} := \Gamma \cup \partial\Gamma$. Die Übergangskanten bzw. -punkte $\bar{\Gamma}_{\text{solid}} \cap \bar{\Gamma}_{\text{fluid}}$ lassen sich ohnehin meist nicht sinnvoll einer der beiden Mengen zuordnen.

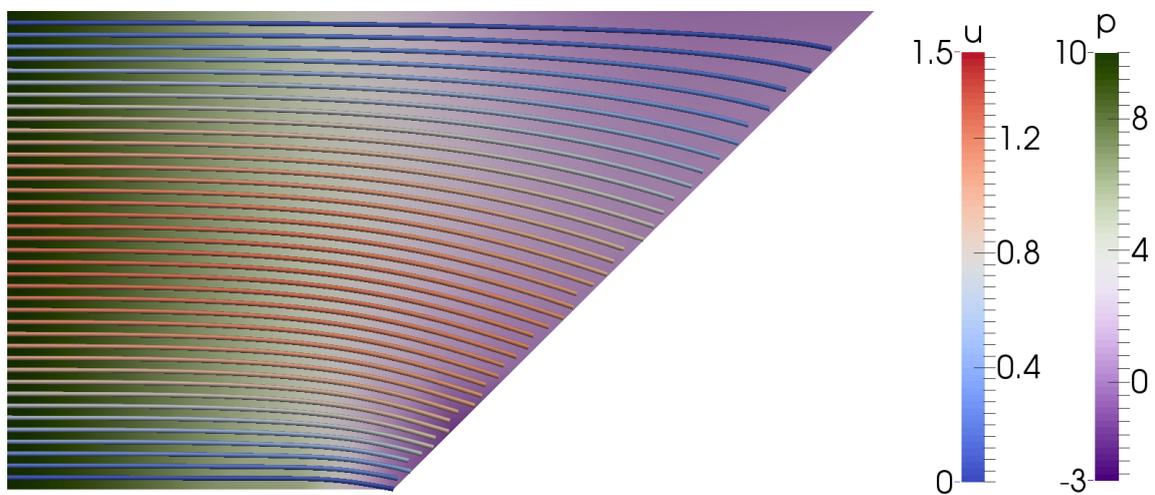
In der Regel gilt

$$\begin{aligned} \bar{\Gamma}_{\text{fluid}} \cup \bar{\Gamma}_{\text{solid}} &= \partial\Omega \text{ und} \\ \Gamma_{\text{fluid}} \cap \Gamma_{\text{solid}} &= \emptyset. \end{aligned}$$

Wir können also sagen, dass Γ_{fluid} und Γ_{solid} eine disjunkte Zerlegung des Randes bilden (bis auf die Abschlüsse).

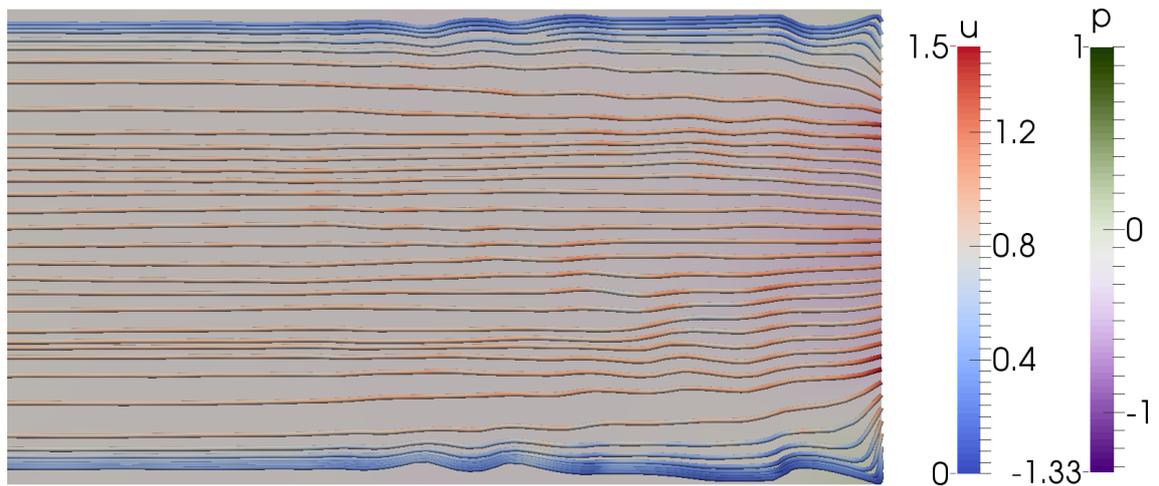


(a): $a(u, v) + b(v, p) + b(u, q) + c(u, u, v) - e_2(u, \nu v) = 0$

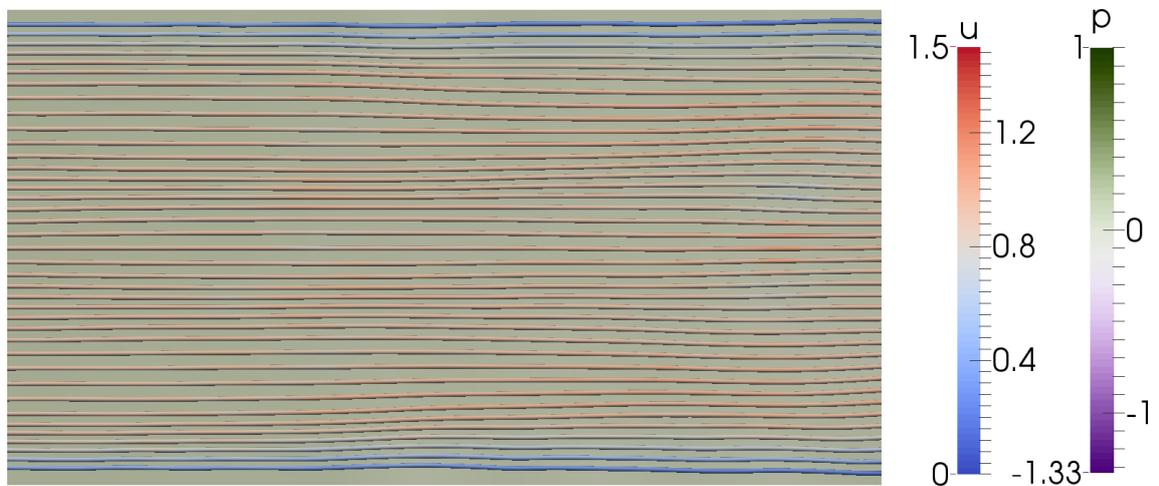


(b): $a(u, v) + b(v, p) + b(u, q) + c(u, u, v) = 0$

Abbildung 3.1: Ausfluss des Fluids aus dem Simulationsgebiet (a) mit bzw. (b) ohne dem Randintegral $\int_{\partial\Omega} n \cdot \nabla u \cdot v$ der Diffusion



(a): $a(u, v) + b(v, p) + b(u, q) + c_s(u, u, v) = 0$



(b): $a(u, v) + b(v, p) + b(u, q) + c_s(u, u, v) + \frac{1}{2}d(u, u \cdot v) = 0$

Abbildung 3.2: Ausfluss des Fluids aus dem Simulationsgebiet (a) ohne bzw. (b) mit dem Randintegral $\frac{1}{2}d(u, u \cdot v)$ der Konvektion

Diese Zerlegung ist keinesfalls a priori gegeben. Etwa bei Problemen mit Wechselwirkung von Fluid und Festkörpern („Fluid-Structure-Interaction“) können diese Grenzen zunächst unbekannt sein und sich erst als Teil der zu berechnenden Lösung ergeben.

Für unsere Betrachtungen wollen wir an dieser Stelle jedoch von einer a priori bekannten Zerlegung ausgehen. Auch sollen in unseren Betrachtungen die Festkörperländer fixiert sein, d. h. etwaige Bewegungen der Festkörper seien von vornherein bekannt.

Beispiel 3.4.1. Unter diesen Voraussetzungen lassen sich die Ränder eines zu simulierenden Strömungsgebietes aufteilen, als Beispiel einer Rohrströmung mit Hindernis siehe etwa Abbildung 3.3. Die Unterteilung erfolgte hier zunächst physikalisch in Festkörper (Γ_{solid}) sowie durchlässig (Γ_{fluid}).

Die Geometrie der Rohrströmung mit Hindernis ist mit dem in [STD⁺96] vorgestellten Problem vergleichbar.

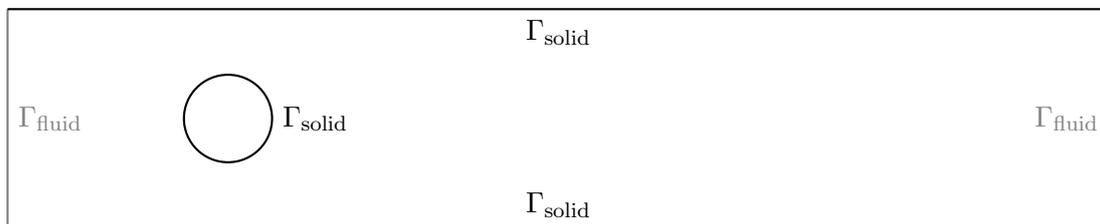


Abbildung 3.3: Exemplarisches Simulationsgebiet: Rohrströmung mit Hindernis in physikalische Ränder unterteilt

3.4.1 Undurchlässiger Rand

Die Bedingungen an den nicht durchlässigen Kontaktflächen können im einfachsten Fall durch Dirichlet-Randbedingungen für die Geschwindigkeit beschrieben werden:

$$u = u_D \quad \text{auf } \Gamma_D.$$

Hierbei ist $\Gamma_D \subseteq \partial\Omega$ eine abgeschlossene Menge (abgeschlossen als Teilmenge des Randes) und $u_D : I \times \Gamma_D \rightarrow \mathbb{R}^d$ gegeben. Mit dieser Randbedingung wird beschrieben, dass das Fluid an der Kontaktfläche die Geschwindigkeit des Festkörpers annimmt, d. h. es entsteht kein Schlupf („no-slip“).

Bemerkung 3.4.1. Um an diesem Rand einen Zu- oder Abfluss zu vermeiden, ist

$$u_D \cdot n = 0 \quad \text{auf } \Gamma_D \cap \Gamma_{\text{solid}}.$$

In stückweise zusammengesetzten Rändern lassen sich verschiedene äußere Normalenvektoren $n(x)$ finden: Für jedes angrenzende glatte Randstück kann der Grenzwert der Normaleneinheitsvektoren bei Annäherung an die kritische Übergangskante bestimmt werden. Die obige Bedingungen gilt in diesen Punkten dann jeweils für alle betreffenden Normaleneinheitsvektoren. Auch später werden noch ähnliche Bedingungen bezüglich des Randes mitsamt zugehörigem Normalenvektor auftauchen. Sie sollen analog in diesen Übergangspunkten für alle Normaleneinheitsvektoren gelten.

Insbesondere aufgrund des schnellen Anstiegs des Geschwindigkeitsfeldes in der Nähe des Randes können diese Kontaktflächen jedoch auch durch Schlupfbedingungen auf einem Randstück $\Gamma_S \subseteq \partial\Omega$ modelliert werden:

$$u \cdot n = 0 \quad \text{und} \quad (2\nu\nabla^s un)_\tau = \alpha(u - u_D) \quad \text{auf } \Gamma_S,$$

wobei $v_\tau := v - (v \cdot n)n$ die Projektion des Vektors v auf die Tangentialebene, die durch den Normaleneinheitsvektor n beschrieben wird. $\alpha \geq 0$ ist ein Parameter, mit dem der Schlupf gesteuert werden kann.

Bemerkung 3.4.2. Für den Schlupfparameter α ergeben sich folgende Spezialfälle:

- $\alpha = 0$ entspricht einer Oberfläche ohne Reibung, für die die Geschwindigkeit des Festkörpers unerheblich ist. Hier ist also maximaler Schlupf möglich.
- Der Grenzfall $\alpha \rightarrow \infty$ entspricht der Dirichlet-Randbedingung, also keinem Schlupf.

Mathematisch wird dieser Randbereich mit Schlupf durch Neumann-Randbedingungen modelliert. Darauf wird weiter unten noch im Detail eingegangen.

Wir haben also im Wesentlichen zwei Möglichkeiten, das Verhalten an den Kontaktflächen zu Festkörpern zu modellieren.

- Ohne Schlupf („no-slip“) als Dirichlet-Randbedingung oder
- Mit Schlupf („slip“) als Neumann-Randbedingung.

Die Frage, welche dieser beiden die physikalischen Gegebenheiten besser wiedergeben, ist durchaus noch Gegenstand der Forschung und hängt auch vom Einzelfall ab.

3.4.2 Durchlässiger Rand

Für den Rand zum Fluid Γ_{fluid} gibt es ebenfalls verschiedene Möglichkeiten der Implementierung.

Eine einfache Vorgehensweise ist auch hier die fixierte Vorgabe eines Einflusses des Fluids. Ist etwa die typische Rohrströmung beim Einfluss in das Simulationsgebiet bekannt, kann diese mit der folgenden Dirichlet-Randbedingung implementiert werden:

$$u = u_D \quad \text{auf } \Gamma_D.$$

Zu beachten ist, dass im Unterschied zum undurchlässigen Rand meist

$$u_D \cdot n \neq 0 \quad \text{auf } \Gamma_D \cap \Gamma_{\text{fluid}}$$

gilt.

Wenn die Flussgeschwindigkeit nicht bekannt ist, so wird auch hier mit Neumann-Randbedingungen gearbeitet. Eine meist sinnvolle Alternative ist die Vorgabe der Spannung des Fluids im Rand. Wir können sie mithilfe des Spannungstensors

$$T(\nu, u, p) := \nu\nabla^s u - p\mathbb{I}$$

bestimmen. Hierbei ist

$$\nabla^s u := \frac{1}{2} \left(\nabla u + (\nabla u)^\top \right).$$

Eine Spannungsvorgabe auf dem Neumann-Rand sieht dann so aus:

$$T(\nu, u, p)n = j_N \quad \text{auf } \Gamma_N. \quad (3.7)$$

Speziell den Fall $j_N \equiv 0$ wollen wir mit „do-nothing“-Rand bezeichnen. Im Allgemeinen kann j_N aber von Zeit und Ort abhängen, d. h. $j_N : I \times \Gamma_N \rightarrow \mathbb{R}^d$.

Eine solche „do-nothing“-Randbedingung wird häufig für Ausflüsse verwendet. Das Fluid wird somit „spannungslos“ aus dem Simulationsgebiet entlassen. Die Annahme eines solchen Verhaltens ist jedoch nicht immer sinnvoll. Oft ist der Ausflussrand ohnehin ein künstlicher Rand. Er existiert nur, um das Gebiet abzuschließen und ein kompaktes Simulationsgebiet zu bekommen. Das Strömungsverhalten soll durch ihn jedoch nicht verändert werden. Die Strömung soll im Idealfall so verlaufen, als gäbe es diesen Rand gar nicht – etwa so, als wäre das Simulationsgebiet größer, aber an dieser Stelle abgeschnitten.

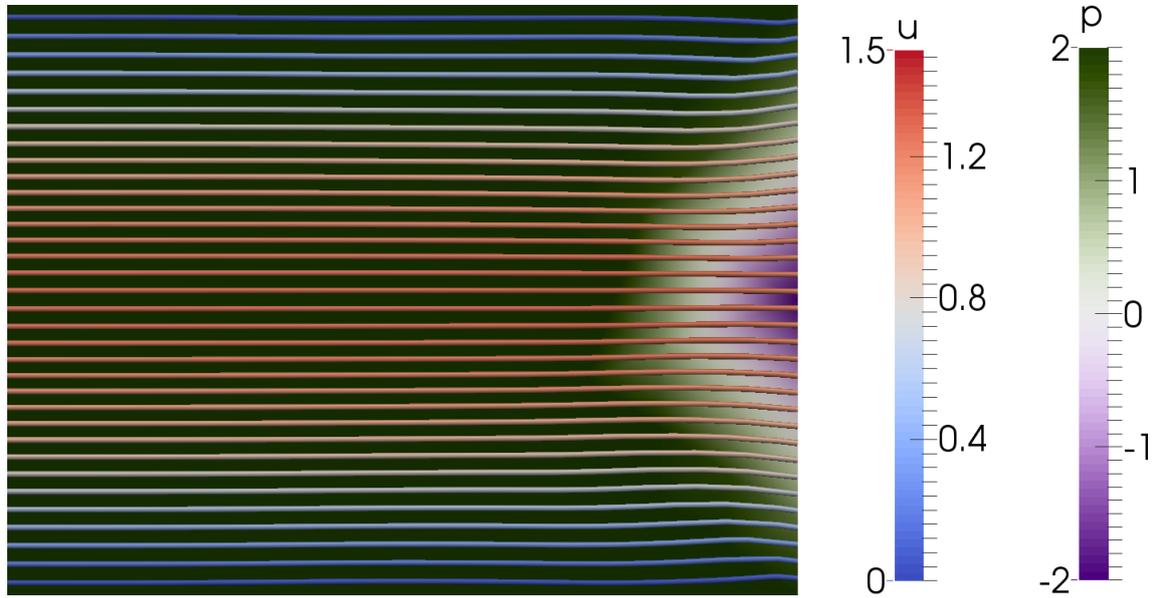
Der Unterschied zwischen „spannungslos“ und „keine Störung“ ist gut in Abbildung 3.4 nachvollziehbar. Bei der Interpretation der Randbedingungen für eine Lösung von (2.2) ergibt sich in (a): $(\nu \nabla^s u - p\mathbb{I})n = 0$ auf Γ_N , in (b): $(\nu \nabla u - p\mathbb{I})n = \frac{1}{2}e_1(u, \nu v) - e_2(u, \nu v)$ auf Γ_N . Während (a) das Fluid spannungsfrei entlässt, wird bei (b) am Ausflussrand die Spannung aufgedrückt, die in einer Rohrströmung vorherrscht. So wird ein künstlicher Abschnitt des Simulationsgebiet in diesem Fall besser modelliert.

Die mathematische Modellierung eines Randes, der die Strömung „nicht stört“, ist nicht einfach, wie etwa der Diskussion in [HRT96] zu entnehmen ist. Mit der konkreten Modellierung eines solchen künstlichen Randes, insbesondere mit solchen mit Rückflüssen, hat sich [Bru00] auseinandergesetzt. Die dort verwendeten Randbedingungen für künstliche („artificial“) Ränder sind

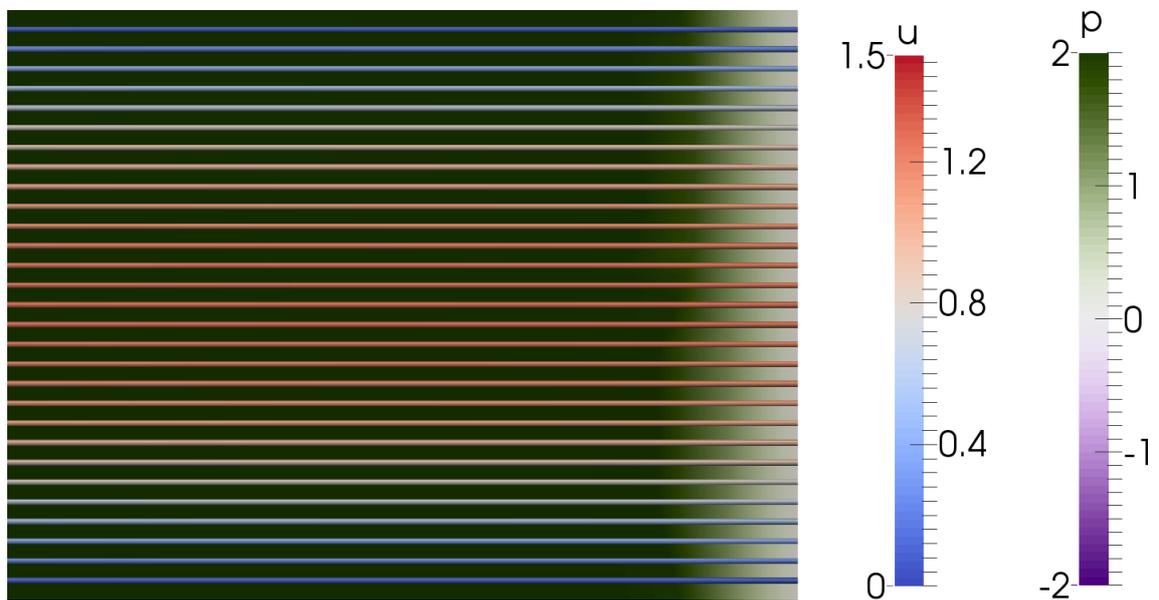
$$T(\nu, u, p)n + \frac{1}{2}(u \cdot n)^-(u - u_{\text{ref}}) = T(\nu_{\text{ref}}, u_{\text{ref}}, p_{\text{ref}})n \quad \text{auf } \Gamma_{\text{art}}, \quad (3.8)$$

wobei $(\nu_{\text{ref}}, u_{\text{ref}}, p_{\text{ref}})$ eine Referenzlösung darstellt, die etwa aus dem stationären Stokes-Problem (2.1) gewonnen werden kann. Bei instationären Problemen kann auch die Lösung eines bereits bestimmten Zeitpunktes verwendet werden. $(s)^- := \frac{1}{2}(s - |s|)$ bezeichnet hierbei den negativen Teil von $s \in \mathbb{R}$.

Die durch die Verwendung dieser Randbedingung verursachten Auswirkungen sind in Abbildung 3.5 zu erkennen. Dort sind oberer und unterer Rand mit homogenen Dirichlet-Randbedingungen vorgegeben, der Einfluss am linken Rand ist ebenfalls durch Dirichlet-Randbedingungen fixiert. Am rechten Ausflussrand wird mit einer Neumann-Randbedingung gerechnet, wobei in (a) do-nothing (3.7) implementiert ist ($j_N = 0$), in (b) hingegen ein künstlicher Rand (3.8) nach Bruneau [Bru00]. Die Viskosität ist $\nu = 2.5 \cdot 10^{-3}$. Die Stromlinien sind gemäß der Geschwindigkeit $\|u\|$ eingefärbt, die Hintergrundfarben im Simulationsgebiet stellen die Druckverteilung dar. In (a) tendieren Verwirbelungen dazu, von dem Rand aufgelöst zu werden. In (b) hingegen erscheint die Strömung am Ausflussrand eher, als wäre es das abgeschnittene Bild eines größeren Simulationsgebietes.



(a): $a(u, v) + b(v, p) + b(u, q) + c(u, u, v) + \frac{1}{2}e_1(u, \nu v) - \frac{1}{2}e_2(u, \nu v) = 0$



(b): $a(u, v) + b(v, p) + b(u, q) + c(u, u, v) + \frac{1}{2}e_1(u, \nu v) - \frac{1}{2}e_2(u, \nu v) = \frac{1}{2}e_1(u, \nu v) - \frac{1}{2}e_2(u, \nu v)$

Abbildung 3.4: Ausfluss des Fluids aus dem Simulationsgebiet an einem Neumann-Rand: (a) do-nothing bzw. (b) vorgegebene Rohrspannung

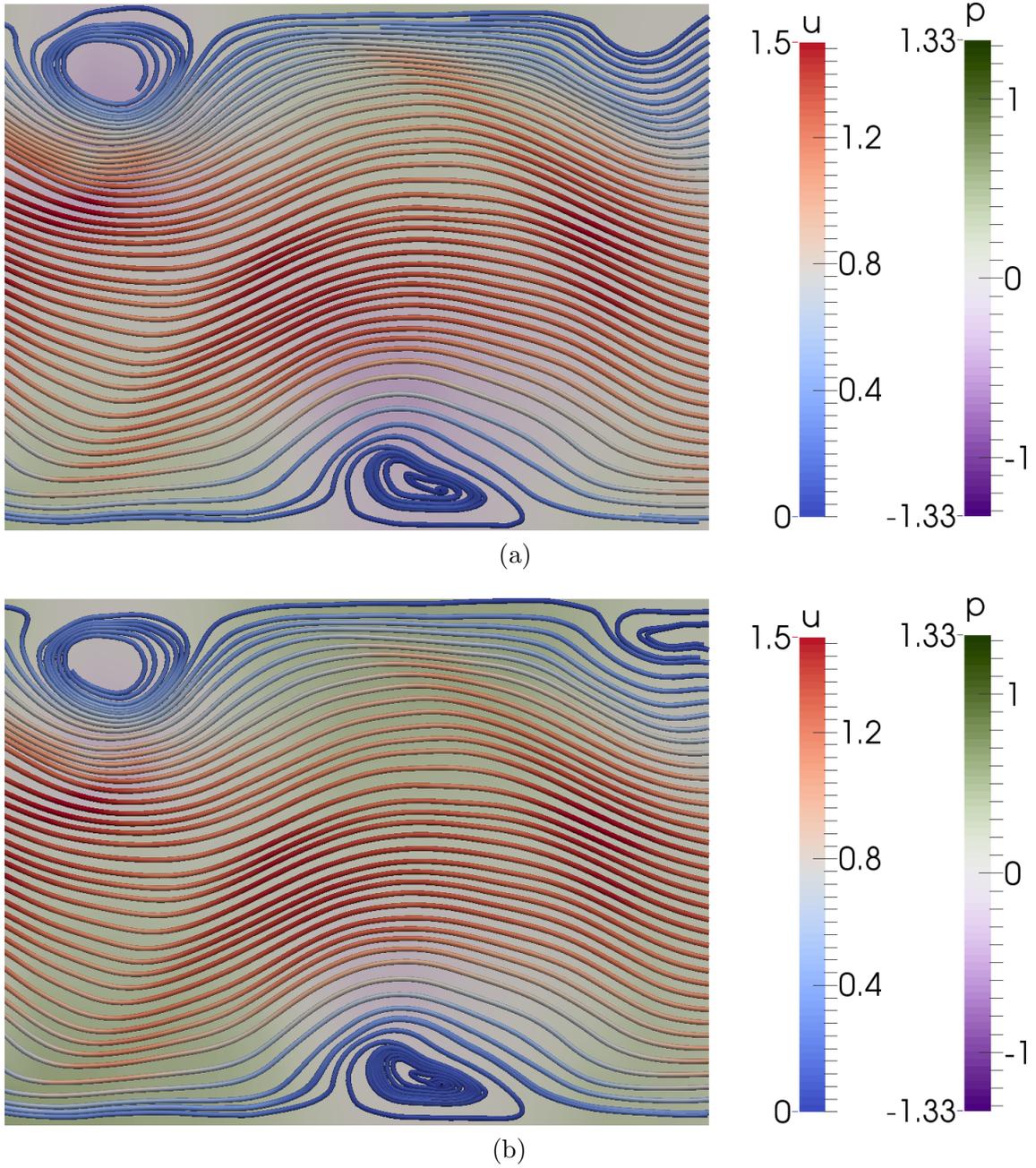


Abbildung 3.5: Ausfluss des Fluids aus einem verwirbelten Simulationsgebiet: Randbedingung (a) do-nothing bzw. (b) künstlich

3.4.3 Mathematische Formulierung

All diese „besonderen“ Randmodellierungen wie schlupfbehaltete oder künstliche Ränder lassen sich ebenfalls als Neumann-Randbedingungen modellieren. Um die Notation nicht unnötig zu erschweren, sollen die passenden Randbedingungen im Folgenden in der rechten Seite j_N vereint sein, dazu lassen wir hier auch eine Abhängigkeit von der Lösung (u, p) zu.

Beispiel 3.4.2. Im Beispiel 3.4.1 von zuvor können wir die Randbedingungen nun mathematisch formulieren, siehe Abbildung 3.6. Für das mathematische Modell sind die Ränder jetzt nach Dirichlet- (Γ_D) bzw. Neumann-Randbedingungen (Γ_N) eingeteilt. Zu beachten ist, dass $u_D \cdot n = 0$ auf Γ_{solid} . Am Einflussrand links kann die Rohrströmung mit $u_D \cdot n \neq 0$ vorgegeben werden.

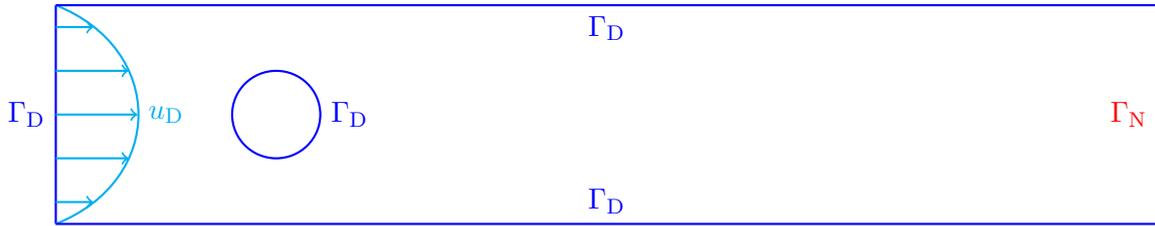


Abbildung 3.6: Randmodellierung für das exemplarisches Simulationsgebiet

Die typischen Randbedingungen sind folglich:

- Dirichlet-Rand Γ_D : Es gelte $u = u_D$ auf Γ_D mit $u_D : \Gamma_D \rightarrow \mathbb{R}^d$.
- Neumann-Rand Γ_N : Es gelte $(\nu \nabla^s u - p \mathbb{I})n = j_N$ auf Γ_N mit $j_N : I \times \Gamma_N \times V \times Q \rightarrow \mathbb{R}^d$.

Dabei soll $\Gamma_D \cup \Gamma_N = \partial\Omega$ sein.

Diese Randbedingungen müssen wir nun variationell formulieren. Dazu sei

$$V^\circ := \{v \in V \mid v|_{\Gamma_D} = 0\} .$$

Wir schränken den Raum der Ansatz- bzw. Testfunktionen also so ein, dass am Dirichlet-Rand keine Veränderung der Funktionswerte möglich ist. Dann suchen wir $(u, p) \in V \times Q$ mit $u - u_D \in V^\circ$, so dass die variationelle Formulierung des Problems für alle $(v, q) \in V^\circ \times Q$ erfüllt ist. Oder anders ausgedrückt: Mit

$$u_D + V^\circ := \{u = u_D + u^\circ \in V \mid u^\circ \in V^\circ\} ,$$

suchen wir $(u, p) \in (u_D + V^\circ) \times Q$, wobei (v, q) beliebig aus $V^\circ \times Q$ ist.

Für die Neumann-Randbedingungen wiederum sind nun die Randintegrale d, e_1, e_2 aus (3.5) relevant. Wegen der Dirichlet-Randbedingungen wird der Testraum V zu V° eingeschränkt. Da jedoch jedes $v \in V^\circ$ auf dem Dirichlet-Rand verschwindet, $v|_{\Gamma_D} \equiv 0$, und in jedem Integranden der Faktor v enthalten ist, ist

$$\begin{aligned} \int_{\Gamma_D} p(u \cdot n) &= 0 & \forall (u, p) \in V^\circ \times Q, \\ \int_{\Gamma_D} ((v \cdot \nabla)u) \cdot n &= 0 & \forall u, v \in V^\circ. \end{aligned}$$

Wegen $\partial\Omega \setminus \Gamma_D = \Gamma_N$ können wir dann

$$\begin{aligned} d^\circ(u, p) &:= \int_{\Gamma_N} p(u \cdot n) & \forall (u, p) \in V^\circ \times Q, \\ e_1^\circ(u, v) &:= \int_{\Gamma_N} ((v \cdot \nabla)u) \cdot n & \forall u, v \in V^\circ, \\ e_2^\circ(u, v) &:= \int_{\Gamma_N} ((n \cdot \nabla)u) \cdot v & \forall u, v \in V^\circ \end{aligned} \quad (3.9)$$

definieren und verwenden.

Mit dem Wissen dieses Kapitels können wir nun das gemischte Randwertproblem

$$\begin{aligned} u \cdot \nabla u - \nabla \cdot (\nu \nabla u - p\mathbb{I}) &= f & \text{in } \Omega, \\ \nabla \cdot u &= 0 & \text{in } \Omega, \\ u &= u_D & \text{auf } \Gamma_D, \\ (\nu \nabla^s u - p\mathbb{I})n &= j_N & \text{auf } \Gamma_N \end{aligned}$$

in die folgende variationelle Formulierung überführen:

Finde $(u, p) \in (u_D + V^\circ) \times Q$ mit

$$\begin{aligned} a(u, v) + b(v, p) + b(u, q) + c(u, u, v) \\ + \frac{1}{2}e_1^\circ(u, \nu v) - \frac{1}{2}e_2^\circ(u, \nu v) &= \langle f, v \rangle + \int_{\Gamma_N} j_N \cdot v \quad \forall (v, q) \in V^\circ \times Q, \end{aligned} \quad (3.10)$$

oder gleichwertig (vgl. (3.6)):

$$\begin{aligned} a_s(u, v) + b(v, p) + b(u, q) + c_s(u, u, v) \\ - \frac{1}{2}e_1^\circ(u, \nu v) - \frac{1}{2}e_2^\circ(u, \nu v) + \frac{1}{2}d^\circ(u, u \cdot v) &= \langle f, v \rangle + \int_{\Gamma_N} j_N \cdot v \quad \forall (v, q) \in V^\circ \times Q. \end{aligned} \quad (3.11)$$

Mit den gemischten Räumen

$$\begin{aligned} \hat{V} &:= V^\circ \times Q, \\ \hat{V}^* &= (V^\circ)^* \times Q^* \end{aligned}$$

sowie den darin enthaltenen Funktionen

$$\begin{aligned} \hat{u} &:= (u, p) \in \hat{V} \\ \hat{u}_D &:= (u_D, 0) \in \hat{V} \\ \hat{v} &:= (v, q) \in \hat{V} \\ \hat{f} &:= (f, g) \in \hat{V}^*, \end{aligned}$$

d. h.

$$\langle \hat{f}, \hat{v} \rangle = \langle f, v \rangle + \langle g, q \rangle \quad \forall \hat{f} \in \hat{V}^* \quad \forall \hat{v} \in \hat{V}$$

wollen wir (3.11) nun auch zusammenfassen zu: Finde $\hat{u} \in \hat{u}_D + \hat{V}$ mit

$$\hat{a}(\hat{u}, \hat{v}) = \langle \hat{f}, \hat{v} \rangle \quad \forall \hat{v} \in \hat{V}.$$

Zu beachten ist, dass hierbei $\hat{a}: \hat{V} \times \hat{V} \rightarrow \mathbb{R}$ nur linear im zweiten Argument, jedoch im ersten nichtlinear ist.

3.5 Existenz und Eindeutigkeit

Zentrale mathematische Fragestellungen sind die nach Lösbarkeit (Existenz einer Lösung) sowie nach Eindeutigkeit oder sogar Wohldefiniertheit (Eindeutigkeit der Lösung und stetige Abhängigkeit von den Daten) des gestellten Problems.

Das dafür betrachtete instationäre Problem können wir variationell wie folgt formulieren: Gesucht ist die Abbildung $[0, T) \ni t \mapsto (u(t), p(t)) = \hat{u}(t) \in \hat{V}$ mit

$$\begin{aligned} \partial_t(u(t), v) + a(u(t), v) + b(v, p(t)) \\ + b(u(t), q) + c(u(t), u(t), v) = \langle f(t), v \rangle \quad \forall t \in (0, T) \forall (v, q) \in \hat{V}, \\ u(0) = u_0. \end{aligned} \quad (3.12)$$

Die spezielle Fragestellung nach Existenz einer starken Lösung für $d = 3$ ($\Omega \subset \mathbb{R}^d$) ist insbesondere als eines der Millennium-Probleme des Clay Mathematics Institute vertraut [Fef00]. In einer derart allgemeinen Formulierung ist hier jedoch noch keine Lösung bekannt. Die Lösungstheorie deckt lediglich einige Spezialfälle ab, über die gesicherte Aussagen erwiesen sind. Für $d = 2$ und $\Omega \in C^\infty$ gilt:

Satz 3.5.1. *Sei $u_0 \in (L^2(\Omega))^2$, $f \in (L^2((0, T) \times \Omega))^2$, $\nabla \cdot u_0 = 0$ und $u_0 \cdot n = 0$ auf $\Gamma = \Gamma_D$. Dann hat (3.12) eine eindeutige Lösung u mit*

$$u_i, \partial_j u_i \in L^2((0, T) \times \Omega), \quad i, j = 1, 2,$$

und $p \in L^2((0, T) \times \Omega)$.

Starke Lösungen existieren unter folgenden Voraussetzungen:

Satz 3.5.2. *Sei $u_0 \in V$, $f \in (L^2((0, T) \times \Omega))^2$. Dann hat (3.12) eine eindeutige Lösung u mit*

$$u_i, \partial_t u_i, \partial_j u_i, \partial_j \partial_k u_i \in L^2((0, T) \times \Omega), \quad i, j, k = 1, 2,$$

und $p, \partial_i p \in L^2((0, T) \times \Omega)$, $i = 1, 2$.

Für $d = 3$ jedoch gilt nur:

- Es existiert eine schwache (nicht glatte) Lösung für alle $t > 0$, aber sie ist nicht notwendig eindeutig.
- Sofern eine starke Lösung für alle $t > 0$ existiert, ist sie eindeutig.
- Eine starke Lösung existiert auf einem gewissen Intervall $(0, T^*)$, wobei T^* von den Daten abhängt.

- Sofern eine starke Lösung auf einem Zeitintervall existiert, dann ist dies auch die einzige schwache Lösung auf diesem Zeitintervall.

Diese Behauptungen sollen an dieser Stelle nicht nachgewiesen werden, sondern sollen nur einen Überblick geben. Für ausführlichere Erläuterungen sei hier auf [MT98] verwiesen.

4 Diskretisierung

Die vorgestellten Probleme sind nur in wenigen Spezialfällen analytisch lösbar. Auch bei komplexeren Geometrien ist eine numerische Approximation nötig. Statt des Problems

$$\text{finde } u \in V \text{ mit } a(u, v) = \langle f, v \rangle \quad \forall v \in V$$

über dem unendlichdimensionalen Funktionsraum V wird ein ähnliches Problem

$$\text{finde } u^h \in V^h \text{ mit } a^h(u^h, v^h) = \langle f^h, v^h \rangle \quad \forall v^h \in V^h$$

über einem endlichdimensionalen Raum V^h (häufig $V^h \subset V$) betrachtet. Es gibt also bestimmte Übergänge in der Formulierung:

$$\begin{array}{lll} V & \rightsquigarrow & V^h, \\ a : V \times V \rightarrow \mathbb{R} & \rightsquigarrow & a^h : V^h \times V^h \rightarrow \mathbb{R}, \\ f \in V^* & \rightsquigarrow & f^h \in (V^h)^*. \end{array}$$

Sofern das diskrete Problem gelöst werden kann, besteht eine gewisse Hoffnung, dass die diskrete Lösung u^h als Approximation für u dient, also $\|u^h - u\|$ für eine hinreichend große Dimension von V^h klein ist. Die Lösbarkeit des diskreten Problems sowie die Größe dieses Fehlers hängen wesentlich von der konkreten Wahl der diskreten Größen V^h , a^h und f^h ab.

4.1 Diskrete Funktionsräume

Gesucht ist eine endliche Menge von Funktionen $\varphi_i : \Omega \rightarrow \mathbb{R}$, $i = 1, \dots, N$, mit deren Linearkombinationen typische Funktionen $v \in V$ (insbesondere die Lösung) gut approximiert werden können. Wenn wir also

$$V^h := \left\{ \sum_{i=1}^N \alpha_i \varphi_i \mid (\alpha_i) \in \mathbb{R}^N \right\}$$

setzen, entspricht das der Forderung, dass

$$\min_{v^h \in V^h} \|v - v^h\| = \min_{(\alpha_i) \in \mathbb{R}^N} \left\| v - \sum_{i=1}^N \alpha_i \varphi_i \right\|$$

klein ist. Die Frage ist folglich, wie diese „typischen“ Funktionen aussehen. Leider kennen wir die Lösung des Problems in der Regel nicht vorher, deswegen können wir diese Frage nicht so einfach beantworten. Allerdings kann aus physikalischem Verhalten bzw. der Formulierung

der Differentialgleichung abgeleitet werden, dass typische Lösungen in einer gewissen Form glatt erscheinen. Es liegt daher nahe, an φ_i die Voraussetzung (stückweiser) Stetigkeit oder Differenzierbarkeit bis zu einem gewissen Grad zu stellen.

Aus Implementierungs- und Aufwandssicht ist es zudem wünschenswert, wenn die Funktionen φ_i möglichst nur lokal wirken. Denn wenn jedes φ_i auf einem Großteil von Ω verschwindet, beschleunigt das die Auswertung von a^h . Außerdem ermöglicht dies die Parallelisierung von Arbeitsschritten, etwa indem Ω in disjunkte Partitionen zerlegt wird, die dann unabhängig voneinander und somit potentiell gleichzeitig bearbeitet werden können.¹

Die Kombination dieser Anforderungen legt die relativ verbreitete Wahl von Splines bzw. stückweise definierter Polynome für V^h nahe. Grundlage hierbei ist eine Zerlegung von Ω in polygonal berandeter Gebiete T_i :

$$\mathcal{T} = \{T_i \mid i = 1, \dots, N_T\} .$$

Für diese Gebiete sind relativ simple Geometrien verbreitet, etwa Drei- oder Vierecke für $d = 2$ und Tetraeder, Prismen oder Quader für $d = 3$. Die Funktionen in $\mathcal{S}_n^{\mathcal{T}}$ sollen auf jedem dieser lokalen Gebiete ein Polynom mit Maximalgrad n sein, also

$$\forall s \in \mathcal{S}_n^{\mathcal{T}} \forall i \in \{1, \dots, N_T\} : s|_{T_i} \in \mathbb{P}_n .$$

Außerdem soll ggf. global noch eine gewisse Glattheit vorherrschen, d. h.

$$\mathcal{S}_{n,r}^{\mathcal{T}} := \mathcal{S}_n^{\mathcal{T}} \cap C^r(\overline{\Omega}) . \quad (4.1)$$

Für $r \geq 0$ werden die Funktionsräume *konforme Elemente* genannt. Im Unterschied dazu gibt es auch *nichtkonforme Elemente*, bei denen die globale Stetigkeit nicht gefordert ist. Wenn etwa die Freiheitsgrade als Funktionswerte in den Schwerpunkten der Seiten festgelegt werden, ist lediglich Stetigkeit in diesen Punkten gesichert, jedoch nicht zwangsläufig in anderen Punkten des Elementrandes. Dieses Element wird auch *Crouzeix-Raviart-Element* genannt und geht auf [CR73] zurück. Um den Umstand zu verdeutlichen, dass zwar in bestimmten Punkten der Elementübergänge Stetigkeit vorherrscht, daraus jedoch noch keine globale Stetigkeit folgt, wird in (4.1) gern $r = -1$ geschrieben.

Die Unterscheidung in konforme und nichtkonforme Elemente erfolgt manchmal auch nach $V^h \subset V$ (konform) oder $V^h \not\subset V$ (nichtkonform). Sofern

$$V \subseteq C(\overline{\Omega}) ,$$

d. h. der Funktionsraum V nur aus stetigen Funktionen besteht (was typischerweise der Fall ist), ist dies gleichbedeutend mit den oben verwendeten Bezeichnungen.

Relativ häufig genutzte Elemente sind etwa

- stückweise konstante Polynome $\mathcal{S}_0^{\mathcal{T}}$,
- stückweise lineare Polynome $\mathcal{S}_{1,-1}^{\mathcal{T}}$, z. B. *Crouzeix-Raviart-Element*,
- stückweise quadratische Polynome $\mathcal{S}_{2,-1}^{\mathcal{T}}$, z. B. *Morley-Element*,

¹Abgesehen von den Rändern, an denen noch ein Austausch stattfinden muss.

- stetige Funktion stückweise linearer Polynome $\mathcal{S}_{1,0}^{\mathcal{T}}$ (*Courant-Dreieck*),
- stetige Funktion stückweise quadratischer Polynome $\mathcal{S}_{2,0}^{\mathcal{T}}$ (*Lagrange-Elemente*).

Die so erzeugten Funktionsräume werden gerne *Finite-Element-Räume* genannt. Dies hängt mit dem Namen *Finite-Elemente-Methode* des Verfahrens zusammen, das hier beschrieben werden soll. Etwas verwirrend kann jedoch die Bezeichnung *Element* dabei sein: Während es im engeren Sinn ein geometrisches, durch das Gitter abgegrenztes Teilstück von Ω bezeichnet (also T_i), hat sich der Begriff auch für die Funktionsräume (etwa $\mathcal{S}_{1,-1}^{\mathcal{T}}$) verbreitet. Streng genommen handelt es sich dabei um die *Finiten Elemente*, jedoch ist bereits an den oben eingeführten Bezeichnungen erkennbar, dass hier gerne abgekürzt wird. Für weitergehende Betrachtungen sei an dieser Stelle auf [Bra07] verwiesen.

4.2 Gitterbasierte Zerlegung

Einen bedeutsamen Einfluss auf einen derartig gewählten Funktionsraum hat die Zerlegung von Ω . Eine wesentliche Voraussetzung an die Existenz einer Zerlegung in polygonal berandete Gebiete ist, dass der Rand von ganz Ω polygonal ist. Ist Ω nicht polygonal berandet, ist auch eine Behandlung mit krummrandigen Zerlegungen denkbar – der wesentliche Unterschied ist die Verwendung einer anders gearteten Bijektion als der affin-linearen. Alternativ genügt es häufig auch, den krummen Rand hinreichend gut durch Annäherung mit polygonalen Rändern zu modellieren.

Wir gehen im Folgenden davon aus, ein polygonal berandetes Gebiet Ω zu betrachten, zu dem eine polygonal berandete Zerlegung \mathcal{T} existiert. Es ist jedoch bekannt, dass die Qualität von V^h sehr von der verwendeten Zerlegung \mathcal{T} abhängen kann. Eine bessere Zerlegung kennzeichnet dabei insbesondere:

- Regelmäßigkeit, d. h. keine stark verzerrten Elemente, sondern möglichst solche, bei denen alle Kanten, Flächen, Winkel ähnlich groß sind.
- Konforme Zerlegung, d. h. keine „hängenden Knoten“. Jede Randfläche bzw. -kante bzw. jeder Eckpunkt inzidiert genau mit dem entsprechenden Teil eines Nachbarn, oder formell: Für alle i, j gilt $\partial T_i \cap \partial T_j = T_i \cap H(\partial T_i \cap \partial T_j)$ mit der affin-linearen Hülle

$$H(M) := \{x + \alpha(y - x) \mid x, y \in M, \alpha \in \mathbb{R}\}$$

einer Menge $M \subseteq \Omega$.

- Gleichartige Elemente, d. h. es existieren affin-lineare oder multilineare Bijektionen zwischen je zwei $T_i, T_j \in \mathcal{T}$.

Unabhängig von der Qualität der Zerlegungen wollen wir später (in den Kapiteln 5 und 6) auch noch nach unterschiedlicher Feinheit von Gittern unterscheiden. Dazu wird die mittlere oder maximale Kantenlänge im Gitter eine Größenordnung vorgeben.

Definition 4.2.1. Wir bezeichnen ein fixiertes „Referenzelement“ als \tilde{T} . Dann können wir für jedes Element T aus \mathcal{T} , dessen Geometrie dem Referenzelement \tilde{T} entspricht, eine affin-lineare Bijektion $F_T : \tilde{T} \rightarrow T$ definieren. F_T heißt die Referenzabbildung und es gilt

$$F_T(\tilde{x}) = A_T \tilde{x} + b_T$$

mit $A_T \in \mathbb{R}^{d \times d}$, $b_T \in \mathbb{R}^d$.

Um den Funktionswert $v^h(x)$ für ein $x \in \Omega$ von einem $v^h \in V^h$ zu bestimmen, können wir also die Referenzabbildung verwenden. Dazu suchen wir uns zunächst das $T \in \mathcal{T}$ mit $x \in T$, und können dann feststellen

$$v^h(x) = \sum_{i=1}^N \tilde{\varphi}_i(F_T^{-1}(x)).$$

Es sind folglich nur Funktionsauswertungen der Ansatzfunktionen $\tilde{\varphi}_i$ auf dem Referenzelement \tilde{T} nötig. Dies ist für den Implementierungsaufwand günstig, da statt der konkreten φ_i als Basis von V^h lediglich die „lokalen“ Ansatzfunktionen $\tilde{\varphi}_i$ auf dem Referenzelement bestimmt werden müssen.

Neben den Elementen der Zerlegung werden auch ihre Ränder relevant sein. Daher verwenden wir $\mu_k(M)$ für das k -dimensionale Lebesgue-Maß von M und bezeichnen mit

$$\mathcal{F} = \{\partial T_i \cap \partial T_j \mid \mu_{d-1}(\partial T_i \cap \partial T_j) > 0, i \neq j\}$$

die Menge der Facetten der Zerlegung und mit

$$\mathcal{F}_\partial = \{\partial T_i \cap \partial \Omega \mid \mu_{d-1}(\partial T_i \cap \partial \Omega) > 0\}$$

die Menge der Randfacetten der Zerlegung \mathcal{T} . Auch für sie wollen wir Abbildungen auf entsprechenden Referenzelementen \tilde{F} (Strecken für $d = 2$, Drei- oder Vierecke für $d = 3$) definieren:

$$F_F : \tilde{F} \rightarrow F \quad \text{affin-linear und bijektiv.}$$

4.3 Diskrete Abbildungen

In der variationellen Formulierung müssen die Ansatz- und Testfunktionen integriert werden. Aufgrund ihrer Konstruktion kann die Integration einer Funktion $g \in V^h$ wie folgt durchgeführt werden:

$$\int_{\Omega} g(x) dx = \sum_{T \in \mathcal{T}} \int_T g(x) dx = \sum_{T \in \mathcal{T}} \int_{\tilde{T}} g(F_T(\tilde{x})) |\det \nabla F_T(\tilde{x})| d\tilde{x} = \sum_{T \in \mathcal{T}} |\det A_T| \int_{\tilde{T}} \tilde{g}_T(\tilde{x}) d\tilde{x}$$

mit $\tilde{g}_T : \tilde{T} \rightarrow \mathbb{R}$, $\tilde{g}_T := g \circ F_T$. $f \circ g$ stellt hierbei die Verkettung der Funktionen dar: $(f \circ g)(x) = f(g(x))$.

Quadraturformeln werden folglich nur für \tilde{T} bzw. \tilde{F} benötigt:

$$\int_{\tilde{T}} \tilde{g}(\tilde{x}) d\tilde{x} \approx Q^{\tilde{T}}(\tilde{g}) = \sum_{\alpha=1}^{r^{\tilde{T}}} w_{\alpha}^{\tilde{T}} \tilde{g}(\tilde{x}_{\alpha}^{\tilde{T}}),$$

$$\int_{\tilde{F}} \tilde{g}(\tilde{x}) d\tilde{x} \approx Q^{\tilde{F}}(\tilde{g}) = \sum_{\alpha=1}^{r^{\tilde{F}}} w_{\alpha}^{\tilde{F}} \tilde{g}(\tilde{x}_{\alpha}^{\tilde{F}}).$$

Definition 4.3.1. Die Quadratur $Q^{\tilde{T}}$ bezeichnen wir als exakt auf einem Polynomraum $\mathbb{P}_s(\tilde{T})$, falls für alle $\tilde{p} \in \mathbb{P}_s(\tilde{T})$ gilt

$$Q^{\tilde{T}}(\tilde{p}) = \int_{\tilde{T}} \tilde{p}(\tilde{x}) d\tilde{x}.$$

Für $g \in C^0(T)$ wird mittels

$$Q^T(g) := \sum_{\alpha=1}^{r^{\tilde{T}}} w_{\alpha}^T g(x_{\alpha}^{\tilde{T}})$$

mit $x_{\alpha}^{\tilde{T}} = F_T(\tilde{x}_{\alpha}^{\tilde{T}})$ und $w_{\alpha}^T = |\det A_T| \tilde{w}_{\alpha}^{\tilde{T}}$ eine Quadratur für $T \in \mathcal{T}$ induziert.

Satz 4.3.1. 1. Ist $\tilde{Q}^{\tilde{T}}$ exakt auf $\mathbb{P}_s(\tilde{T})$, dann ist die induzierte Quadratur Q^T exakt auf $\mathbb{P}_s(T)$.

2. Ist \tilde{Q} exakt auf $\mathbb{P}_s(\tilde{\Omega})$ und sei $g \in C^{\infty}(\Omega)$, $Q(g) := \sum_{T \in \mathcal{T}} Q^T(g)$, so gilt

$$\left| \int_{\Omega} g(x) dx - Q(g) \right| \leq Ch^{s+1},$$

(mit z. B. $h := \max_{T \in \mathcal{T}} \text{diam}(T)$ und C einer von h unabhängigen Konstanten).

Beweis. Siehe etwa [OD06], Satz 1.11. □

Entsprechendes gilt natürlich auch für F bzw. \tilde{F} statt T bzw. \tilde{T} .

Kanonisch ergibt sich also für

$$\begin{aligned} a(u, v) &= \int_{\Omega} \alpha_{\text{vol}}(u, v)(x) dx + \int_{\partial\Omega} \alpha_{\text{bnd}}(u, v)(x) dx + \int_{\Gamma_{\text{int}}} \alpha_{\text{skel}}(u, v)(x) dx \\ &= \sum_{T \in \mathcal{T}} \int_T \alpha_{\text{vol}}(u, v)(x) dx + \sum_{F \in \mathcal{F}_{\partial}} \int_F \alpha_{\text{bnd}}(u, v)(x) dx + \sum_{F \in \mathcal{F}} \int_F \alpha_{\text{skel}}(u, v)(x) dx \\ &= \sum_{T \in \mathcal{T}} |\det A_T| \int_{\tilde{T}} \tilde{\alpha}_{\text{vol}}^T(u, v)(\tilde{x}) d\tilde{x} \\ &\quad + \sum_{F \in \mathcal{F}_{\partial}} |\det A_F| \int_{\tilde{F}} \tilde{\alpha}_{\text{bnd}}^F(u, v)(\tilde{x}) d\tilde{x} + \sum_{F \in \mathcal{F}} |\det A_F| \int_{\tilde{F}} \tilde{\alpha}_{\text{skel}}^F(u, v)(\tilde{x}) d\tilde{x} \end{aligned}$$

die Modellierung

$$\begin{aligned} a^h(u, v) &= \sum_{T \in \mathcal{T}} \sum_{\alpha=1}^{r^{\tilde{T}}} w_{\alpha}^T \tilde{\alpha}_{\text{vol}}^T(u, v)(\tilde{x}_{\alpha}^{\tilde{T}}) \\ &\quad + \sum_{F \in \mathcal{F}_{\partial}} \sum_{\alpha=1}^{r^{\tilde{F}}} w_{\alpha}^F \tilde{\alpha}_{\text{bnd}}^F(u, v)(\tilde{x}_{\alpha}^{\tilde{F}}) + \sum_{F \in \mathcal{F}} \sum_{\alpha=1}^{r^{\tilde{F}}} w_{\alpha}^F \tilde{\alpha}_{\text{skel}}^F(u, v)(\tilde{x}_{\alpha}^{\tilde{F}}), \end{aligned}$$

so dass wir hier nur den Quadraturfehler gemacht haben: Bei Verwendung von Quadraturen, die exakt auf \mathbb{P}_s sind, sowie $u, v \in C^{s+1}(\Omega)$ ergibt sich dann

$$\left| a(u, v) - a^h(u, v) \right| \leq Ch^{s+1}.$$

Es sei nun $\{\varphi_i \mid i \in \mathcal{I}\}$ eine Basis von V^h . \mathcal{I} ist dabei eine Indexmenge, etwa $\mathcal{I} \subset \mathbb{N}$. Dann können wir statt $u, v \in V$ Funktionen $u^h, v^h \in V^h$ aus dem diskreten Funktionsraum mit

$$u^h = \sum_{i \in \mathcal{I}} u_i^h \varphi_i \quad \text{und} \quad v^h = \sum_{i \in \mathcal{I}} v_i^h \varphi_i$$

verwenden und es ergibt sich

$$\begin{aligned} \tilde{\alpha}_{\text{vol}}^T \left(\sum_{i \in \mathcal{I}} u_i^h \varphi_i, \sum_{i \in \mathcal{I}} v_i^h \varphi_i \right) (\tilde{x}) &= \tilde{\alpha}_{\text{vol}}^T \left(\sum_{i \in \mathcal{I}_T} u_i^h \varphi_i, \sum_{i \in \mathcal{I}_T} v_i^h \varphi_i \right) (\tilde{x}) \\ &= \tilde{\alpha}_{\text{vol}}^{\tilde{T}} \left(\sum_{i \in \mathcal{I}_T} u_i^h \tilde{\varphi}_i, \sum_{i \in \mathcal{I}_T} v_i^h \tilde{\varphi}_i \right) (\tilde{x}). \end{aligned}$$

Dabei sind

$$\begin{aligned} \mathcal{I}_T &:= \{i \in \mathcal{I} \mid \text{int}(\text{supp } \varphi_i) \cap T \neq \emptyset\}, \\ \mathcal{I}_F &:= \{i \in \mathcal{I} \mid \text{supp } \varphi_i \cap F \neq \emptyset\} \end{aligned}$$

die Indizes jener Basisfunktionen von V^h , die für das Gebiet T bzw. die Facette F relevant sind (d. h. nicht verschwinden). $\text{int } M$ bezeichnet hierbei die inneren Punkte von M . Bei einer analogen Vorgehensweise für $\tilde{\alpha}_{\text{bnd}}^F$ bzw. $\tilde{\alpha}_{\text{skel}}^F$ ergibt sich

$$\begin{aligned} a^h(u^h, v^h) &= \sum_{T \in \mathcal{T}} \sum_{\alpha=1}^{r^{\tilde{T}}} w_\alpha^T \tilde{\alpha}_{\text{vol}}^{\tilde{T}} \left(\sum_{i \in \mathcal{I}_T} u_i^h \tilde{\varphi}_i, \sum_{i \in \mathcal{I}_T} v_i^h \tilde{\varphi}_i \right) (\tilde{x}_\alpha^{\tilde{T}}) \\ &\quad + \sum_{F \in \mathcal{F}} \sum_{\alpha=1}^{r^{\tilde{F}}} w_\alpha^F \tilde{\alpha}_{\text{bnd}}^{\tilde{F}} \left(\sum_{i \in \mathcal{I}_F} u_i^h \tilde{\varphi}_i, \sum_{i \in \mathcal{I}_F} v_i^h \tilde{\varphi}_i \right) (\tilde{x}_\alpha^{\tilde{F}}) \\ &\quad + \sum_{F \in \mathcal{F}_\partial} \sum_{\alpha=1}^{r^{\tilde{F}}} w_\alpha^F \tilde{\alpha}_{\text{skel}}^{\tilde{F}} \left(\sum_{i \in \mathcal{I}_F} u_i^h \tilde{\varphi}_i, \sum_{i \in \mathcal{I}_F} v_i^h \tilde{\varphi}_i \right) (\tilde{x}_\alpha^{\tilde{F}}). \end{aligned}$$

Zu Beachten ist, dass jetzt lediglich Auswertungen von $\alpha_{\text{vol}}^{\tilde{T}}, \alpha_{\text{bnd}}^{\tilde{F}}, \alpha_{\text{skel}}^{\tilde{F}}$ auf den Referenzgebieten \tilde{T} bzw. \tilde{F} nötig sind. In der Regel sind diese Funktionen linear im zweiten Argument, so dass gilt

$$\begin{aligned} a^h(u^h, v^h) &= \sum_{j \in \mathcal{I}_T} v_j^h \sum_{T \in \mathcal{T}} \sum_{\alpha=1}^{r^{\tilde{T}}} w_\alpha^T \tilde{\alpha}_{\text{vol}}^{\tilde{T}} \left(\sum_{i \in \mathcal{I}_T} u_i^h \tilde{\varphi}_i, \tilde{\varphi}_j \right) (\tilde{x}_\alpha^{\tilde{T}}) \\ &\quad + \sum_{j \in \mathcal{I}_F} v_j^h \sum_{\alpha=1}^{r^{\tilde{F}}} \left(\sum_{F \in \mathcal{F}} w_\alpha^F \tilde{\alpha}_{\text{bnd}}^{\tilde{F}} \left(\sum_{i \in \mathcal{I}_F} u_i^h \tilde{\varphi}_i, \tilde{\varphi}_j \right) (\tilde{x}_\alpha^{\tilde{F}}) \right. \\ &\quad \left. + \sum_{F \in \mathcal{F}_\partial} w_\alpha^F \tilde{\alpha}_{\text{skel}}^{\tilde{F}} \left(\sum_{i \in \mathcal{I}_F} u_i^h \tilde{\varphi}_i, \tilde{\varphi}_j \right) (\tilde{x}_\alpha^{\tilde{F}}) \right). \end{aligned}$$

Entsprechend bekommen wir für

$$\begin{aligned}
-\langle f, v \rangle &= \int_{\Omega} \lambda_{\text{vol}}(v)(x) dx + \int_{\partial\Omega} \lambda_{\text{bnd}}(v)(x) dx + \int_{\Gamma_{\text{int}}} \lambda_{\text{skel}}(v)(x) dx \\
&= \sum_{T \in \mathcal{T}} \int_T \lambda_{\text{vol}}(v)(x) dx + \sum_{F \in \mathcal{F}_{\partial}} \int_F \lambda_{\text{bnd}}(v)(x) dx + \sum_{F \in \mathcal{F}} \int_F \lambda_{\text{skel}}(v)(x) dx \\
&= \sum_{T \in \mathcal{T}} |\det A_T| \int_{\tilde{T}} \tilde{\lambda}_{\text{vol}}^T(v)(\tilde{x}) d\tilde{x} \\
&\quad + \sum_{F \in \mathcal{F}_{\partial}} |\det A_F| \int_{\tilde{F}} \tilde{\lambda}_{\text{bnd}}^F(v)(\tilde{x}) d\tilde{x} + \sum_{F \in \mathcal{F}} |\det A_F| \int_{\tilde{F}} \tilde{\lambda}_{\text{skel}}^F(v)(\tilde{x}) d\tilde{x}
\end{aligned}$$

die Modellierung

$$\begin{aligned}
-\langle f^h, v \rangle &= \sum_{T \in \mathcal{T}} \sum_{\alpha=1}^{r^{\tilde{T}}} w_{\alpha}^T \tilde{\lambda}_{\text{vol}}^T(v)(\tilde{x}_{\alpha}^{\tilde{T}}) \\
&\quad + \sum_{F \in \mathcal{F}_{\partial}} \sum_{\alpha=1}^{r^{\tilde{F}}} w_{\alpha}^F \tilde{\lambda}_{\text{bnd}}^F(v)(\tilde{x}_{\alpha}^{\tilde{F}}) + \sum_{F \in \mathcal{F}} \sum_{\alpha=1}^{r^{\tilde{F}}} w_{\alpha}^F \tilde{\lambda}_{\text{skel}}^F(v)(\tilde{x}_{\alpha}^{\tilde{F}}),
\end{aligned}$$

und schließlich

$$\begin{aligned}
-\langle f^h, v^h \rangle &= \sum_{j \in \mathcal{I}_T} v_i^h \sum_{T \in \mathcal{T}} \sum_{\alpha=1}^{r^{\tilde{T}}} w_{\alpha}^T \tilde{\lambda}_{\text{vol}}^T(\tilde{\varphi}_j)(\tilde{x}_{\alpha}^{\tilde{T}}) \\
&\quad + \sum_{j \in \mathcal{I}_F} v_i^h \sum_{\alpha=1}^{r^{\tilde{F}}} \left(\sum_{F \in \mathcal{F}} w_{\alpha}^F \tilde{\lambda}_{\text{bnd}}^F(\tilde{\varphi}_j)(\tilde{x}_{\alpha}^{\tilde{F}}) + \sum_{F \in \mathcal{F}_{\partial}} w_{\alpha}^F \tilde{\lambda}_{\text{skel}}^F(\tilde{\varphi}_j)(\tilde{x}_{\alpha}^{\tilde{F}}) \right).
\end{aligned}$$

Genau diese Zerlegungen und Umformungen werden so in dem Programmpaket Dune/PDELab [BBD⁺13b, BBD⁺13a] implementiert.

Um das diskrete Problem

$$\text{finde } u^h \in V^h \text{ mit } a^h(u^h, v^h) = \langle f^h, v^h \rangle \quad \forall v^h \in V^h$$

zu lösen, müssen lediglich

- eine Gitterzerlegung \mathcal{T} vorgegeben werden,
- zu verwendende Basisfunktionen auf \tilde{T} definiert bzw. aus den bereits implementierten ausgewählt werden,
- die lokalen Operatoren $\tilde{\alpha}_{\text{vol}}^{\tilde{T}}$, $\tilde{\alpha}_{\text{bnd}}^{\tilde{F}}$, $\tilde{\alpha}_{\text{skel}}^{\tilde{F}}$, $\tilde{\lambda}_{\text{vol}}^{\tilde{T}}$, $\tilde{\lambda}_{\text{bnd}}^{\tilde{F}}$, $\tilde{\lambda}_{\text{skel}}^{\tilde{F}}$ implementiert werden.

Der diskrete Funktionsraum V^h kann dann leicht konstruiert werden und gemäß der Umformungen von oben kann dann eine Lösung für

$$\begin{aligned}
a^h(u^h, v^h) - \langle f^h, v^h \rangle &= \sum_{T \in \mathcal{T}} \sum_{\alpha=1}^{r^{\tilde{T}}} w_{\alpha}^T \left(\tilde{\alpha}_{\text{vol}}^T(u^h, v^h) + \tilde{\lambda}_{\text{vol}}^T(v^h) \right) (\tilde{x}_{\alpha}^{\tilde{T}}) \\
&+ \sum_{F \in \mathcal{F}_{\partial}} \sum_{\alpha=1}^{r^{\tilde{F}}} w_{\alpha}^F \left(\tilde{\alpha}_{\text{bnd}}^F(u^h, v^h) + \tilde{\lambda}_{\text{bnd}}^F(v^h) \right) (\tilde{x}_{\alpha}^{\tilde{F}}) \\
&+ \sum_{F \in \mathcal{F}} \sum_{\alpha=1}^{r^{\tilde{F}}} w_{\alpha}^F \left(\tilde{\alpha}_{\text{skel}}^F(u^h, v^h) + \tilde{\lambda}_{\text{skel}}^F(v^h) \right) (\tilde{x}_{\alpha}^{\tilde{F}}) = 0
\end{aligned} \tag{4.2}$$

gesucht werden.

Wir bezeichnen mit

$$(a^h)'(u; w, v) := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left[a^h(u + \varepsilon w; v) - a^h(u; v) \right].$$

die Richtungsableitung von a^h nach dem ersten Argument. Da es sich bei V^h um endlichdimensionale Funktionsräume handelt, stellen wir $(u^h)^{(k)} \in V^h$ dar als

$$(u^h)^{(k)} = \sum_{i=1}^N (u^h)_i^{(k)} \varphi_i.$$

Es kann dann versucht werden, das Problem (4.2) mit dem Gleichungssystem

$$\sum_{j=1 \dots N} (a^h)' \left((u^h)^{(k)}; \varphi_j, \varphi_i \right) s_j^{(k)} = \langle f^h, \varphi_i \rangle \quad i = 1, \dots, N \tag{4.3}$$

und der Korrektur

$$(u^h)^{(k+1)} := (u^h)^{(k)} - s^{(k)} \tag{4.4}$$

gemäß Newton-Verfahren iterativ zu lösen. Falls a^h linear im ersten Argument ist, führt dies bereits nach einer Iteration und beliebigem Startwert $u^{h,(0)}$ zur Lösung. Ansonsten können diese Schritte für $k = 0, 1, \dots$ iteriert werden. Die Konvergenz des Verfahrens hängt im Wesentlichen von der Gutartigkeit der Nichtlinearität sowie des Startwertes ab.

Durch die Art der Diskretisierung ist die Koeffizientenmatrix („Steifigkeitsmatrix“) im linearen Gleichungssystem (4.3) schwach besetzt, so dass hier auch iterative (parallelisierbare) Löser verwendet werden können. Bei starker Nichtlinearität ist es ggf. ratsam, beim Updateschritt (4.4) eine Dämpfung einzubauen.

4.4 Zeitdiskretisierung

Im vorangegangenen Abschnitt wurde die Diskretisierung stationärer elliptischer partieller Differentialgleichungen vorgestellt. Wir können sie aber auch dazu verwenden, die instationären

parabolischen Probleme zu diskretisieren.

$$\text{für } t \in I \text{ finde } u(t) \in V \text{ mit } \partial_t \underbrace{(u(t), v)}_{=:m(u(t),v,t)} + \underbrace{a(u(t), v) - \langle f, v \rangle}_{=:r(u(t),v,t)} = 0 \quad \forall v \in V$$

m kann hier genauso wie zuvor bereits r mithilfe lokaler Operatoren zu m^h diskretisiert werden, einzige Änderung ist die mögliche Abhängigkeit von der hier als Parameter zu betrachtenden Zeit t .

Zur Zeitdiskretisierung unterteilen wir zunächst das Zeitintervall in diskrete Zeitpunkte

$$0 = t_0 < t_1 < \dots < t_N = T.$$

Die Idee ist es nun, ∂_t so zu diskretisieren, dass wir insbesondere für jeden Zeitschritt t_n eine Lösung u_n^h bestimmen. Dazu gibt es verbreitete Verfahren, etwa die s -stufigen Einschrittmethoden, siehe Algorithmus 4.1.

Algorithm 4.1 Ein Schritt einer allgemeinen s -stufigen Einschrittmethode

Input: $t_n, t_{n+1} \in \mathbb{R}, u_n^h \in V^h, A = \begin{pmatrix} a_{10} & \dots & a_{1s} \\ \vdots & & \vdots \\ a_{s0} & \dots & a_{ss} \end{pmatrix}, B = \begin{pmatrix} b_{10} & \dots & b_{1s} \\ \vdots & & \vdots \\ b_{s0} & \dots & b_{ss} \end{pmatrix} \in \mathbb{R}^{s \times (s+1)},$

$$d = (d_0, \dots, d_s)^\top \in \mathbb{R}^{s+1}$$

Output: $u_{n+1}^h \in V^h$

$$u_{(0)}^h \leftarrow u_n^h$$

$$k \leftarrow t_{n+1} - t_n$$

for $i = 1, \dots, s$ **do**

Finde $u_{(i)}^h \in V^h$ als Approximation für $u(t_n + kd_i)$ mit

$$\sum_{j=0}^s \left(a_{ij} m^h(u_{(j)}^h, v, t_n + kd_j) + b_{ij} r^h(u_{(j)}^h, v, t_n + kd_j) \right) = 0 \quad \forall v \in V^h$$

end for

$$u_{n+1}^h \leftarrow u_{(s)}^h$$

Als stark S-stabiles Verfahren zweiter Ordnung wird etwa das Verfahren von Alexander [Ale77] verwendet. Es hat die Parameter

$$A = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \alpha & 0 \\ 0 & 1 - \alpha & \alpha \end{pmatrix}, \quad d = (0, \alpha, 1)^\top$$

mit $\alpha := 1 - \sqrt{2}/2$.

Eine weitere Klasse gut geeigneter Zeitdiskretisierungen sind die sogenannten Rosenbrock-Methoden. Die Klasse der s -stufigen Rosenbrock-Verfahren zur Zeitdiskretisierung der impliziten Differentialgleichung

$$\begin{aligned} M\dot{u} &= f(t, u), \\ u(t_0) &= u_0 \end{aligned}$$

sieht wie folgt aus:

$$Mk_i := \tau f(t_n + \alpha_i \tau, u_n + \sum_{j=1}^{i-1} \alpha_{ij} k_j) + \tau f_u(t_n, u_n) \sum_{j=1}^i \gamma_{ij} k_j + \tau^2 \gamma_i f_t(t_n, u_n), \quad i = 1, \dots, s,$$

$$u_{n+1} := u_n + \sum_{i=1}^s b_i k_i,$$

wobei $\tau := t_{n+1} - t_n$ die Schrittweite, α_{ij} , γ_{ij} sowie b_i die Parameter des Verfahrens und $\alpha_i := \sum_{j=1}^{i-1} \alpha_{ij}$, $\gamma_i := \sum_{j=1}^{i-1} \gamma_{ij}$ Abkürzungen sind.

In [RA08], Beispiel 4 wurde experimentell gezeigt, dass die dort vorgestellten Rosenbrock-Methoden zur Lösung der Navier-Stokes-Gleichungen gut geeignet sind. Insbesondere liefern die dort vorgestellten (und aufwendigeren) Rosenbrock-Methoden qualitativ bessere Ergebnisse als die einfacheren klassischen Methoden wie das implizite Euler-Verfahren, das Crank-Nicolson-Schema oder das θ -Verfahren.

Für eine genauere Beschreibung der Bedingungen an die Parameter sowie verschiedene konkrete Wahlen sei an dieser Stelle lediglich auf die entsprechenden Abschnitte von [RA08] verwiesen.

4.5 Diskretisierung des gemischten Problems

Um (3.11) zu diskretisieren, müssen zunächst die Funktionsräume V bzw. Q diskretisiert werden. Die Diskretisierung erfolgt also für zwei Funktionsräume, oder anders betrachtet handelt es sich um zwei Diskretisierungen. Wir werden mit V^h den diskreten Funktionsraum zur Diskretisierung von V und zur Darstellung des Geschwindigkeitsfeldes bezeichnen, und Q^h den diskreten Funktionsraum zur Diskretisierung von Q und zur Darstellung der Druckverteilung. Die variationelle Formulierung des gemischten Problems erfolgt dann auf dem zusammengesetzten Funktionsraum

$$\hat{V}^h = V^h \times Q^h.$$

Wir wollen voraussetzen, dass wir konforme Finite Elemente verwenden, d. h. $V^h \subset V$ und $Q^h \subset Q$. Dies hat den Vorteil, dass direkt eine kanonische Definition von \hat{a}^h bzw. \hat{f}^h aus \hat{a} bzw. \hat{f} möglich ist.

Bei der Kombination der Funktionsräume V^h und Q^h ist zu beachten, dass auch hier die Inf-Sup-Bedingung (3.2) (nun für die diskreten Räume V^h, Q^h) erfüllt sein soll.

Eine weit verbreitete Wahl des Funktionsraumpaars ist die der Taylor-Hood-Elemente der Ordnung $k \in \mathbb{N}$ [TH73]:

$$V_{\text{TH},k}^h := (\mathcal{S}_{n+1,n}^{\mathcal{T}})^d \cap H_0^1(\Omega)^d,$$

$$Q_{\text{TH},k}^h := \mathcal{S}_{n,0}^{\mathcal{T}} \cap L_0^2(\Omega),$$

$$\hat{V}_{\text{TH},k}^h := V_{\text{TH},k}^h \times Q_{\text{TH},k}^h.$$

Haben wir etwa eine Gitterzerlegung aus Dreiecken ($d = 2$) bzw. Tetraedern ($d = 3$), dann sprechen wir auch von P_2/P_1 Elementen bei $\hat{V}^h = \hat{V}_{\text{TH},1}^h$ bzw. den klassischen Taylor-Hood-

Elementen. Anhand der Definition der Taylor-Hood-Elemente wäre es denkbar, dass die noch einfachere Elementpaarung P_1/P_0 ebenfalls geeignet ist – jedoch wird von ihnen die diskrete Inf-Sup-Bedingung verletzt. Eine Vielzahl von Arbeiten hat bereits mögliche Elementpaarungen untersucht, an dieser Stelle wollen wir dies jedoch nicht vertiefen, sondern uns mit den klassischen Taylor-Hood-Elementen begnügen.

Anschließend implementieren wir die lokalen Operatoren für $\hat{u}^h = (\tilde{u}^h, \tilde{p}^h)$ und $\hat{v}^h = (\tilde{v}^h, \tilde{q}^h)$ gemäß

$$\begin{aligned}\hat{\alpha}_{\text{vol}}^T(\hat{u}^h, \hat{v}^h)(\tilde{x}) &:= 2\nu \nabla^s \tilde{u}^h(\tilde{x}) : \nabla \tilde{v}^h(\tilde{x}) + \nabla \cdot \tilde{v}^h(\tilde{x}) \tilde{p}^h(\tilde{x}) + \nabla \cdot \tilde{u}^h(\tilde{x}) \tilde{q}^h(\tilde{x}) \\ &\quad + \frac{1}{2} \left(\tilde{u}^h(\tilde{x}) \cdot \nabla \tilde{u}^h(\tilde{x}) \cdot \tilde{v}^h(\tilde{x}) - \tilde{u}^h(\tilde{x}) \cdot \nabla \tilde{v}^h(\tilde{x}) \cdot \tilde{u}^h(\tilde{x}) \right), \\ \hat{\alpha}_{\text{bnd}}^F(\hat{u}^h, \hat{v}^h)(\tilde{x}, n) &:= \frac{1}{2} (\tilde{u}^h(\tilde{x}) \cdot \tilde{v}^h(\tilde{x})) (\tilde{u}^h(\tilde{x}) \cdot n) \\ &\quad - \frac{1}{2} \nu n \nabla \tilde{u}^h(\tilde{x}) \tilde{v}^h(\tilde{x}) - \frac{1}{2} \nu \tilde{v}^h(\tilde{x}) \nabla \tilde{u}^h(\tilde{x}) n, \\ \hat{\lambda}_{\text{vol}}^T(\hat{v}^h)(\tilde{x}) &:= -\tilde{f}^h(\tilde{x}) \cdot \tilde{v}^h(\tilde{x}) \\ \hat{\lambda}_{\text{bnd}}^T(\hat{v}^h)(\tilde{x}, n) &:= -\tilde{j}_N^h(\tilde{x}) \cdot n \tilde{q}^h(\tilde{x}).\end{aligned}$$

Dabei ist stets für das aktuelle Element $T \in \mathcal{T}$

$$\tilde{u}^h = u^h \circ F_T, \quad \tilde{p}^h = p^h \circ F_T, \quad \tilde{v}^h = v^h \circ F_T, \quad \tilde{q}^h = q^h \circ F_T$$

bzw. für das aktuelle $F \in \mathcal{F}$:

$$\tilde{u}^h = u^h \circ F_F, \quad \tilde{p}^h = p^h \circ F_F, \quad \tilde{v}^h = v^h \circ F_F, \quad \tilde{q}^h = q^h \circ F_F.$$

4.6 Turbulenz und Multiskalen

Ein wesentliches Problem sind turbulente Strömungen. Sie werden charakterisiert durch eine kleine Viskosität $\nu \ll 1$. Es treten dabei Verwirbelungen in der Strömung auf, die von sehr unterschiedlicher Größe sind.

Die kleinen Verwirbelungen aufzulösen erfordert eine feine Diskretisierung. Mit ihr zu rechnen wiederum bedeutet einen Mehraufwand an Rechenzeit und Speicherkapazität. Unter diesen praktischen Gesichtspunkten erscheint eine lediglich grobskalige Diskretisierung wünschenswert.

Jedoch transportieren auch die kleineren Verwirbelungen Energie und tragen zum Entstehen größerer Wirbel bei oder verändern das Strömungsbild. Selbst wenn die kleinskaligen Verwirbelungen also für den Anwender unwichtig erscheinen, können sie oftmals nicht einfach ignoriert werden, sofern ein akkuraten Gesamtergebnis von Relevanz ist.

Ein vielversprechender Ausweg aus dem Dilemma ist die Idee, die feinen Verwirbelungen nicht zu berechnen, sondern nur ihren Effekt auf die (berechneten) größeren Wirbel („Large Eddies“) zu modellieren. Dieser Ansatz wird entsprechend *Large Eddy Simulation* genannt. Details dazu sind in [Sag02] zu finden.

Ein weiterer (noch gitterorientierterer) Ansatz ist die sogenannte *Variationelle Multiskalenmethode*. Die Effekte der (nicht berechneten) feinen Skalen werden durch ein Ersatzmodell nachgebildet. Dieses Ersatzmodell hat häufig die Form einer zusätzlichen Viskosität, etwa

$$\hat{a}_m(\hat{u}, \hat{v}) = (\nu_T \nabla^s u, \nabla v).$$

Die künstliche Viskosität ν_T ist dabei in der Regel nicht konstant, sondern kann insbesondere von der Lösung \hat{u} abhängen, siehe etwa [MP94, Joh06]. Für seine genaue Form gibt es aber viele verschiedene konkurrierende Ansätze. Ein Überblick über die geläufigen Techniken ist in [SDT06, Rem10] zu finden.

5 A-posteriori-Abschätzung der Navier-Stokes-Gleichung sowie des dualen Problems zur Modellgewinnung

Durch die Diskretisierung sowie der konkreten Berechnung werden Fehler gemacht, sodass wir nur eine Approximation \hat{u}^h an die exakte Lösung \hat{u} bekommen. Es gibt also einen Fehler zwischen der tatsächlich berechneten und der eigentlich gesuchten Lösung. Dieser Fehler ist prinzipbedingt und kann nicht eliminiert werden. Jedoch kann der Versuch unternommen werden, die Modellierung so zu gestalten, dass der Fehler in einer gewissen Hinsicht klein ist.

Konkret wird folgende Herangehensweise verwendet: Die Lösung $\hat{u} = (u, p) \in V \times Q$ etwa der inkompressiblen Navier-Stokes-Gleichung beinhaltet das gesamte Strömungsverhalten und die Druckverteilung im Berechnungsgebiet Ω . Häufig ist der Anwender aber vor allem an einer von dieser Lösung abgeleiteten Kenngröße interessiert, z. B. dem Strömungswiderstand oder dem Auftrieb. Wir können sie etwa mithilfe des Funktional

$$\hat{j}: \hat{V} \rightarrow \mathbb{R}$$

beschreiben. Zum Beispiel kann mit

$$\hat{j}(\hat{u}) = \int_{\Gamma_j} n \cdot (\nu \nabla^s u - p \mathbb{I}) \psi$$

die durch die Strömung $\hat{u} = (u, p)$ verursachte Kraft in Richtung $\psi \in \mathbb{R}^d$ auf ein vom Rand $\Gamma_j \subset \Gamma$ umschlossenes Objekt beschrieben werden, vgl. [HJ10]. Die daraus resultierenden skalaren Größen können – je nach Wahl von ψ – typische physikalische Kenngrößen sein:

- Zeigt ψ in Richtung der Hauptströmung in der Röhre, so beschreibt \hat{j} den *Strömungswiderstand* des Hindernisses.
- Zeigt ψ „nach oben“, so beschreibt \hat{j} den *Auftrieb* desselben Objekts.

Die Idee ist jetzt, dass wir bei der Approximation $\hat{u}^h \approx \hat{u}$ berücksichtigen, dass insbesondere der Fehler in dieser abgeleiteten Größe klein ist, also

$$\left| \hat{j}(\hat{u}^h) - \hat{j}(\hat{u}) \right| < \varepsilon.$$

Anders gesagt: Wir lassen Fehler zu, die diese Zielgröße nicht so stark verändern. Dies kann sich positiv auf den Berechnungsaufwand auswirken.

5.1 ε -schwache Lösungen

Um verschiedene Approximationen an die Lösung des eigentlichen Problems diskutieren zu können, wollen wir zunächst die Problemformulierung abschwächen. Das bislang betrachtete Problem

$$\hat{a}(\hat{u}, \hat{v}) = \langle \hat{f}, \hat{v} \rangle \quad \forall \hat{v} \in \hat{V}$$

schreiben wir nun mithilfe von

$$\hat{R}(\hat{u}, \hat{v}) := \hat{a}(\hat{u}, \hat{v}) - \langle \hat{f}, \hat{v} \rangle \quad \forall \hat{u}, \hat{v} \in \hat{V}$$

als

$$\hat{R}(\hat{u}, \hat{v}) = 0 \quad \forall \hat{v} \in \hat{V}.$$

Dann bezeichnen wir \hat{u} als ε -schwache Lösung des Problems, wenn gilt

$$|\hat{R}(\hat{u}, \hat{v})| \leq \varepsilon \|\hat{v}\| \quad \forall \hat{v} \in \hat{V}.$$

5.2 Das duale Problem

Grundlage der a-posteriori-Abschätzung ist das zu dem bisher betrachteten (nachfolgend „primal“ bezeichnet) duale Problem.

Das (nichtlineare) primale Problem

$$\hat{a}(\hat{u}, \hat{v}) = \langle \hat{f}, \hat{v} \rangle \quad \forall \hat{v} \in \hat{V}$$

verwendet die Form $\hat{a}: \hat{V} \times \hat{V} \rightarrow \mathbb{R}$, die im zweiten Argument linear ist, im Ersten jedoch nichtlinear sein kann. Wie bereits zuvor eingeführt bezeichnet

$$\hat{a}'(\hat{u}; \hat{w}, \hat{v}) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [\hat{a}(\hat{u} + \varepsilon \hat{w}; \hat{v}) - \hat{a}(\hat{u}; \hat{v})]$$

die Richtungsableitung von a nach dem ersten Argument. Es ist $\hat{a}': \hat{V} \times \hat{V} \times \hat{V} \rightarrow \mathbb{R}$ ebenfalls nichtlinear im ersten Argument, aber linear in den anderen beiden Argumenten.

Ebenso wollen wir auch für das Funktional $\hat{j}: \hat{V} \rightarrow \mathbb{R}$ die Richtungsableitung

$$\hat{j}'(\hat{u}; \hat{w}) := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [\hat{j}(\hat{u} + \varepsilon \hat{w}) - \hat{j}(\hat{u})].$$

bestimmen. Sie wird verwendet, um das duale Problem zu formulieren: Finde $\hat{z} \in \hat{V}$, so dass

$$\hat{a}'(\hat{u}; \hat{v}, \hat{z}) = \hat{j}'(\hat{u}; \hat{v}) \quad \forall \hat{v} \in \hat{V},$$

wobei $\hat{u} \in \hat{V}$ als fester Parameter zu betrachten ist. Hervorzuheben ist, dass das duale Problem für jeden Parameter $\hat{u} \in \hat{V}$ linear ist.

Dann ergibt sich für ε -schwache Lösungen \hat{u} und \hat{w} des primalen Problems:

$$\hat{j}(\hat{w}) - \hat{j}(\hat{u}) \approx \hat{j}'(\hat{u}; \hat{w} - \hat{u}) = \hat{a}'(\hat{u}; \hat{w} - \hat{u}, \hat{z}) \approx \hat{a}(\hat{w}, \hat{z}) - \hat{a}(\hat{u}, \hat{z}).$$

Die Lösung $\hat{z} \in \hat{V}$ des dualen Problems wird auch als „Einflussfunktion“ für das Funktional \hat{j} bezeichnet.

Da \hat{u}, \hat{w} ε -schwache Lösungen sind, gilt insbesondere

$$|\hat{a}(\hat{w}, \hat{z}) - \hat{a}(\hat{u}, \hat{z})| = \left| \hat{R}(\hat{w}, \hat{z}) - \hat{R}(\hat{u}, \hat{z}) \right| \leq 2\varepsilon \|\hat{z}\|.$$

Je nach Art der Nichtlinearität können die Abschätzungen sogar noch zu Gleichungen verbessert werden. Das gelingt bei den Navier-Stokes-Gleichungen, indem als Parameter für die linke Seite des dualen Problems beide ε -schwachen Lösungen \hat{u} und \hat{w} verwendet werden:

Mit

$$\hat{a}'(\hat{u}, \hat{w}; \hat{v}, \hat{z}) := (u \cdot \nabla v, z) + (v \cdot \nabla w, z) + (\nabla \cdot z, q) + (\nabla \cdot v, s) + (\nu \nabla v, \nabla z)$$

(vgl. [HJ10]) bekommen wir als duales Problem

$$\hat{z} \in \hat{V} : \quad \hat{a}'(\hat{u}, \hat{w}; \hat{v}, \hat{z}) = \hat{j}'(\hat{u}; \hat{v}) \quad \forall \hat{v} \in \hat{V}$$

und den Zusammenhang

$$\hat{j}(\hat{w}) - \hat{j}(\hat{u}) \approx \hat{j}'(\hat{u}; \hat{w} - \hat{u}) = \hat{a}'(\hat{u}, \hat{w}; \hat{w} - \hat{u}, \hat{z}) = \hat{a}(\hat{w}, \hat{z}) - \hat{a}(\hat{u}, \hat{z}).$$

Sofern das Funktional \hat{j} linear ist, ist sogar die erste Approximation eine Gleichung. Insgesamt erhalten wir dann den Zusammenhang

$$\left| \hat{j}(\hat{w}) - \hat{j}(\hat{u}) \right| = |\hat{a}(\hat{w}, \hat{z}) - \hat{a}(\hat{u}, \hat{z})| \leq 2\varepsilon \|\hat{z}\|. \quad (5.1)$$

Mithilfe der Lösung \hat{z} des dualen Problems haben wir also eine Größe parat um den maximalen Fehler in dem Funktional \hat{j} abzuschätzen. Zu beachten ist, dass das duale Problem – auch bei nichtlinearem primalen Problem – linear und daher im Vergleich wesentlich einfacher zu lösen ist.

5.3 Modellierungskorrektur durch a-posteriori-Abschätzung

In diesem Abschnitt soll zunächst auf die gemischte Formulierung mit $\hat{u} = (u, p)$ verzichtet werden, sondern zwecks Übersichtlichkeit von einem einfacheren Problem ausgegangen werden. Die Übertragung der Idee auf unsere konkrete Anwendung erfolgt dann im anschließenden Abschnitt.

Im Folgenden wollen wir die Betrachtungen für ein Problem anstellen, bei dem die linke Seite zur konkreten Rechnung vereinfacht wird. Das kann durch Modellierung oder vereinfachende Implementierung geschehen. Grundlagen dieser Überlegungen sind etwa bei [BE03] oder [Ang13] zu finden.

Das zu betrachtende nichtlineare Problem wird beschrieben durch

$$u \in V : \quad a(u; v) = \langle f, v \rangle \quad \forall v \in V.$$

Hierbei ist f ein gegebenes lineares Funktional. Die Form $a: V \times V \rightarrow \mathbb{R}$ sei auch hier linear im zweiten Argument.

Nun sei zusätzlich ein nichtlineares Funktional $j: V \rightarrow \mathbb{R}$ gegeben, mit dem wir das duale Problem wie folgt formulieren:

$$z \in V : \quad a'(u; w, z) = j'(u; w) \quad \forall w \in V.$$

Zur konkreten Rechnung wird V durch verschiedene diskrete Räume V^H, V^h angenähert, wobei V^h die bessere (feinere) Approximation sein soll.

Beispiel 5.3.1. Als einfaches Beispiel soll zunächst eine gewöhnliche Differentialgleichung mit Randbedingungen betrachtet werden:

$$\begin{aligned} u''(x) + \pi^2 u(x) &= 0 \quad \forall x \in (-1, 1) \\ u(\pm 1) &= -1 \end{aligned}$$

Dieses Problem ist analytisch eindeutig zu lösen, es gilt:

$$u(x) = \cos(\pi x)$$

Bei der approximierenden Diskretisierung soll der Funktionswert bei $x = 0$ möglichst korrekt getroffen werden. Dazu wählen wir das Funktional j als

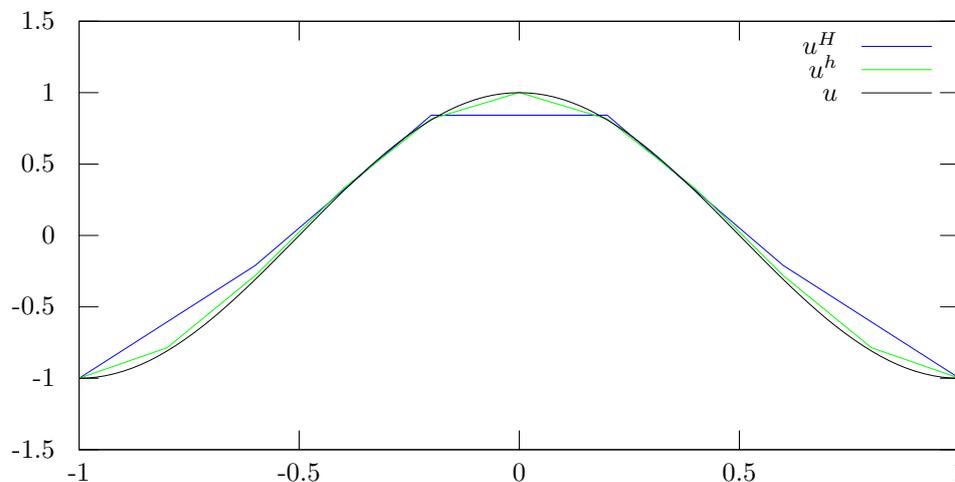
$$j(u) := u(0).$$

Wir diskretisieren nun das Problem mithilfe der Finite-Elemente-Methode (FEM) mit stückweise linearen Ansatzfunktionen auf einem gleichabständigen Gitter.

- u^H bezeichne die FEM-Lösung bei 4 Freiheitsgraden (Gitterstellen $\pm 0.6, \pm 0.2$).
- u^h bezeichne die FEM-Lösung bei 9 Freiheitsgraden (Gitterstellen $\pm 0.8, \pm 0.6, \pm 0.4, \pm 0.2, 0$).

Der Ansatzraum für die feinskalige Lösung wird also durch einen Gitterverfeinerungsschritt (h -Refinement) erreicht.

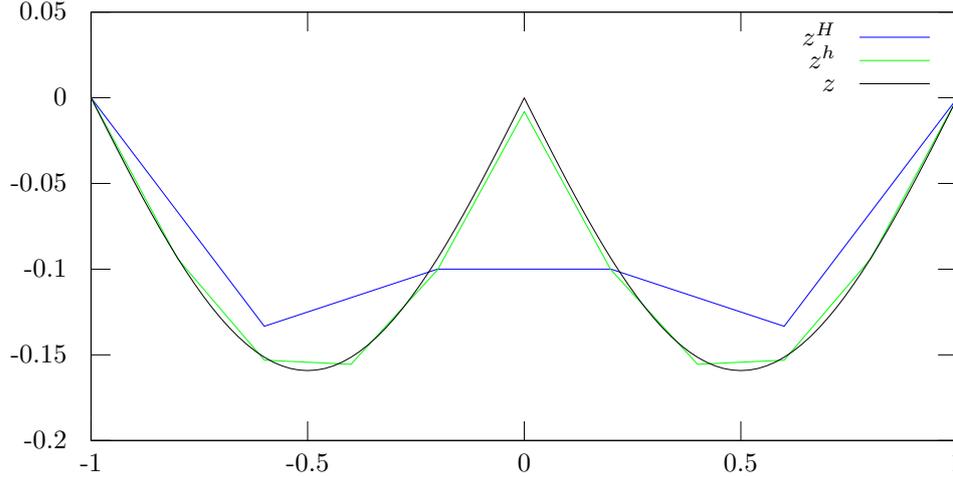
Die Lösungen der primalen Probleme sind dann:



Wir stellen fest, dass es bei der grobskaligen Approximation gegenüber der exakten Lösung eine deutliche Differenz im Funktional j gibt, wohingegen die feinskalige Lösung hier bereits eine ziemlich gute Approximation liefert.

$$j(u) = 1, \quad j(u^h) = 1.0012, \quad j(u^H) = 0.84206$$

Die Lösungen der zugeordneten dualen Probleme sind dann:



Als Zusammenhang (5.1) ergibt sich hier

$$\left| j(u^h) - j(u^H) \right| \approx 0.16 \leq \left| a(u^h, z^h) - a(u^H, z^h) \right| \approx 0.16 \leq 2\varepsilon \left\| z^h \right\| \approx 0.72\varepsilon.$$

Dies verdeutlicht bereits an dieser Stelle, dass verschiedene Diskretisierungen zu unterschiedlichen Funktionalwerten führen.

Im Unterschied zu diesem Beispiel kommt häufig in der Diskretisierung nicht dieselbe Form $a: V \times V \rightarrow \mathbb{R}$ wie bei dem kontinuierlichen Problem zum Einsatz, sondern eine modellierungsbehaftete modifizierte Version.

Dazu verwenden wir noch die Form $a_m: V \times V \rightarrow \mathbb{R}$, die ebenfalls linear im zweiten Argument, aber möglicherweise nichtlinear im ersten Argument ist. a_m beschreibt eine Modellierung, die zusätzlich bei einer (vereinfachten) Rechnung verwendet wird. Entsprechend bestimmen wir nun die primale Lösung $u_m^H \in V^H$ und die duale Lösung $z_m^H \in V^H$ des vereinfachten Problems im diskreten Raum V^H

$$u_m^H \in V^H : \quad a(u_m^H; v) + a_m(u_m^H; v) = \langle f, v \rangle \quad \forall v \in V^H, \quad (5.2)$$

$$z_m^H \in V^H : \quad a'(u_m^H; w, z_m^H) + a'_m(u_m^H; v) = j'(u_m^H; w) \quad \forall w \in V^H, \quad (5.3)$$

sowie die primale Lösung $u^h \in V^h$ und die duale Lösung $z^h \in V^h$ des Ausgangsproblems im diskreten Raum V^h :

$$\begin{aligned} u^h \in V^h : & \quad a(u^h; v) = \langle f, v \rangle \quad \forall v \in V^h, \\ z^h \in V^h : & \quad a'(u^h; w, z^h) = j'(u^h; w) \quad \forall w \in V^h. \end{aligned}$$

Diese variationellen Probleme sollen nun als Optimierungsprobleme formuliert werden. Das Tupel der primalen sowie dualen Lösung wollen wir im Folgenden mit

$$x_m^H = (u_m^H, z_m^H) \in X^H := V^H \times V^H$$

bzw.

$$x^h = (u^h, z^h) \in X^h := V^h \times V^h$$

bezeichnen. Wir definieren die folgenden Funktionale im Raum X :

$$\begin{aligned} L(x) &:= j(u) + \langle f, z \rangle - a(u, z), \\ L_m(x) &:= L(x) - a_m(u, z). \end{aligned}$$

Die Ableitung von L bzw. L_m auf eine Testfunktion $y = (w, v)$ angewendet ergibt

$$\begin{aligned} L'(x; y) &= j'(u; w) - a'(u; w, z) + \langle f, v \rangle - a(u, v), \\ L'_m(x; y) &= j'(u; w) - a'(u; w, z) - a'_m(u; w, z) + \langle f, v \rangle - a(u, v) - a_m(u, v) \\ &= L'(x; y) - (a'_m(u; w, z) + a_m(u, v)). \end{aligned}$$

Die gesuchten Probleme werden also offensichtlich durch kritische Punkte $x = (u, z)$, $x_m = (u_m, z_m)$ von L bzw. L_m gelöst:

$$\begin{aligned} x \in X &: & L'(x; y) &= 0 \quad \forall y \in X, \\ x_m \in X &: & L'_m(x_m; y) &= 0 \quad \forall y \in X, \\ x^h \in X^h &: & L'(x^h; y) &= 0 \quad \forall y \in X^h, \\ x_m^H \in X^H &: & L'_m(x_m^H; y) &= 0 \quad \forall y \in X^H. \end{aligned}$$

Außerdem haben wir den Zusammenhang

$$\begin{aligned} j(u^h) &= L(x^h), \\ j(u_m^H) &= L_m(x_m^H). \end{aligned}$$

Wir definieren zusätzlich noch die Residuen

$$\begin{aligned} \rho(u_m^H; v) &:= \langle f, v \rangle - a(u_m^H, v), \\ \rho^*(u_m^H; z_m^H, w) &:= j'(u_m^H; w) - a'(u_m^H; w, z_m^H) \end{aligned}$$

und stellen fest, dass

$$\begin{aligned} L'(x; y) &= \rho^*(u; z, w) + \rho(u, v), \\ L'_m(x; y) &= \rho^*(u; z, w) - a'_m(u; w, z) + \rho(u, v) - a_m(u, v). \end{aligned}$$

Satz 5.3.1. *Sofern a , a_m sowie j hinreichend differenzierbar (im ersten Argument) sind und $X^H \subseteq X$, dann gilt*

$$\begin{aligned} j(u) - j(u_m^H) &= a_m(u_m^H, z_m^H) \\ &+ \frac{1}{2} [\rho^*(u_m^H; z_m^H, e_u) + \rho(u_m^H, e_z)] - \frac{1}{2} R, \end{aligned}$$

wobei $e := (e_u, e_z) := (u - u_m^H, z - z_m^H)$ und

$$R := \int_0^1 \sigma(1 - \sigma)L'''(x_m^H + \sigma e; e, e, e)d\sigma.$$

Beweis.

$$\begin{aligned} j(u) - j(u_m^H) &= L(x) - L_m(x_m^H) \\ &= a_m(u_m^H; z_m^H) + \int_0^1 L'(x_m^H + \sigma e; e)d\sigma. \end{aligned}$$

Durch Anwendung der Trapezregel erhalten wir

$$\int_0^1 L'(x_m^H + \sigma e; e)d\sigma = \frac{1}{2} [L'(x; e) + L'(x_m^H; e) - R]$$

Da $X^H \subseteq X^h$ und folglich $e \in X^h$ ist schließlich

$$L'(x; e) + L'(x_m^H; e) = L'(x_m^H; e) = \rho(u_m^H; e_z) + \rho^*(u_m^H; z_m^H, e_u).$$

□

Es folgt:

$$\begin{aligned} j(u) - j(u_m^H) &\approx \eta^H + \eta^m, \\ \eta^H &:= \frac{1}{2} [\rho(u_m^H; e_z) + \rho^*(u_m^H; z_m^H, e_u)] \\ &= \frac{1}{2} [\langle f, z - z_m^H \rangle - a(u_m^H, z - z_m^H) \\ &\quad + j'(u_m^H; u^h - u_m^H) - a'(u_m^H; u^h - u_m^H, z_m^H)], \\ \eta^m &:= a_m(u_m^H, z_m^H). \end{aligned} \tag{5.4}$$

Beispiel 5.3.2. (Fortführung von Beispiel 5.3.1) Statt

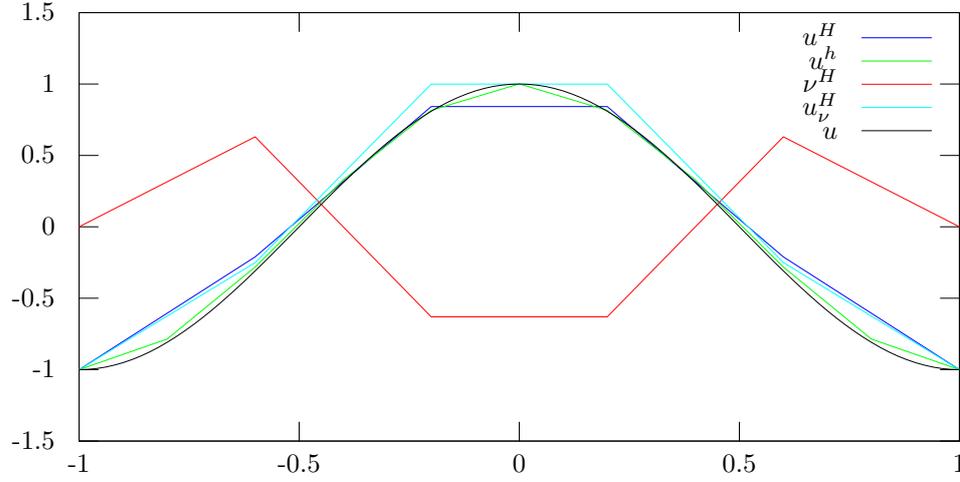
$$\int_{-1}^1 [u^{H'} v' - \pi^2 u^H v] = 0 \quad \forall v \in V^H,$$

der gewöhnlichen FEM-Formulierung lösen wir

$$\int_{-1}^1 [(1 + \nu^H) u_\nu^{H'} v' - \pi^2 u_\nu^H v] = 0 \quad \forall v \in V^H$$

zu einem vorgegebenen ν^H . Das bedeutet, dass wir eine Modellierung $a_m(u, v) = \nu^H u' v'$ verwenden. Hierbei ist zu beachten, dass ν^H nicht nur als feste Zahl, sondern als über $x \in \Omega$ veränderliche Funktion interpretiert werden kann.

Sei nun ν^H wie in der nachfolgenden Abbildung dargestellt, dann ergeben sich entsprechend andere Werte für das Funktional der von ν^H abhängenden Lösung u_ν^H :



$$j(u) = 1, \quad j(u^H) = 0.84206, \quad j(u_\nu^H) = 1$$

Das hier verwendete ν^H eignet sich also gut für die Problemstellung, jedoch ist nicht klar, woher ein gutes ν^H zu erhalten ist.

Als strukturelle Idee wollen wir versuchen, die Approximation (5.4) zu nutzen, um das Modell a_m zu verbessern. Die Verbesserung soll dabei darauf ausgerichtet sein, $|j(u) - j(u_m^H)|$ klein zu halten.

$$j(\hat{u}) - j(\hat{u}^H) + \frac{1}{2}R = 0 = \eta^H + \eta^m \quad (5.5)$$

Voraussetzung für die folgenden Schritte ist, dass die Formen a , a_m sowie f bzw. j von folgender Form sind:

$$\begin{aligned} a(u, v) &= \int_{\Omega} a^{\text{vol}}(u, v) + \int_{\partial\Omega} a^{\text{bnd}}(u, v), \\ a_m(u, v) &= \int_{\Omega} a_m^{\text{vol}}(u, v) + \int_{\partial\Omega} a_m^{\text{bnd}}(u, v), \\ \langle f, v \rangle &= \int_{\Omega} f^{\text{vol}}(v) + \int_{\partial\Omega} f^{\text{bnd}}(v), \\ j(v) &= \int_{\Omega} j^{\text{vol}}(v) + \int_{\partial\Omega} j^{\text{bnd}}(v) \end{aligned} \quad (5.6)$$

mit geeigneten Abbildungen

$$a^{\text{vol}}, a^{\text{bnd}}, a_m^{\text{vol}}, a_m^{\text{bnd}}: V \times V \rightarrow S \quad \text{sowie} \quad f^{\text{vol}}, f^{\text{bnd}}, j^{\text{vol}}, j^{\text{bnd}}: V \rightarrow S_{\partial},$$

wobei

$$S := \{f \mid f: \Omega \rightarrow \mathbb{R}\} \quad \text{und} \quad S_{\partial} := \{f \mid f: \partial\Omega \rightarrow \mathbb{R}\}.$$

Es sei N ein Hilbertraum mit Ansatzfunktionen $\nu: \Omega \rightarrow \mathbb{R}$ für alle $\nu \in N$. Die konstante Funktion 1 sei in N .

Definition 5.3.1. Wir bilden eine *Lokalisierung* der Funktionen a, f, j, a_m aus (5.6) mithilfe von N :

$$\begin{aligned} a(\mu, u, v) &:= \int_{\Omega} \mu a^{\text{vol}}(u, v) + \int_{\partial\Omega} \mu a^{\text{bnd}}(u, v), \\ \langle f_{\mu}, v \rangle &:= \int_{\Omega} \mu f^{\text{vol}}(v) + \int_{\partial\Omega} \mu f^{\text{bnd}}(v), \\ j(\mu, v) &:= \int_{\Omega} \mu j^{\text{vol}}(v) + \int_{\partial\Omega} \mu j^{\text{bnd}}(v) \end{aligned}$$

sowie

$$\begin{aligned} a_m(\nu, \mu, u, v) &:= \int_{\Omega} \nu \mu a_m^{\text{vol}}(u, v) + \int_{\partial\Omega} \nu \mu a_m^{\text{bnd}}(u, v) \\ a_m(\nu, u, v) &:= a_m(\nu, 1, u, v) \end{aligned}$$

für alle $\nu, \mu \in N, u, v \in V$.

Bemerkung 5.3.1. Die Bezeichnungen der Abbildungen werden hier mehrfach verwendet. Jedoch lässt die Anzahl der Argumente stets den Schluss auf die gemeinte Abbildung zu. Es ist

$$\begin{aligned} a(1, u, v) &= a(u, v), & a(0, u, v) &= 0, \\ \langle f_1, v \rangle &= \langle f, v \rangle, & \langle f_0, v \rangle &= 0, \\ j(1, v) &= j(v), & j(0, v) &= 0, \\ a_m(1, 1, u, v) &= a_m(u, v) & \text{sowie } a_m(0, \mu, u, v) &= a_m(\mu, 0, u, v) = 0 \end{aligned}$$

für alle $\mu \in N$ und $u, v \in V$.

Bemerkung 5.3.2. Auch für die partiellen Ableitungen gilt dann

$$\begin{aligned} j'(\mu, v; w) &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [j(\mu, v + \varepsilon w) - j(\mu, v)], \\ a'(\mu, u; w, v) &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [a(\mu, u + \varepsilon w, v) - a(\mu, u, v)], \\ a'_m(\nu, \mu, u; w, v) &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [a_m(\nu, \mu, u + \varepsilon w, v) - a_m(\nu, \mu, u, v)]. \end{aligned}$$

Sie sind also wohldefiniert.

Die grobskaligen (modellbehafteten) Probleme sollen nun mithilfe der Lokalisierung verbessert werden. Dafür starten wir mit einem Startparameter $\nu \in N^H$, etwa zu Beginn $\nu = 1$ oder $\nu = 0$. Das primale Problem ist dann

$$u_{\nu}^H \in V^H : \quad a(1, u_{\nu}^H; v) + a_m(\nu, 1, u_{\nu}^H; v) = \langle f_1, v \rangle \quad \forall v \in V^H$$

und das duale Problem ist

$$z_{\nu}^H \in V^H : \quad a'(1, u_{\nu}^H; w, z_m^H) + a'_m(\nu, 1, u_{\nu}^H; v) = j'(1, u_{\nu}^H; w) \quad \forall w \in V^H.$$

Bemerkung 5.3.3. Für $\nu = 0$ wirkt keine Modellierung, für $\nu = 1$ ist die Modellierung wie in (5.2)-(5.3) beschrieben. In diesen Problemen ist stets $\mu = 1$.

Wir bezeichnen das Lösungspaar wieder mit

$$x_\nu^H = (u_\nu^H, z_\nu^H) \in X^H$$

und definieren ein entsprechend vom Parameter ν abhängiges Funktional über X :

$$L_\nu(x) := L(x) - a_m(\nu, u, z).$$

Es gilt wegen

$$L'_\nu(x; y) = \rho^*(u; z, w) - a'_m(\nu, u; w, z) + \rho(u, v) - a_m(\nu, u, v)$$

der Zusammenhang

$$\begin{aligned} x_\nu \in X &: & L'_\nu(x_\nu; y) &= 0 \quad \forall y \in X, \\ x_\nu^H \in X^H &: & L'_\nu(x_\nu^H; y) &= 0 \quad \forall y \in X^H \end{aligned}$$

sowie

$$j(u_\nu^H) = L_\nu(x_\nu^H).$$

Satz 5.3.2. *Sofern a , a_m sowie j hinreichend differenzierbar (bezüglich u) sind und $X^H \subseteq X^h$, dann gilt*

$$\begin{aligned} j(u^h) - j(u_\nu^H) &= \frac{1}{2} [a_m(\nu, u_\nu^H, z_\nu^H) - a'_m(\nu, u_\nu^H; u_\nu^H, z_\nu^H)] \\ &\quad + \frac{1}{2} [\rho^*(u_\nu^H; z_\nu^H, u^h) + \rho(u_\nu^H, z^h)] - \frac{1}{2}R. \end{aligned}$$

wobei $e := (e_u, e_z) := (u^h - u_\nu^H, z^h - z_\nu^H)$ und

$$R := \int_0^1 \sigma(1 - \sigma) L'''(x_\nu^H + \sigma e; e, e, e) d\sigma.$$

Beweis.

$$\begin{aligned} j(u^h) - j(u_\nu^H) &= L(x^h) - L_\nu(x_\nu^H) \\ &= L(x^h) - L(x_\nu^H) + a_m(\nu, u_\nu^H, z_\nu^H) \\ &= a_m(\nu, u_\nu^H, z_\nu^H) + \int_0^1 L'(x_\nu^H + \sigma e; e) d\sigma. \end{aligned}$$

Durch Anwendung der Trapezregel gilt

$$\int_0^1 L'(x_\nu^H + \sigma e; e) d\sigma = \frac{1}{2} [L'(x_\nu^H; e) + L'(x^h; e) - R].$$

Da $X^H \subseteq X^h$ und daher $e \in X^h$ ist, folgt

$$L'(x^h; e) = 0$$

und

$$\begin{aligned}
L'(x_\nu^H; e) &= L'_\nu(x_\nu^H; e) + a'_m(\nu, u_\nu^H; e_u, z_\nu^H) + a_m(\nu, u_\nu^H, e_z) \\
&= L'_\nu(x_\nu^H; x^h) + a'_m(\nu, u_\nu^H; e_u, z_\nu^H) + a_m(\nu, u_\nu^H, e_z) \\
&= L'(x_\nu^H; x^h) - a'_m(\nu, u_\nu^H; u^h, z_\nu^H) - a_m(\nu, u_\nu^H, z^h) + a'_m(\nu, u_\nu^H; e_u, z_\nu^H) + a_m(\nu, u_\nu^H, e_z) \\
&= L'(x_\nu^H; x^h) - a'_m(\nu, u_\nu^H; u_\nu^H, z_\nu^H) - a_m(\nu, u_\nu^H, z_\nu^H).
\end{aligned}$$

Zusammengefasst ergibt sich

$$\begin{aligned}
&\int_0^1 L'(x_\nu^H + \sigma e; e) d\sigma \\
&= \frac{1}{2} [L'(x_\nu^H; e) - R] \\
&= \frac{1}{2} [L'(x_\nu^H; x^h) - a'_m(\nu, u_\nu^H; u_\nu^H, z_\nu^H) - a_m(\nu, u_\nu^H, z_\nu^H) - R] \\
&= \frac{1}{2} [\rho^*(u_\nu^H; z_\nu^H, u^h) + \rho(u_\nu^H, z^h) - a'_m(\nu, u_\nu^H; u_\nu^H, z_\nu^H) - a_m(\nu, u_\nu^H, z_\nu^H) - R].
\end{aligned}$$

□

Folgerung 5.3.1. *Es gelten die Voraussetzungen und Bezeichnungen von Satz 5.3.2. Ist a_m sogar linear in u , so ist*

$$j(u^h) - j(u_\nu^H) = \frac{1}{2} [\rho(u_\nu^H, z^h) + \rho^*(u_\nu^H; z_\nu^H, u^h)] - \frac{1}{2} R.$$

Beweis. Nachrechnen mit

$$a'_m(\nu, w; u, v) = a_m(\nu, u, v) \quad \forall u, v, w \in V \forall \nu \in N.$$

□

Ein einfacher Ansatz wäre es, den Parameter ν losgelöst zu betrachten, d. h. unabhängig von den anderen Größen. Für unsere Modellierungsform

$$a_m(\nu, u, v) = \int_\Omega \nu a_m^{\text{vol}}(u, v) + \int_{\partial\Omega} \nu a_m^{\text{bnd}}(u, v)$$

hieße das,

$$a'_m(\nu, u; w, v) = \int_\Omega \nu a_m^{\text{vol}'}(u; w, v) + \int_{\partial\Omega} \nu a_m^{\text{bnd}'}(u; w, v).$$

Sind a_m^{vol} sowie a_m^{bnd} jedoch linear im ersten Argument, d. h.

$$a_m^{\text{vol}'}(u; w, v) = a_m^{\text{vol}}(w, v) \quad \text{und} \quad a_m^{\text{bnd}'}(u; w, v) = a_m^{\text{bnd}}(w, v),$$

dann ist auch

$$a'_m(\nu, u; w, v) = \int_\Omega \nu a_m^{\text{vol}}(w, v) + \int_{\partial\Omega} \nu a_m^{\text{bnd}}(w, v) = a_m(\nu, w, v),$$

also a_m linear.

Wählen wir jedoch an dieser Stelle eine Abhängigkeit des Parameters ν von u und v , so erhalten wir

$$a_m(\nu(u, v), u, v) = \int_{\Omega} \nu(u, v) a_m^{\text{vol}}(u, v) + \int_{\partial\Omega} \nu(u, v) a_m^{\text{bnd}}(u, v)$$

und

$$\begin{aligned} \frac{\partial}{\partial u} a_m(\nu(u, v), u, v)[w] &= \int_{\Omega} \left[\nu(u, v) a_m^{\text{vol}'}(u; w, v) + \nu'(u; w, v) a_m^{\text{vol}}(u, v) \right] \\ &\quad + \int_{\partial\Omega} \left[\nu(u, v) a_m^{\text{bnd}'}(u; w, v) + \nu'(u; w, v) a_m^{\text{bnd}}(u, v) \right]. \\ &= a_m'(\nu(u, v), u; w, v) + a_m(\nu'(u; w, v), u, v). \end{aligned}$$

Für ein lineares a_m gilt dann entsprechend

$$\frac{\partial}{\partial u} a_m(\nu(u, v), u, v)[w] = a_m(\nu(u, v), w, v) + a_m(\nu'(u; w, v), u, v).$$

Für die parametrisierte Modellierung können wir also entsprechend festhalten:

$$\begin{aligned} j(u^h) - j(u_\nu^H) + \frac{1}{2}R &= \eta_\nu^H + \eta_\nu^m + \eta_\nu^m, \\ \eta_\nu^H &:= \frac{1}{2} \left[\rho(u_\nu^H, z^h) + \rho^*(u_\nu^H; z_\nu^H, u^h) \right] \\ &= \frac{1}{2} \left[\langle f, z^h \rangle - a(u_\nu^H, z^h) + j'(u_\nu^H; u^h) - a'(u_\nu^H; u^h, z_\nu^H) \right], \\ \eta_\nu^m &:= \frac{1}{2} \left[a_m(\nu(u_\nu^H, z_\nu^H), u_\nu^H, z_\nu^H) - a_m(\nu(u_\nu^H, z_\nu^H), u_\nu^H; u_\nu^H, z_\nu^H) \right], \\ \eta_\nu^m &:= -\frac{1}{2} a_m(\nu'(u_\nu^H; u_\nu^H, z_\nu^H), u_\nu^H, z_\nu^H). \end{aligned} \tag{5.7}$$

Um eine Modellierungskorrektur zu erhalten, wollen wir nach ν auflösen. Bei dem Ansatz der Unabhängigkeit von ν und (u, v) wäre in (5.7) $\eta_\nu^m = 0$ und bei linearem a_m ist nach Folgerung 5.3.1 auch $\eta_\nu^m = 0$ – folglich sind alle ν -Terme verschwunden.

Mit diesem zweiten Ansatz der Abhängigkeit ist jedoch bei linearem a_m wiederum $\eta_\nu^m = 0$, jedoch $\eta_\nu^m \neq 0$. Allerdings haben wir dann nur eine (implizite) Formulierung bezüglich ν' , und nicht wie gewünscht explizit nach ν . An dieser Stelle wäre eine (approximative) Vereinfachung wünschenswert.

Die Abhängigkeit zwischen ν und (u, v) muss dazu noch weiter spezifiziert werden. Wenn wir die Abbildung

$$F(\kappa, u, v, x) := f^{\text{vol}}(v)(x) - a^{\text{vol}}(u, v)(x) - \kappa a_m^{\text{vol}}(u, v)(x) \quad \forall \kappa \in \mathbb{R}^+ \forall u, v \in V \forall x \in \Omega$$

betrachten, so kennzeichnet die Lösung u_ν^H des groben modellbehafteten Problems, dass

$$\begin{aligned} F(\nu(u_\nu^H, v)(x), u_\nu^H, v, x) &= f^{\text{vol}}(v)(x) - a^{\text{vol}}(u_\nu^H, v)(x) - \nu(u_\nu^H, v)(x) a_m^{\text{vol}}(u_\nu^H, v)(x) \\ &= 0 \quad \forall v \in V^H. \end{aligned}$$

Dies können wir nutzen, um eine implizite Formulierung für $\nu: V \times V \rightarrow N$ zu erhalten:

Satz 5.3.3. $F: \mathbb{R}^+ \times V \times V \times \Omega \rightarrow \mathbb{R}$ sei einmal stetig differenzierbar und für ein $\nu^*: V \rightarrow N$ sowie $u^* \in V$ und $D \subset V \times \Omega$ gelte

$$F(\nu^*(v)(x), u^*, v, x) = 0 \quad \forall (v, x) \in D.$$

Ist nun

$$F'(\nu^*(v)(x); 1, u^*, v, x) \neq 0 \quad \forall (v, x) \in D,$$

dann existiert eine Umgebung $U \subset V$ von $u^* \in U$ sowie eine Abbildung $\nu: U \times V \rightarrow N$ mit

$$F(\nu(u, v)(x), u, v, x) = 0 \quad \forall (u, v, x) \in U \times D$$

und es gilt

$$\nu'(u; w, v)(x) = -\frac{F'(\nu(u, v)(x), u; w, v, x)}{F'(\nu(u, v)(x); 1, u, v, x)} \quad \forall (v, x) \in D.$$

Beweis. Satz über implizite Funktionen: Wir erhalten für jedes $(v, x) \in D$:

$$\begin{aligned} 0 &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [F(\nu(u + \varepsilon w, v)(x), u + \varepsilon w, v, x) - F(\nu(u, v)(x), u, v, x)] \\ &= F'(\nu(u, v)(x); \nu'(u; w, v)(x), u, v, x) + F'(\nu(u, v)(x), u; w, v, x) \quad \forall (v, x) \in D. \end{aligned}$$

Wird dies nach $\nu'(u; w, v)(x)$ aufgelöst, so ergibt sich

$$\nu'(u; w, v)(x) = -\frac{F'(\nu(u, v)(x), u; w, v, x)}{F'(\nu(u, v)(x); 1, u, v, x)} \quad \forall (v, x) \in D.$$

□

Wir verwenden unser F wie oben definiert und erhalten

$$\nu'(u; w, v)(x) = -\frac{a^{\text{vol}'}(u; w, v)(x) + \nu(u, v)(x)a_m^{\text{vol}'}(u; w, v)(x)}{a_m^{\text{vol}}(u, v)(x)} \quad \forall (u, w, v, x) \in V \times V \times D.$$

Diese Gleichung stimmt nur für $(v, x) \in D$. Wir wollen die Gleichung an dieser Stelle auf alle $(v, x) \in V \times \Omega$ beziehen, und den Fehler, der in den Punkten $(v, x) \in (V \times \Omega) \setminus D$ entsteht (insbesondere auch den in $x \in \partial\Omega$ durch Weglassen der Randintegrale), mit Q bezeichnen:

$$a_m(\nu'(u; w, v), u, v) = -a'(u; w, v) - a'_m(\nu(u, v), u; w, v) + Qa_m(u, v) \quad \forall (u, w, v) \in V \times V \times V.$$

Unter Verwendung dieser Annahmen ist

$$\begin{aligned} \eta_{\nu'}^m &= -\frac{1}{2}a_m(\nu'(u_\nu^H; u_\nu^H, z_\nu^H), u_\nu^H, z_\nu^H) \\ &= \frac{1}{2} [a'(u_\nu^H; u_\nu^H, z_\nu^H) + a'_m(\nu(u_\nu^H, z_\nu^H), u_\nu^H; u_\nu^H, z_\nu^H) - Qa_m(u_\nu^H, z_\nu^H)]. \end{aligned}$$

Wenn wir zusätzlich annehmen, dass a_m linear und $Q \approx 0$, können wir $\eta_{\nu'}^m$ lokal approximieren durch

$$\begin{aligned}\eta_{\nu'}^m &= \frac{1}{2} [a'(u_{\nu'}^H; u_{\nu'}^H, z_{\nu'}^H) + a_m(\nu(u_{\nu'}^H, z_{\nu'}^H), u_{\nu'}^H, z_{\nu'}^H) - Qa_m(u_{\nu'}^H, z_{\nu'}^H)] \\ &\approx \frac{1}{2} [a'(u_{\nu'}^H; u_{\nu'}^H, z_{\nu'}^H) + a_m(\nu(u_{\nu'}^H, z_{\nu'}^H), u_{\nu'}^H, z_{\nu'}^H)].\end{aligned}$$

Folgerung 5.3.2. *Es gelten die Voraussetzungen der Sätze 5.3.2 sowie 5.3.3. Ist $a_m(u, v)$ linear, dann gilt*

$$\begin{aligned}j(u^h) - j(u_{\nu'}^H) &= \frac{1}{2} [\rho^*(u_{\nu'}^H; z_{\nu'}^H, u^h) + \rho(u_{\nu'}^H, z^h) + a'(u_{\nu'}^H; u_{\nu'}^H, z_{\nu'}^H)] \\ &\quad + \frac{1}{2} a_m(\nu(u_{\nu'}^H, z_{\nu'}^H), u_{\nu'}^H, z_{\nu'}^H) - \frac{1}{2} (Qa_m(u_{\nu'}^H, z_{\nu'}^H) + R).\end{aligned}$$

Beweis. Nachrechnen mit $a'_m(u; w, v) = a_m(w, v)$ bzw. $a'_m(\nu, u; w, v) = a_m(\nu, w, v)$. □

Wir setzen nun

$$\begin{aligned}d_{u^H, z^H}(\nu, \mu) &:= \frac{1}{2} a_m(\nu, \mu, u^H, z^H) \quad \forall \nu, \mu \in N, \\ \langle f_{u^H, u^h, z^H, z^h}, \mu \rangle &:= \frac{1}{2} [a(\mu, u^H, z^h) - \langle f_{\mu}, z^h \rangle \\ &\quad + a'(\mu, u^H; u^h - u^H, z^H) - j'(\mu, u^H; u^h)] \quad \forall \mu \in N.\end{aligned}$$

Das variationelle Problem, finde $\nu \in N$ mit

$$d_{u^H, z^H}(\nu, \mu) = \langle f_{u^H, u^h, z^H, z^h}, \mu \rangle \quad \forall \mu \in N,$$

entspricht dann der Lokalisierung der Gleichung

$$\eta_{\nu'}^m + \eta_{\nu'}^m = -\eta_{\nu'}^H$$

unter Verwendung der obigen Annahmen und Approximationen.

Konkret können wir dies natürlich wieder nur approximativ in einem diskreten Unterraum N^H von N lösen. Finde $\nu^H \in N^H$ mit

$$d_{u_{\nu'}^H, z_{\nu'}^H}(\nu^H, \mu) = \langle f_{u_{\nu'}^H, u^h, z_{\nu'}^H, z^h}, \mu \rangle \quad \forall \mu \in N^H$$

Insgesamt sind also (nach Zeitdiskretisierung in jedem Zeitschritt) folgende Probleme zu lösen:

$$u^h \in V^h : \quad a(u^h, v) = \langle f, v \rangle \quad \forall v \in V^h, \quad (5.8)$$

$$z^h \in V^h : \quad a'(u^h; w, z^h) = j'(u^h; w) \quad \forall w \in V^h, \quad (5.9)$$

$$u_{\nu'}^H \in V^H : \quad a(u_{\nu'}^H, v) + a_m(\nu^H, u_{\nu'}^H, v) = \langle f, v \rangle \quad \forall v \in V^H, \quad (5.10)$$

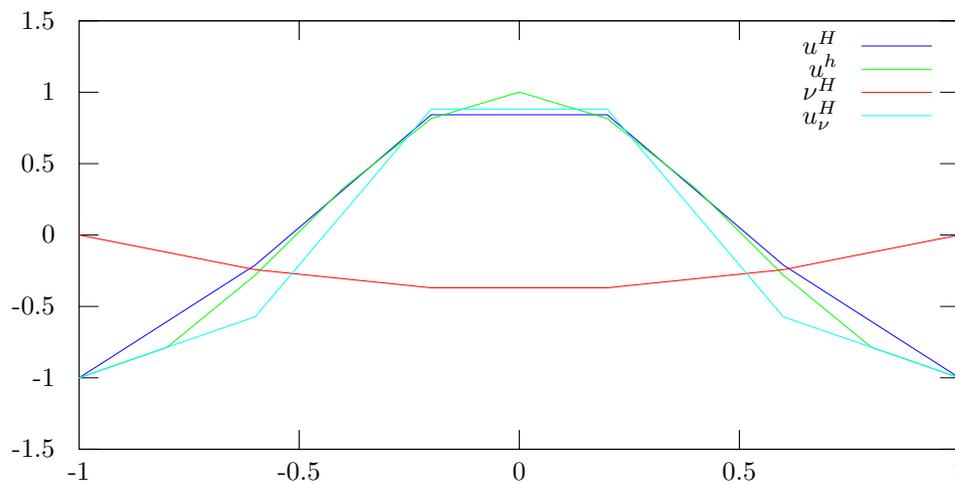
$$z_{\nu'}^H \in V^H : \quad a'(u_{\nu'}^H; w, z_{\nu'}^H) + a'_m(\nu^H, u_{\nu'}^H; w, z_{\nu'}^H) = j'(u_{\nu'}^H; w) \quad \forall w \in V^H, \quad (5.11)$$

$$\nu^H \in N^H : \quad d_{u_{\nu'}^H, z_{\nu'}^H}(\nu^H, \mu) = \langle f_{u_{\nu'}^H, u^h, z_{\nu'}^H, z^h}, \mu \rangle \quad \forall \mu \in N^H. \quad (5.12)$$

Zu beachten ist, dass die Gleichungen (5.10)-(5.12) voneinander abhängen. Es handelt sich folglich um ein (nichtlineares) Variationsgleichungssystem. Hier kann z. B. iterativ vorgegangen werden: Zunächst wird ein ν^H (etwa $\nu^H = 0$ oder $\nu^H = 1$) vorgegeben, dann können die obigen Gleichungen der Reihe nach gelöst werden. Ggf. wird nun das Lösen der letzten drei Gleichungen iteriert, sofern noch wesentliche Abweichungen entstehen.

Mit jeder Iteration von $u_m^H \rightsquigarrow \bar{u}_m^H$ dürfte sich der Fehler in j verringern: $|j(u^h) - j(\bar{u}_m^H)|$ sollte nun kleiner sein als $|j(u^h) - j(u_m^H)|$. Allerdings ergeben sich Fehler etwa aufgrund der weggelassenen Restglieder R, Q , der Diskretisierung oder anderen Approximationen.

Beispiel 5.3.3. (Fortführung von Beispiel 5.3.1, 5.3.2) Nach einem einzigen Schritt der Gleichung (5.12) sowie der dann aus (5.10) resultierenden primalen Lösung erhalten wir



$$j(u) = 1, \quad j(u^H) = 0.84206, \quad j(u_\nu^H) = 0.88172$$

Der Funktionalwert verbessert sich bereits und nähert sich mit weiteren Approximationen weiter der Referenz an.

5.4 Skalen- und Modellierungsabschätzung für die Navier-Stokes-Gleichungen

Wir wollen nun wieder die Formulierungen des gemischten Problems verwenden, etwa $\hat{u} = (u, p)$. Nach einer Zeitdiskretisierung haben wir zur Beschreibung der stationären Probleme:

$$\begin{aligned} \hat{V} &= V \times Q, \\ X &:= \hat{V} \times \hat{V}, \\ \hat{a}(\hat{u}, \hat{v}) &:= (2\nu_0 \nabla^s u, \nabla v) - (p, \nabla \cdot v) - (\nabla \cdot u, q) + \frac{1}{2} [(u \cdot \nabla u, v) - (u \cdot \nabla v, u)] \\ &\quad + \frac{1}{2} \int_{\Gamma_N} [(u \cdot v)(u \cdot n) - (2\nu_0 (n \cdot \nabla^s) u) \cdot v], \\ \hat{a}'(\hat{u}; \hat{w}, \hat{v}) &= (2\nu_0 \nabla^s w, \nabla v) - (r, \nabla \cdot v) - (\nabla \cdot w, q) \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} [(w \cdot \nabla u, v) + (u \cdot \nabla w, v) - (w \cdot \nabla v, u) - (u \cdot \nabla v, w)] \\
& + \frac{1}{2} \int_{\Gamma_N} [(w \cdot v)(u \cdot n) + (u \cdot v)(w \cdot n) - (2\nu_0(n \cdot \nabla^s)w) \cdot v], \\
\hat{a}_m(\hat{u}, \hat{v}) & := (2\nabla^s u, \nabla v) - \int_{\Gamma_N} ((n \cdot \nabla^s)u) \cdot v, \\
\hat{a}'_m(\hat{u}; \hat{w}, \hat{v}) & = (2\nabla^s w, \nabla v) - \int_{\Gamma_N} ((n \cdot \nabla^s)w) \cdot v, \\
\langle \hat{f}, \hat{v} \rangle & := \langle f, v \rangle + \int_{\Gamma_N} j_N \cdot v, \\
\hat{j}(\hat{u}) & := \int_{\Gamma_j} (n \cdot (\nu_0 \nabla^s u - p\mathbb{I})) \cdot \psi, \\
\hat{j}'(\hat{u}; \hat{w}) & = \int_{\Gamma_j} (n \cdot (\nu_0 \nabla^s w - p\mathbb{I})) \cdot \psi,
\end{aligned}$$

wobei $\hat{u} = (u, p)$, $\hat{v} = (v, q)$, $\hat{w} = (w, r)$, $\hat{z} = (z, s) \in \hat{V} = V \times Q$ und $\nu_0 > 0$.

Hierbei ist das Funktional \hat{j} so gewählt, dass der Strömungswiderstandskoeffizient des von Γ_j umschlossenen Hindernisses gemessen wird.

Wir betrachten das primale Problem

$$\hat{u}_m^H \in \hat{V}^H : \quad \hat{a}(\hat{u}_m^H, \hat{v}) + \hat{a}_m(\hat{u}_m^H, \hat{v}) = \langle \hat{f}, \hat{v} \rangle \quad \forall \hat{v} \in \hat{V}^H$$

sowie das duale Problem

$$\hat{z}_m^H \in \hat{V}^H : \quad \hat{a}'(\hat{u}_m^H; \hat{w}, \hat{z}_m^H) + \hat{a}'_m(\hat{u}_m^H; \hat{w}, \hat{z}_m^H) = \hat{j}'(\hat{u}_m^H; \hat{w}) \quad \forall \hat{w} \in \hat{V}^H.$$

Entsprechend wird als Referenzlösung

$$\hat{u}^h \in \hat{V}^h : \quad \hat{a}(\hat{u}^h; \hat{v}) = \langle \hat{f}, \hat{v} \rangle \quad \forall \hat{v} \in \hat{V}^h$$

sowie das duale Problem

$$\hat{z}^h \in \hat{V}^h : \quad \hat{a}'(\hat{u}^h; \hat{w}, \hat{z}^h) = \hat{j}'(\hat{u}^h; \hat{w}) \quad \forall \hat{w} \in \hat{V}^h$$

auf einem feineren Gitter gelöst.

Nach Folgerung 5.3.2 ist nun

$$\begin{aligned}
\hat{j}(\hat{u}^h) - \hat{j}(\hat{u}_\nu^H) & \approx \frac{1}{2} \left[\langle \hat{f}, \hat{z}^h \rangle - \hat{a}(\hat{u}_\nu^H, \hat{z}^h) + \hat{j}'(\hat{u}_\nu^H; \hat{u}^h) - \hat{a}'(\hat{u}_\nu^H; \hat{u}^h - \hat{u}_\nu^H, \hat{z}_\nu^H) \right] \\
& + \frac{1}{2} \hat{a}_m(\nu(\hat{u}_\nu^H, \hat{z}_\nu^H), \hat{u}_\nu^H, \hat{z}_\nu^H).
\end{aligned}$$

Um Indikatoren für die Korrektur von ν zu bekommen, werden die einzelnen Terme lokal betrachtet.

5.5 Modellierungskorrektur für die Navier-Stokes-Gleichungen

Die Voraussetzungen (5.6) sind für $\hat{a}, \hat{a}_m, \hat{j}$ und \hat{f} erfüllt. Es ist also möglich – wie oben skizziert – mithilfe der Fehlerabschätzung eine Korrektur des Modells \hat{a}_m durchzuführen. Grundidee dafür ist die Lokalisierung der Gleichung

$$j(\hat{u}) - j(\hat{u}_m^H) \approx 0 = \eta^H + \eta_\nu^m + \eta_{\nu'}^m,$$

wodurch der Fehler im Funktional j klein werden soll.

Dazu wählen wir einen Ansatzraum N , etwa $N := Q$ und führen die Lokalisierungen ein:

$$\begin{aligned} \hat{a}(\mu, \hat{u}, \hat{v}) &:= (2\nu_0 \nabla^s u, \mu \nabla v) - (\mu p, \nabla \cdot v) - (\nabla \cdot u, \mu q) + \frac{1}{2} [(u \cdot \nabla u, \mu v) - (u \cdot \nabla v, \mu u)] \\ &\quad + \frac{1}{2} \int_{\Gamma_N} \mu [(u \cdot v)(u \cdot n) - (2\nu_0 (v \cdot \nabla^s) u) \cdot n], \\ \langle \hat{f}_\mu, \hat{v} \rangle &:= \langle f, \mu v \rangle + \int_{\Gamma_N} \mu j_N \cdot v, \\ \hat{j}(\mu, \hat{u}) &:= \int_{\Gamma_j} \mu ((n \cdot \nabla^s) u) \cdot \psi, \end{aligned}$$

sowie

$$\hat{a}_m(\nu, \mu, \hat{u}, \hat{v}) := (2\nu \nabla^s u, \mu \nabla v) - \int_{\Gamma_N} 2((\mu v \cdot \nabla^s) u v) \cdot n$$

für alle $\nu, \mu \in N$ und $u, v \in V$.

Dann ist auch

$$\begin{aligned} \hat{a}'(\mu, \hat{u}; \hat{w}, \hat{v}) &:= (2\nu_0 \nabla^s w, \mu \nabla v) - (\mu r, \nabla \cdot v) - (\nabla \cdot w, \mu q) \\ &\quad + \frac{1}{2} [(w \cdot \nabla u, \mu v) + (u \cdot \nabla w, \mu v) - (w \cdot \nabla v, \mu u) - (u \cdot \nabla v, \mu w)] \\ &\quad + \frac{1}{2} \int_{\Gamma_N} \mu [(w \cdot v)(u \cdot n) + (u \cdot v)(w \cdot n) - (2\nu_0 (v \cdot \nabla^s) w) \cdot n], \\ \hat{a}'_m(\nu, \mu, \hat{u}; \hat{w}, \hat{v}) &:= (2\nu \nabla^s w, \mu \nabla v) - \int_{\Gamma_N} 2((\mu v \cdot \nabla^s) w v) \cdot n, \end{aligned}$$

d. h. insbesondere ist \hat{a}_m linear.

Nun können wir die Größen für das variationelle Problem der lokalen Korrektur formulieren:

$$\begin{aligned} d_{\hat{u}^H, \hat{z}^H}(\nu, \mu) &:= \frac{1}{2} (2\nu \nabla^s u^H, \mu \nabla z^H) - \int_{\Gamma_N} 2((\mu z^H \cdot \nabla^s) u^H \nu) \cdot n \quad \forall \nu, \mu \in N, \\ \langle f_{\hat{u}^H, \hat{u}^h, \hat{z}^H, \hat{z}^h}, \mu \rangle &:= \frac{1}{2} \left[\hat{a}(\mu, \hat{u}^H, \hat{z}^h) - \langle \hat{f}_\mu, \hat{z}^h \rangle \right. \\ &\quad \left. + \hat{a}'(\mu, \hat{u}^H; \hat{u}^h - \hat{u}^H, \hat{z}^H) - \hat{j}'(\mu, \hat{u}^H; \hat{u}^h) \right] \quad \forall \mu \in N \end{aligned}$$

und bestimmen $\nu^H \in N^H$ durch

$$d_{\hat{u}_\nu^H, \hat{z}_\nu^H}(\nu^H, \mu) = \left\langle f_{\hat{u}_\nu^H, \hat{u}^h, \hat{z}_\nu^H, \hat{z}^h}, \mu \right\rangle \quad \forall \mu \in N^H.$$

Insgesamt sind also nach Zeitdiskretisierung folgende Probleme zu lösen:

$$\begin{aligned} \hat{u}^h \in \hat{V}^h : & \quad \hat{a}(\hat{u}^h, \hat{v}) = \left\langle \hat{f}, \hat{v} \right\rangle \quad \forall \hat{v} \in \hat{V}^h, \\ \hat{z}^h \in \hat{V}^h : & \quad \hat{a}'(\hat{u}^h; \hat{w}, \hat{z}^h) = \hat{j}'(\hat{u}^h; \hat{w}) \quad \forall \hat{w} \in \hat{V}^h, \\ \hat{u}_\nu^H \in \hat{V}^H : & \quad \hat{a}(\hat{u}_\nu^H, \hat{v}) + \hat{a}_m(\nu^H, \hat{u}_\nu^H, \hat{v}) = \left\langle \hat{f}, \hat{v} \right\rangle \quad \forall \hat{v} \in \hat{V}^H, \\ \hat{z}_\nu^H \in \hat{V}^H : & \quad \hat{a}'(\hat{u}_\nu^H; \hat{w}, \hat{z}_\nu^H) + \hat{a}'_m(\nu^H, \hat{u}_\nu^H; \hat{w}, \hat{z}_\nu^H) = \hat{j}'(\hat{u}_\nu^H; \hat{w}) \quad \forall \hat{w} \in \hat{V}^H, \\ \nu^H \in N^H : & \quad d_{\hat{u}_\nu^H, \hat{z}_\nu^H}(\nu^H, \mu) = \left\langle f_{\hat{u}_\nu^H, \hat{u}^h, \hat{z}_\nu^H, \hat{z}^h}, \mu \right\rangle \quad \forall \mu \in N^H. \end{aligned}$$

Die unteren drei Gleichungen können dann bei Bedarf durch Iteration gelöst werden.

6 Mehrskalige Rechnung zur Parameterbestimmung

Der im letzten Kapitel vorgestellte Ansatz zur Bestimmung einer künstlichen Viskosität für die stationären inkompressiblen Navier-Stokes-Gleichungen soll nun konkret umgesetzt werden. Ziel ist es, den Parameter $\nu \in N$ der Navier-Stokes-Gleichung so einzustellen, dass der Fehler in einem vorgegebenen Funktional j klein ist.

Um auch instationäre Probleme behandeln zu können, wird dazu zunächst eine Semidiskretisierung der Zeit durchgeführt. In jedem Zeitschritt $t = 0, t_1, \dots, t_N = T$ wird dann der Algorithmus 6.1 angewendet.

Algorithm 6.1 Mehrskalige Parameterbestimmung

Input: $\varepsilon_{\text{tol}}, i_{\text{max}}, \nu_{\text{min}}, \nu_{\text{max}}, d_{\text{damp}}, \nu_0^H$
Output: $u^h, p^h, z^h, s^h, u^H, p^H, z^H, s^H, \nu^H$

Löse das feinskalige primale Problem $\rightarrow (u^h, p^h)$.

Löse das feinskalige duale Problem $\rightarrow (z^h, s^h)$.

$i \leftarrow 0$

repeat

if $i > 0$ **then**

 Verwende $\hat{u}_{i-1}^H, \hat{u}_{i-1}^h, \hat{z}_{i-1}^H$ und \hat{z}_{i-1}^h zur Bestimmung von $\rightarrow \nu_i^H$.

 Glätte Änderung $\nu_i^H \leftarrow \min(\nu_{\text{min}}, \max(\nu_{\text{max}}, \nu_0 + d_{\text{damp}}(\nu_i^H - \nu_{i-1}^H) + \nu_{i-1}^H)) - \nu_0$.

end if

 Verwende ν_i^H zur Lösung des grobskaligen primalen Problems $\rightarrow (u_i^H, p_i^H)$.

 Verwende ν_i^H zur Lösung des grobskaligen dualen Problems $\rightarrow (z_i^H, s_i^H)$.

 Setze $r_i = (u^h - u_i^H, p^h - p_i^H)$ und $\rho_i = (z^h - z_i^H, s^h - s_i^H)$.

if $i > 0$ **and** $|j(u_i^H) - j(u^h)| > |j(u_{i-1}^H) - j(u^h)|$ **then**

return $(\nu^H, u^H, p^H) \leftarrow (\nu_{i-1}^H, u_{i-1}^H, p_{i-1}^H)$.

end if

 Zähle Durchläufe $i \leftarrow i + 1$.

until $\|r_{i-1}\| + \|\rho_{i-1}\| < \varepsilon_{\text{tol}}$ **or** $i \geq i_{\text{max}}$

return $(\nu^H, u^H, p^H) \leftarrow (\nu_{i-1}^H, u_{i-1}^H, p_{i-1}^H)$.

Zu Beginn wird als erste künstliche Viskosität ν_0^H die konstante Funktion 0 verwendet. In späteren Zeitschritten wird die letzte künstliche Viskosität des vorangegangenen Zeitschritts übernommen.

Das Problem der gekoppelten Gleichungen wird nun angegangen, indem die Gleichungen iterativ nacheinander gelöst werden. Dies wird in einer Schleife solange wiederholt, bis

- entweder die Abweichungen zwischen grobskaliger und feinskaliger Lösung hinreichend klein sind ($\|r_{i-1}\| + \|\rho_{i-1}\| < \varepsilon_{\text{tol}}$),

- die Korrektur keinen positiven Effekt auf das Funktional hat (d. h. größere Abweichung, $|j(u_i^H) - j(u^h)| > |j(u_{i-1}^H) - j(u^h)|$) oder
- die maximale Iterationsanzahl erreicht wird $i \geq i_{\max}$.

Die dabei bestimmten ν^H werden jeweils vor der Verwendung als künstliche Viskosität noch einer Glättung unterzogen:

$$\nu_i^H \leftarrow \min(\nu_{\min}, \max(\nu_{\max}, \nu_0 + d_{\text{damp}}(\nu_i^H - \nu_{i-1}^H) + \nu_{i-1}^H)) - \nu_0.$$

Dabei wird sichergestellt, dass

- die Änderung $\nu_i^H - \nu_{i-1}^H$ mit einem Dämpfungsfaktor d_{damp} versehen wird und
- die zukünftig verwendete Viskosität $\nu_0 + \nu_i^H$ zwischen ν_{\min} und ν_{\max} liegt.

Die Bestimmung von ν_i^H aus \hat{u}_{i-1}^H , \hat{u}_{i-1}^h , \hat{z}_{i-1}^H und \hat{z}_{i-1}^h erfolgte zuvor wie im letzten Kapitel ausgeführt im Wesentlichen durch:

$$\nu_i^H \in N^H : \quad d_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\nu_i^H, \mu) = \left\langle f_{\hat{u}_{i-1}^H, \hat{u}_{i-1}^h, \hat{z}_{i-1}^H, \hat{z}_{i-1}^h}, \mu \right\rangle \quad \forall \mu \in N^H. \quad (6.1)$$

Es ist jedoch zu beachten, dass es sich hierbei um ein inverses Problem handelt. Je nach Problem benötigen wir zum Lösen eine Regularisierung, etwa die Methode der kleinsten Quadrate [BG06]: Finde $\nu_i^H \in N^H$, so dass

$$\begin{aligned} & \left(\mathcal{L}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\nu_i^H), \mathcal{L}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) \right) \\ & + \int_{\Gamma_N} \mathcal{R}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\nu) \mathcal{R}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) = \left\langle f_{\hat{u}_{i-1}^H, \hat{u}_{i-1}^h, \hat{z}_{i-1}^H, \hat{z}_{i-1}^h}^{\text{LS}}, \mathcal{L}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) \right\rangle \\ & \quad \quad \quad + \int_{\Gamma_N} g_{\hat{u}_{i-1}^H, \hat{u}_{i-1}^h, \hat{z}_{i-1}^H, \hat{z}_{i-1}^h}^{\text{LS}} \mathcal{R}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) \quad \forall \mu \in N^H \end{aligned}$$

mit

$$\begin{aligned} \mathcal{L}_{\hat{u}^H, \hat{z}^H}(\nu) & := \nu(\nabla^s u^H : \nabla z^H), \\ \mathcal{R}_{\hat{u}^H, \hat{z}^H}(\nu) & := -2\nu((z^H \cdot \nabla^s)u^H) \cdot n, \\ \left\langle f_{\hat{u}^H, \hat{u}^h, \hat{z}^H, \hat{z}^h}^{\text{LS}}, \mu \right\rangle & := \frac{1}{2} \left\langle 2\nu_0(\nabla^s u^H : \nabla z^h) - (\nabla \cdot z^h)p^H - (\nabla \cdot u^H)s^h, \mu \right\rangle \\ & \quad + \frac{1}{4} \left\langle (u^H \cdot \nabla u^H) \cdot z^h - (u^H \cdot \nabla z^h) \cdot u^H, \mu \right\rangle - \frac{1}{2} \left\langle f, \mu z^h \right\rangle \\ & \quad + \left\langle \nu_0(\nabla^s(u^h - u^H) : \nabla z^H), \mu \right\rangle \\ & \quad - \frac{1}{2} \left\langle (\nabla \cdot z^H)(p^h - p^H) + (\nabla \cdot (u^h - u^H))s^H, \mu \right\rangle \\ & \quad + \frac{1}{4} \left\langle ((u^h - u^H) \cdot \nabla u^H) \cdot z^H + (u^H \cdot \nabla(u^h - u^H)) \cdot z^H, \mu \right\rangle \\ & \quad - \frac{1}{4} \left\langle ((u^h - u^H) \cdot \nabla z^H) \cdot u^H + (u^H \cdot \nabla z^H) \cdot (u^h - u^H), \mu \right\rangle, \end{aligned}$$

$$\begin{aligned}
g_{\hat{u}^H, \hat{u}^h, \hat{z}^H, \hat{z}^h}^{\text{LS}} &:= \frac{1}{4} \left[(u^H \cdot z^h)(u^H \cdot n) - (2\nu_0(z^h \cdot \nabla^s)u^H) \cdot n \right] - \frac{1}{2} j_N \cdot z^h \\
&+ \frac{1}{4} \left[((u^h - u^H) \cdot z^H)(u^H \cdot n) + (u^H \cdot z^H)((u^h - u^H) \cdot n) \right] \\
&- \frac{1}{2} \nu_0((z^H \cdot \nabla^s)(u^h - u^H)) \cdot n - \frac{1}{2} \chi_{\Gamma_j}((n \cdot \nabla^s)u^h) \cdot \psi.
\end{aligned}$$

Für $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{d \times d}$ und eine Menge $M \subseteq \Omega$ ist dabei $A : B := \sum_{1 \leq i, j \leq d} a_{ij} b_{ij}$ und

$$\chi_M(x) := \begin{cases} 1 & x \in M, \\ 0 & \text{sonst.} \end{cases}$$

Durch die quadratischen Ausdrücke wird die Systemmatrix positiv definit. Allerdings kann ihre Kondition immer noch schlecht sein, so dass wir zum Lösen des inversen Problems zusätzlich noch eine Tichonov-Regularisierung einsetzen [HW90]. Das variationelle Problem lautet dann: Finde $\nu_i^H \in N^H$, so dass

$$\begin{aligned}
& \left(\mathcal{L}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\nu_i^H), \mathcal{L}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) \right) + \alpha(\nu, \mu) \\
& + \int_{\Gamma_N} \mathcal{R}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\nu) \mathcal{R}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) = \left\langle f_{\hat{u}_{i-1}^H, \hat{u}_{i-1}^h, \hat{z}_{i-1}^H, \hat{z}_{i-1}^h}^{\text{LS}}, \mathcal{L}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) \right\rangle \\
& + \int_{\Gamma_N} g_{\hat{u}_{i-1}^H, \hat{u}_{i-1}^h, \hat{z}_{i-1}^H, \hat{z}_{i-1}^h}^{\text{LS}} \mathcal{R}_{\hat{u}_{i-1}^H, \hat{z}_{i-1}^H}(\mu) \quad \forall \mu \in N^H. \quad (6.2)
\end{aligned}$$

Mithilfe eines Parameters $\alpha > 0$ wird hierbei die Systemmatrix stärker zu einer Identitätsabbildung modifiziert. Dies hat neben einer besseren Kondition der Systemmatrix zur Folge, dass die Lösungen ν^H kleiner werden und nicht so stark von Rundungsfehlern beeinflusst sind. Hierbei sind wir daran interessiert, den Parameter nur so groß wie nötig, aber so klein wie möglich zu halten. Da insbesondere der Grad der nötigen Regularisierung in jedem Zeitschritt variiert, wird eine adaptive Parametersuche eingesetzt, die in Algorithmus 6.2 beschrieben ist.

Im ersten Zeitschritt wird $\alpha_{\text{old}} = 1$ gesetzt, später wird jeweils der im zuletzt berechneten Zeitschritt bestimmte Regularisierungsparameter α verwendet.

6.1 Duales instationäres Problem

Das oben vorgestellte Verfahren basiert darauf, dass zunächst eine Zeitdiskretisierung durchgeführt wird und jeweils für die einzelnen stationären (primalen) Probleme das duale (ebenfalls stationäre) Problem betrachtet wird. Tatsächlich ist das betrachtete Problem jedoch zeitabhängig, d. h. diese Vorgehensweise ist zumindest fragwürdig. Gehen wir vom instationären primalen Problem aus, so ergibt sich dazu ein duales Problem, das ebenfalls instationär ist. Dieses instationäre Problem kann zwar wiederum zeitdiskretisiert werden, so dass wir eine Folge von stationären Problemen erhalten. Allerdings geht als Parameter jeweils die gesamte (instationäre) Lösung des primalen Problems ein. Dies erweist sich unter praktischen Gesichtspunkten als ungünstig, denn zusätzlich zum erhöhten Speicherbedarf zur Darstellung der

Algorithm 6.2 Regularisierung

Input: $\hat{u}_m^H, \hat{u}_m^h, \hat{z}_m^H, \hat{z}_m^h, \alpha_{\text{old}}$ **Output:** ν, α

Versuche (6.1) zu lösen

if (6.1) war nicht lösbar **then** $\alpha \leftarrow \alpha_{\text{old}}$ $s \leftarrow 0$ **while** noch keine Lösung ν gefunden wurde oder $s = 0$ **do**

Versuche (6.2) zu lösen

if (6.2) war lösbar **then****if** $s > 0$ **then****return** ν **end if** $\alpha \leftarrow \frac{\alpha}{10}$ **else** $\alpha \leftarrow 10 \cdot \alpha$ $s \leftarrow 1$ **end if****end while****end if**

instationären Lösung kommt noch, dass sämtliche Ergebnisse des Lösungsweges des temporalen Problems, wie die Schrittweitensteuerung oder die Viskositätsmodellierung, reproduzierbar und wieder abrufbar für die nächste Iteration sind.

Dieser Mehraufwand war ausschlaggebend für die Wahl des vorgestellten Verfahrens. Auch wenn mit ihm die temporalen Effekte im dualen Problem nicht bzw. nur sehr eingeschränkt abgebildet werden können, so ist zumindest eine Implementierung mit vertretbarem Aufwand erreichbar. Es ist zu beachten, dass dabei gewisse Fehler zu erwarten sind, sofern die temporalen Effekte beim dualen Problem nicht unerheblich sind.

Grundsätzlich denkbar ist hingegen auch, das vorgestellte Verfahren für die instationären Probleme zu formulieren und unter den oben aufgeführten praktischen Nachteilen zu möglicherweise besseren Resultaten zu kommen. Dies konnte jedoch aufgrund nicht verfügbarer Infrastruktur nicht in dieser Arbeit untersucht werden.

Sofern für die instationäre Lösung gewisse Eigenschaften erwartet werden, können wir in dieser Fragestellung möglicherweise auch einen Mittelweg wählen: In [BBT11] wird ein Ansatz vorgestellt, bei dem das duale Problem zeitlich gemittelt wird, so dass nur ein stationäres Problem zu lösen ist. Dies ist unter der Voraussetzung einer (quasi-)periodischen Lösung nachvollziehbar und in dem dort verfolgten Verfahren zur Gitteradaption hilfreich. Eine Anwendung dieser Idee auf die oben formulierte Kopplung der Probleme zur Modellgewinnung ist ebenfalls denkbar.

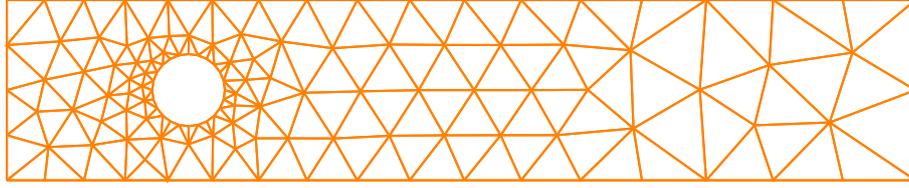


Abbildung 6.1: Geometrie und Gitter des Modellproblems A

6.2 Modellprobleme

6.2.1 Modellproblem A

Die Geometrie des zu betrachtenden Strömungsproblems ist mit der aus [STD⁺96] vergleichbar, wenngleich hier eine höhere Reynolds-Zahl (1300 statt 100) verwendet wird.

Es wird ein zweidimensionales Strömungsgebiet mit einem Hindernis in Form eines Kreises betrachtet. Das betrachtete Gebiet hat die Ausmaße 5×1 . Das Hindernis hat einen Radius von 0.2, seinen Mittelpunkt 1 nach dem Einflusrand und ist zentriert im Strömungsrohr.

Am oberen und unteren Rand des Rohrs sowie am Hindernis herrschen Solid-Wall-Randbedingungen. Die Strömung wird am linken Rand durch Dirichlet-Einflussbedingungen vorgegeben, am rechten Rand gilt die künstliche Neumann-Randbedingung (3.8) nach [Bru00]. Für die Geometrie von Gebiet und Gitter sei auf Abbildung 6.1 verwiesen. Das grobe Gitter kennzeichnet sich durch eine Gitterweite von $0.03 \lesssim H \lesssim 0.6$, die Triangulierung besteht aus 198 Dreieckselementen. Das feine Gitter ergibt sich durch Halbierung der Kantenlänge aller Elemente, d. h. einem Schritt h -Refinements.

Als Viskosität wurde hier $\nu_0 = 3 \cdot 10^{-4}$ gewählt. Die Reynolds-Zahl beträgt somit etwa $\text{Re} = \frac{0.4 \cdot |u|}{\nu_0} \approx 1300$ nachdem sich die Strömung entwickelt hat, wobei hier der Durchmesser des Strömungshindernisses als charakteristische Länge verwendet wurde.

Das Funktional j , bezüglich dessen die Rechnung optimiert werden soll, ist bei dieser Strömung der Strömungswiderstand des Hindernisses,

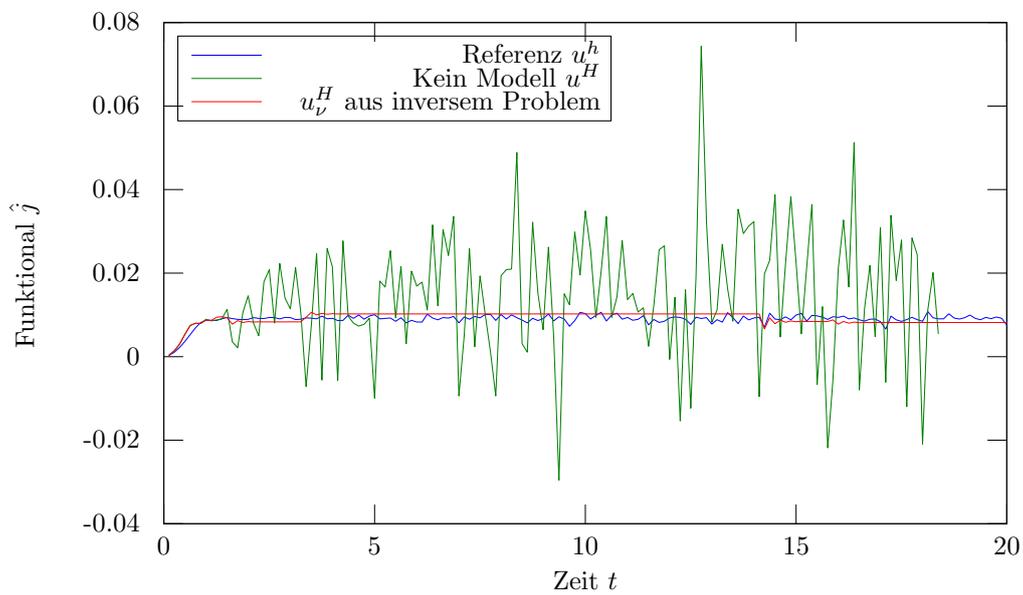
$$\hat{j}(\hat{u}) = \int_{\Gamma_{\text{obs}}} n \cdot (\nu \nabla^s u - p \mathbb{I}) \psi,$$

wobei Γ_{obs} die Berandung des kreisförmigen Hindernisses im Strömungsrohr ist und $\psi = (1, 0)^\top$ die Hauptströmungsrichtung.

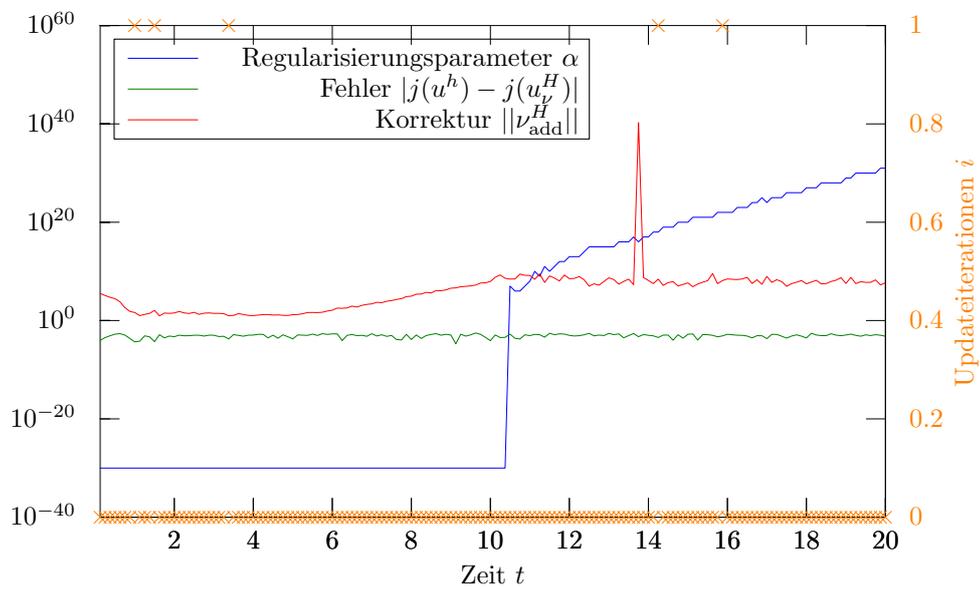
Es wird der Zeitraum $I = (0, 20)$ betrachtet, wobei für $t \in (0, 1)$ die Einflusrandbedingung aufgebaut wird.

Beim Lösen des inversen Problems wurden die Parameter $\nu_{\min} = 3 \cdot 10^{-5}$, $\nu_{\max} = 3 \cdot 10^{-1}$ und $d_{\text{damp}} = 1$ verwendet.

In Abbildung 6.2 sind die Ergebnisse der Berechnung dieses Modellproblems dargestellt. In (a) ist zu erkennen, wie instabil die Strömung bei grober Diskretisierung und ohne künstliche Viskosität ist. Dahingegen ist die Strömung bei feiner Diskretisierung relativ gleichförmig. Mit der künstlichen Viskosität aus dem inversen Problem stellt sich ein noch gleichförmigeres



(a)



(b)

Abbildung 6.2: Entwicklung (a) des Funktionals sowie (b) der Parameter des inversen Problems bei Modellproblem A

Funktional ein, dessen Funktionswert sich stets an dem der feinen Diskretisierung orientiert. Gut zu erkennen sind die ruckartigen Schwankungen, wenn die künstliche Viskosität „nachjustiert“ wird. Hervorzuheben ist, dass diese Updates fünf Mal stattgefunden haben, wie aus Abbildung 6.2(b) zu entnehmen ist. Obwohl die (späteren) Änderungen jeweils einen deutlichen Einfluss auf den Funktionalwert j haben (vgl. 6.2(a)), ist dies nicht in dem Fehler im Funktional $|j(u^h) - j(h_\nu^U)|$ zu erkennen. Interessant ist, dass der Regularisierungsparameter α ab etwa $t \approx 10$ langsam aber gleichmäßig ansteigt.

6.2.2 Modellproblem B

Das hier betrachtete Strömungsproblem orientiert sich an [MKM99].

Es wird ein zweidimensionales Strömungsgebiet vom Format $2\pi \times 1$ betrachtet.

Am oberen und unteren Rand des Rohrs herrschen Solid-Wall-Randbedingungen. Die Strömung wird durch eine Druckdifferenz zwischen linkem und rechtem Rand vorgegeben. Für die Geometrie von Gebiet und Gitter sei auf Abbildung 6.3 verwiesen. Für das grobe Gitter gilt hier eine Gitterweite von $0.03 \lesssim H \lesssim 0.1$, die Triangulierung besteht aus 8192 Dreieckselementen. Ebenfalls wird die feine Diskretisierung durch einen h -Refinement-Schritt erreicht.

Die Viskosität der Strömung betrage $\nu_0 = 4.6225 \cdot 10^{-5}$. Die Reynolds-Zahl liegt hier bei $\text{Re} = \frac{1 \cdot |\bar{u}|}{\nu_0} \approx 5250$ bzw. $\text{Re}_\tau = \frac{\delta \cdot u_\tau}{\nu_0} \approx 590$ nach Entwicklung der Strömung. δ ist der Röhrenradius, u_τ die Schubspannungsgeschwindigkeit und $|\bar{u}|$ die mittlere Strömungsgeschwindigkeit in der Mitte der Röhre.

Als Funktional j , bezüglich dessen die Rechnung optimiert werden soll, ist bei dieser Strömung der Durchfluss in Hauptströmungsrichtung,

$$\hat{j}(\hat{u}) = \int_M u \cdot \psi,$$

wobei $M = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [-0.5, 0.5]$ so gewählt ist, dass die Randeffekte bei Zu- bzw. Abfluss möglichst vernachlässigt werden. $\psi = (1, 0)^\top$ ist die Hauptströmungsrichtung.

Es wird der Zeitraum $I = [0, 5]$ betrachtet, wobei für $t \in [0, 1]$ die Randbedingung aus einem unbewegten Strömungsgebiet aufgebaut wird.

Beim Lösen des inversen Problems wurden die Parameter $\nu_{\min} = 10^{-5}$, $\nu_{\max} = 10^{-2}$ und $d_{\text{damp}} = 10^{-8}$ verwendet.

Die Abbildung 6.4 zeigt das Funktional j sowie die Ergebnisse bei Anwendung des vorgestellten Algorithmus auf Modellproblem B. In (a) ist gut zu erkennen, wie die beiden unstabilierten Rechnungen voneinander abweichen: Während bei der groben Diskretisierung der Funktio-

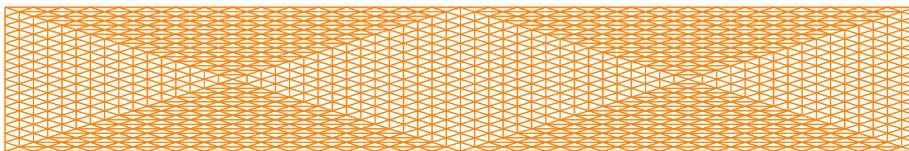
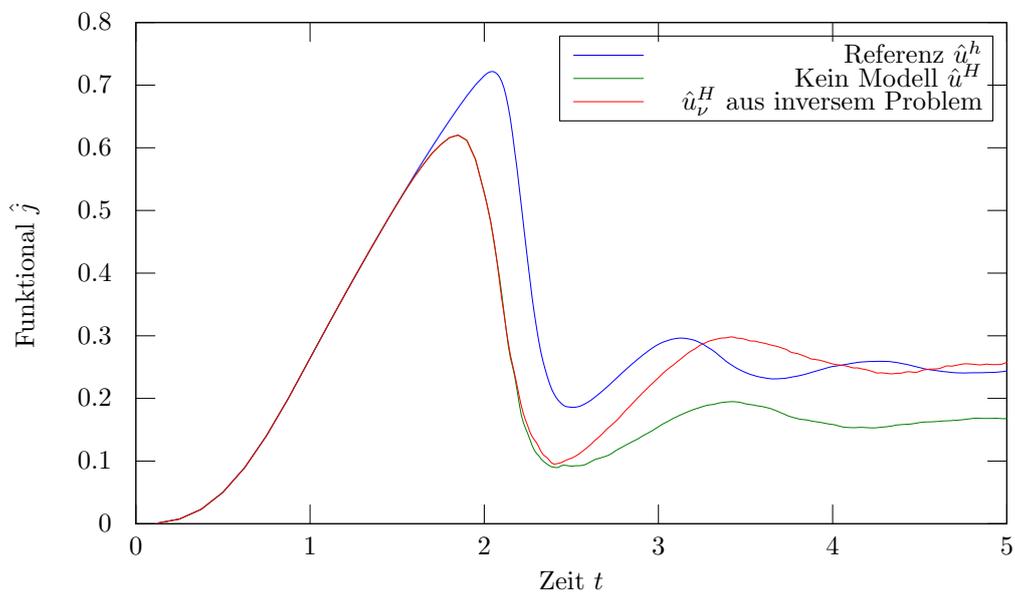
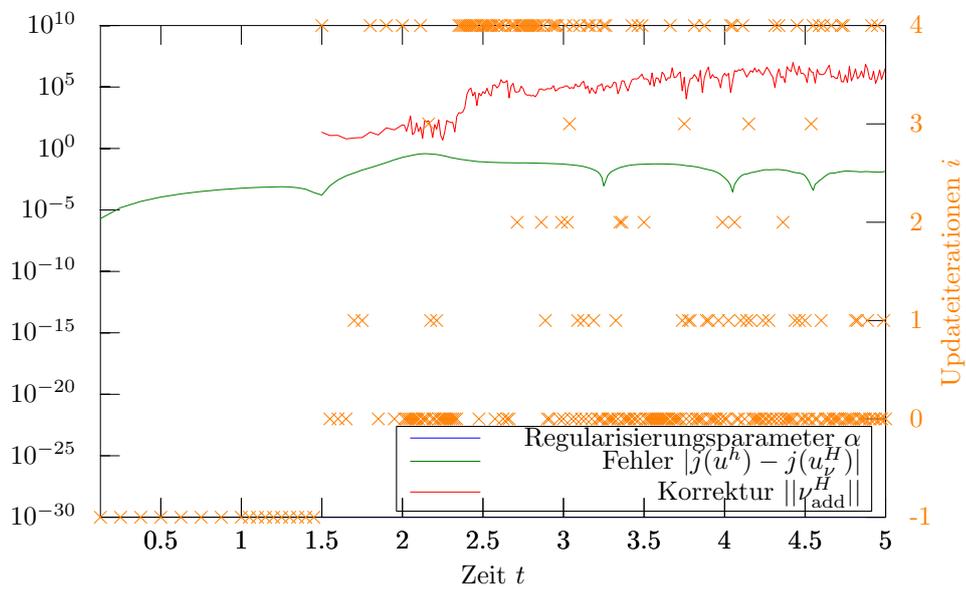


Abbildung 6.3: Geometrie und Gitter des Modellproblems B



(a)



(b)

Abbildung 6.4: Entwicklung (a) des Funktionals sowie (b) der Parameter des inversen Problems bei Modellproblem B

nalwert bereits früher einbricht, baut sich die Strömung bei der feinen Diskretisierung länger auf. Auch der sich nach Strömungsaufbau einstellende Durchsatz unterscheidet sich bei den Rechnungen: Bei der feineren Diskretisierung werden hier höhere Werte bestimmt als bei der Groben.

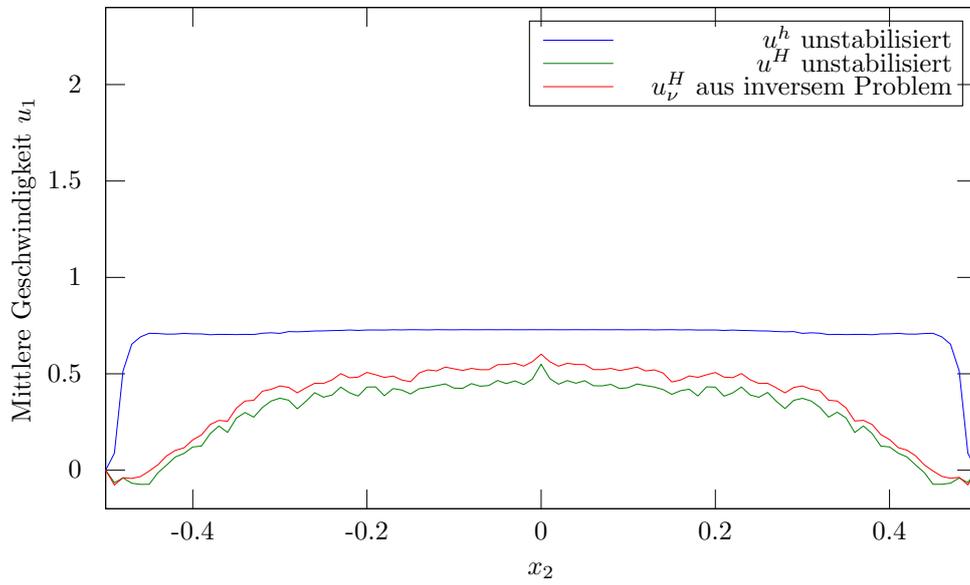
Da die Funktionale bei grob- und feinskaliger Diskretisierung zu Beginn sehr gut übereinstimmen, wird die Korrektur mithilfe des inversen Problems erst ab $t \geq t_{\text{start}} := 1.5$ durchgeführt. Dies erspart Rechenzeit, da in diesen Zeitschritten ohnehin keine wesentliche Korrektur nötig und möglich ist. In Abbildung 6.4(b) ist dies durch eine Updateiteration von $i = -1$ gekennzeichnet.

Für die grobskalige Diskretisierung mit Stabilisierung aus dem inversen Problem ist in 6.4(a) zu erkennen, dass sie sich zunächst ähnlich der unstabilisierten grobskaligen verhält. Der Einbruch der Anfangsströmung verläuft weitgehend gleich. Erst später, beim eingeschwungenen Durchsatz, orientiert sich der Funktionalwert an dem der feinskaligen Lösung und nähert sich ihr an. Auch wenn die Schwingungsphasen nicht übereinstimmen, können wir doch festhalten, dass der durchschnittliche Level des Funktionals hier sehr gut getroffen wird.

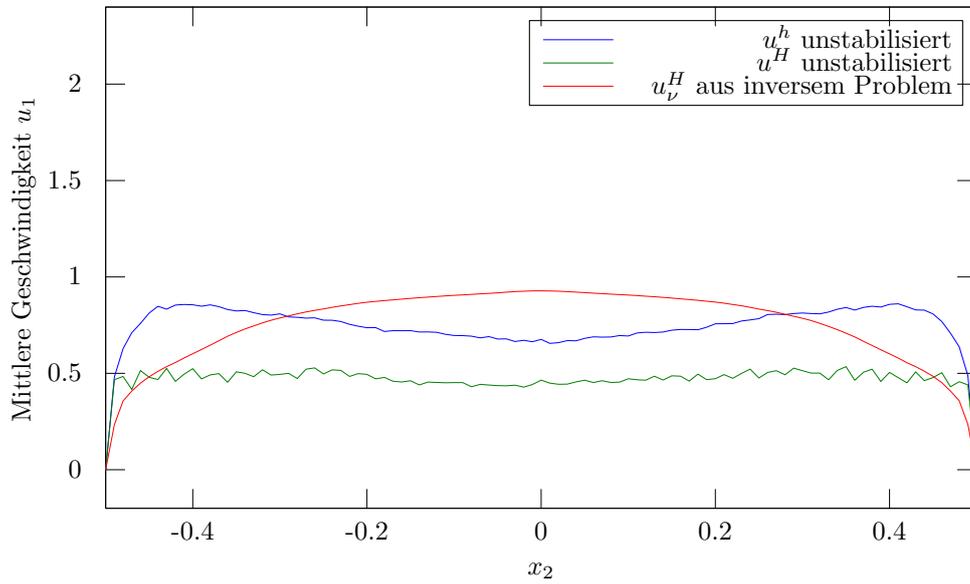
Bemerkenswert ist, dass bei der Berechnung zu Modellproblem B stets nur eine minimale Regularisierung von $\alpha = 10^{-30}$ nötig ist. Außerdem ist gut zu erkennen, dass im Unterschied zu Modellproblem A deutlich häufiger und mehr an der künstlichen Viskosität korrigiert wird. Es gibt zwar etliche Zeitpunkte, zu denen keine Korrektur nötig ist ($i = 0$), oder wenige Korrekturen ausreichen ($1 \leq i \leq 3$), aber auch etliche Zeitpunkte, an denen die Korrektur abgebrochen wird ($i = i_{\text{max}} = 4$). Dies dürfte aber auch mit der extrem starken Dämpfung d_{damp} in Zusammenhang stehen, die bei der Korrektur genutzt wird.

Zusätzlich zum Wert des Funktionals können wir uns bei diesem Modellproblem den Strömungsverlauf im Röhrenquerschnitt anschauen, siehe Abbildung 6.5. Dort werden zwar zum Teil klare Differenzen deutlich, jedoch muss auch berücksichtigt werden, dass in dem zur Optimierung verwendeten Funktional lediglich der Gesamtdurchfluss einbezogen ist, nicht jedoch das genaue Strömungsprofil.

In Abbildung 6.6 sind vergleichbare Strömungsprofile aus Rechnungen von [DAG⁺11] aufgetragen. Es handelt sich um die Ergebnisse von sieben unabhängigen Gruppen mit unterschiedlichen Codes. Jede der Gruppen hat ein favorisiertes Viskositätsmodell ausgewählt und verwendet. Im Vergleich zu den Ergebnissen aus Abbildung 6.5(b) fällt auf, dass keines der Profile ein derartiges Rauschen wie die unstabilisierte grobskalige Lösung aufweist. Überraschenderweise scheint das Profil aus dem inversen Problem sogar besser zu den Vergleichsdaten zu passen als jenes aus der feinskaligen Rechnung. Zwischen den Ergebnissen der einzelnen Gruppen variiert jedoch auch der Level der Geschwindigkeiten deutlich. Dieser Effekt wird durch das Funktional quantifiziert und durch den hier vorgestellten Ansatz zu verringern versucht.



(a)



(b)

Abbildung 6.5: Geschwindigkeitsquerschnitte bei Modellproblem B über den Zeitraum (a) $t \in (2.25, 2.75)$ sowie (b) $t \in (2.75, 5)$

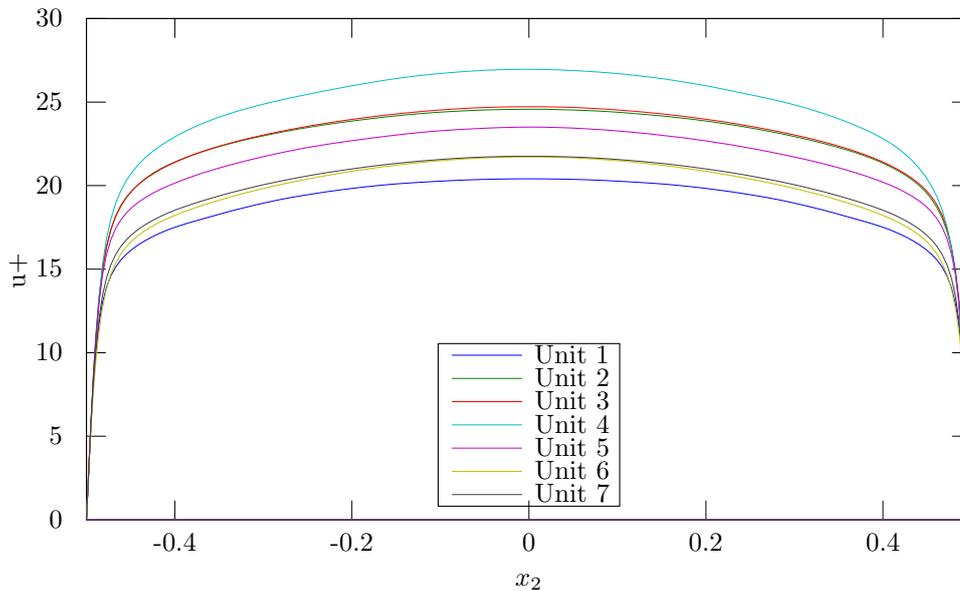


Abbildung 6.6: Vergleichsströmungsprofile aus [DAG+11]

6.2.3 Modellproblem C

Das nun betrachtete Strömungsproblem orientiert sich an [KMM87].

Es wird ein dreidimensionales Strömungsgebiet vom Format $2\pi \times 1 \times \pi$ betrachtet.

Am oberen, unteren, vorderen sowie hinteren Rand des Rohrs mit rechteckigem Querschnitt herrschen Solid-Wall-Randbedingungen. Die Strömung wird durch eine Druckdifferenz zwischen linkem und rechtem Rand vorgegeben. Das heißt, die Hauptströmung ist in x_1 -Richtung. Für die Geometrie von Gebiet und Gitter sei auf Abbildung 6.7 verwiesen. Für das grobe Gitter gilt hier eine mittlere Gitterweite von $H \approx 1.07$, die Triangulierung besteht aus 352 Tetraederelementen. Ebenfalls wird die feine Diskretisierung durch einen h -Refinement-Schritt erreicht.

Die Viskosität der Strömung betrage $\nu_0 = 1.5151 \cdot 10^{-4}$, so dass hier eine Reynolds-Zahl von $Re = \frac{1 \cdot \overline{|u|}}{\nu_0} \approx 2000$ bzw. $Re_\tau = \frac{\delta \cdot u_\tau}{\nu_0} \approx 180$ nach Entwicklung der Strömung vorliegt. δ ist dabei der Röhrenradius, u_τ die Schubspannungsgeschwindigkeit und $\overline{|u|}$ die mittlere

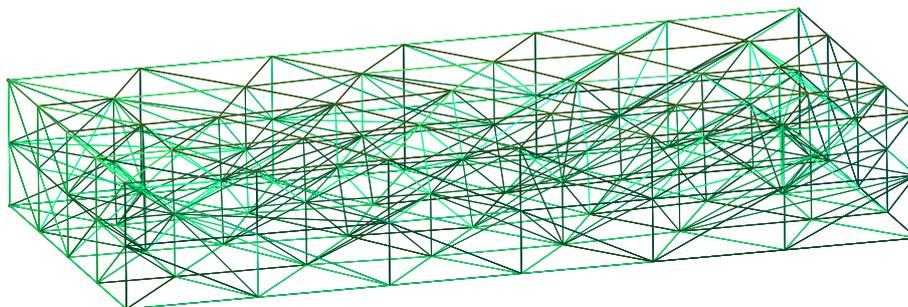


Abbildung 6.7: Geometrie und Gitter des Modellproblems C

Strömungsgeschwindigkeit in der Mitte der Röhre.

Als Funktional j , bezüglich dessen die Rechnung optimiert werden soll, ist bei dieser Strömung der Durchfluss in Hauptströmungsrichtung,

$$\hat{j}(\hat{u}) = \int_M u \cdot \psi,$$

wobei $M = [-\frac{\pi}{2}, \frac{\pi}{2}] \times [-0.5, 0.5] \times [-\frac{\pi}{4}, \frac{\pi}{4}]$ so gewählt ist, dass die Randeffekte bei Zu- bzw. Abfluss möglichst ignoriert werden. $\psi = (1, 0, 0)^\top$ ist die Hauptströmungsrichtung.

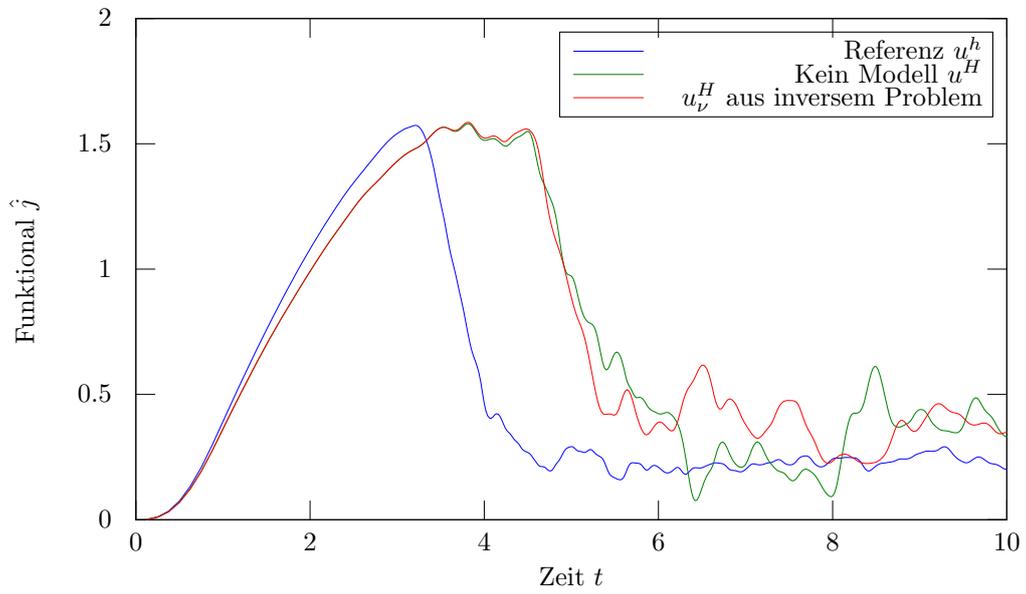
Es wird der Zeitraum $I = [0, 10]$ betrachtet, wobei für $t \in [0, 1]$ die Randbedingung aufgebaut wird.

Beim Lösen des inversen Problems wurden die Parameter $\nu_{\min} = 10^{-5}$, $\nu_{\max} = 10^{-3}$ und $d_{\text{damp}} = 10^{-1}$ verwendet.

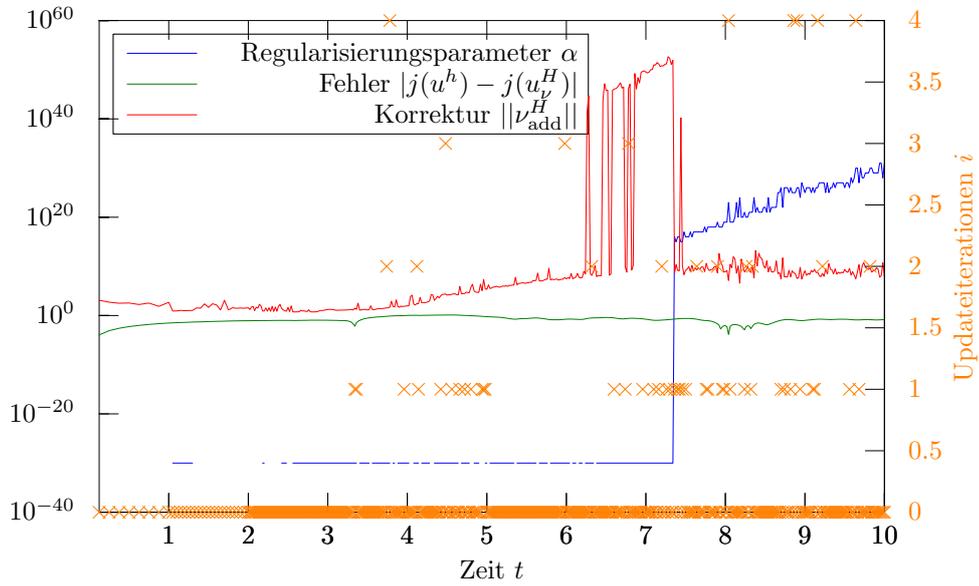
Die Abbildung 6.8 zeigt das Funktional j sowie die Ergebnisse bei Anwendung des vorgestellten Algorithmus auf Modellproblem C.

Auch in diesem Modellproblem können wir uns zusätzlich zum Wert des Funktionals wieder den Strömungsverlauf im Röhrenquerschnitt anschauen, siehe Abbildung 6.9. Deutlich fällt die unterschiedliche Gitterstruktur der grob- und feinskaligen Rechnung auf. Die Diskretisierung ist hier also bei weitem nicht ausreichend, um sämtliche turbulente Skalen aufzulösen. Dennoch ist zumindest in Abbildung 6.9(a) zu erkennen, wie sich das Strömungsprofil bei der Rechnung mit inversem Problem trotz grober Skalenauflösung denen der feinskaligen Rechnung schneller annähern, als dies ohne Stabilisierung der Fall ist.

Schließlich wurden in Tabelle 6.1 die wesentlichen Parameter der verschiedenen Modellproblem-Rechnungen gegenübergestellt.

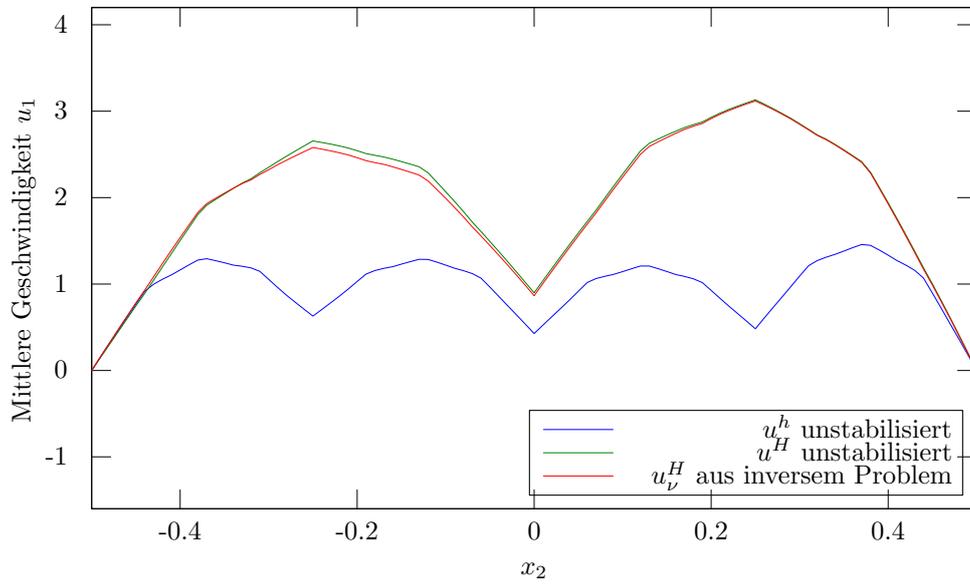


(a)

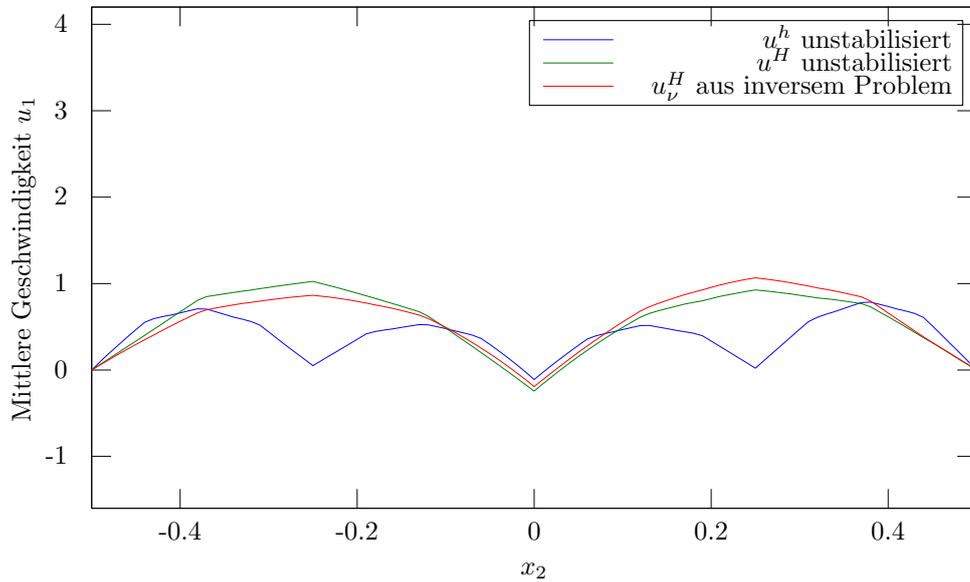


(b)

Abbildung 6.8: Entwicklung (a) des Funktionals sowie (b) der Parameter des inversen Problems bei Modellproblem C



(a)



(b)

Abbildung 6.9: Geschwindigkeitsquerschnitte bei Modellproblem C über den Zeitraum (a) $t \in (3, 6)$ sowie (b) $t \in (8, 10)$

	Modellproblem A	Modellproblem B	Modellproblem C
Räumliches Simulationsgebiet Ω	$([-1, 4] \times [-\frac{1}{2}, \frac{1}{2}]) \setminus B_{\frac{1}{5}}(0)$	$[-\pi, \pi] \times [-\frac{1}{2}, \frac{1}{2}]$	$[-\pi, \pi] \times [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$
Zeitintervall I	$[0, 20]$	$[0, 5]$	$[0, 10]$
Viskosität ν_0	$3 \cdot 10^{-4}$	$4.6225 \cdot 10^{-5}$	$1.5151 \cdot 10^{-4}$
Grobskalige Gitterzerlegung	198 Dreiecke	8192 Dreiecke	352 Simplicizes
Grobskalige Gitterweite	$0.03 \lesssim H \lesssim 0.6$	$0.03 \lesssim H \lesssim 0.1$	$H \approx 1.07$
Diskretisierungselemente	P_2/P_1	P_2/P_1	P_2/P_1
Grobskalige Freiheitsgrade	801	36831	1142
Feinskalige Gitterzerlegung	792 Dreiecke	32768 Dreiecke	2816 Simplicizes
Feinskalige Gitterweite	$0.015 \lesssim h \lesssim 0.15$	$0.015 \lesssim h \lesssim 0.05$	$h \approx 0.54$
Diskretisierungselemente	P_2/P_1	P_2/P_1	P_2/P_1
Feinskalige Freiheitsgrade	3386	147391	10408
Minimale Viskosität ν_{\min}	$3 \cdot 10^{-5}$	10^{-5}	10^{-5}
Maximale Viskosität ν_{\max}	$3 \cdot 10^{-1}$	10^{-2}	10^{-3}
Dämpfung d_{damp}	1	10^{-8}	0.1

Tabelle 6.1: Parameter der Modellprobleme

7 Modellierung der Stabilisierung als Viskositätsmodell

Die in den vorherigen Kapiteln vorgestellte Vorgehensweise hat den entscheidenden Nachteil, dass sie aufwendig ist. Das liegt daran, dass zur Generierung der künstlichen Viskosität ν^H insgesamt 5 variationelle Probleme gelöst werden müssen.

Für praktische Rechnungen ist es daher wünschenswert, eine Art „Abkürzung“ zu finden: Wir möchten gerne eine künstliche Viskosität ν^H zur Stabilisierung der groben Skalen verwenden, ohne jedes Mal die zusätzlichen Probleme (insbesondere das inverse Problem) lösen zu müssen.

Um dies zu erreichen wollen wir die Grundidee verfolgen, diese Abkürzung durch *maschinelles Lernen* herzustellen:

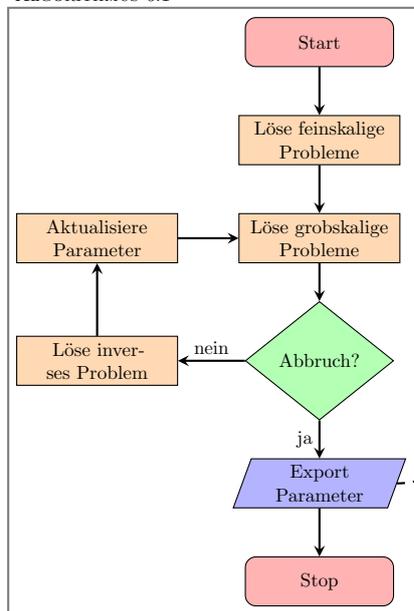
1. Zunächst wird in einer zeitlich unkritischen Umgebung die in dem vorausgegangenen Kapiteln vorgestellte Vorgehensweise angewendet. Dadurch erhalten wir Daten $(u^h, z^h, u_\nu^H, z_\nu^H, \nu^H)$, die wir zur späteren Verwendung speichern (*Trainingsdaten*).
2. Basierend auf den (aufwendig gewonnenen) Trainingsdaten wird nun eine Approximation m gesucht, die $m(u_\nu^H) \approx \nu^H$ für die Trainingsdaten möglichst gut erfüllt.
3. Diese Approximation wird schließlich auf die eigentliche Rechnung angewendet. In der Hoffnung, dass $m(u_\nu^H) \approx \nu^H$ auch für diese *Testdaten* gilt, haben wir dadurch mit m eine Abkürzung des aufwendigen Verfahrens gewonnen.

In Abbildung 7.1 ist diese Grundidee schematisch dargestellt. Die bisher betrachtete Vorgehensweise ist im Wesentlichen die Anwendung von Algorithmus 6.1. Da dessen Ausführung jedoch sehr aufwendig ist (vgl. Abschnitt 7.1), wird versucht lediglich den Algorithmus ganz rechts in der Grafik zu verwenden: Wir lösen das grobskalige Problem unter Verwendung einer Modellierung. Diese Modellierung ist bei der vorliegenden Betrachtung der Strömungssimulation konkret ein Viskositätsmodell.

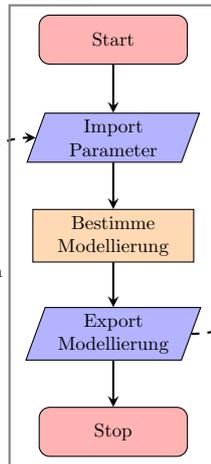
Der Kern dieses Kapitels beschäftigt sich damit, aus den Trainingsdaten eine passende Modellierung zu gewinnen. Dieser Prozess stellt gewissermaßen das Bindeglied zwischen dem bereits betrachteten (aufwendigen) Verfahren (Algorithmus 6.1) sowie dem angestrebten Verfahren. Die Grundüberlegung der Aufwandsreduktion ist dabei, dass sowohl der Algorithmus des inversen Problems wie auch jener zur Bestimmung einer Modellierung als Vorbereitung zu verstehen sind, d. h. sie können etwa im Vorfeld durchgeführt werden. Hier können insbesondere auch Vereinfachungen gegenüber dem tatsächlichen Problem vorgenommen werden, etwa eine Reduzierung des betrachteten Zeitintervalls. Außerdem ist es denkbar, eine einmal gewonnene Modellierung für andere ähnliche Probleme wiederzuverwenden. Lediglich der Algorithmus zur Anwendung der gewonnenen Modellierung trägt dann zur eigentlichen Laufzeit bei, da hier das komplette zu betrachtende Problem behandelt wird.

Der Fokus soll an dieser Stelle jedoch nicht bei geeigneten Vereinfachungen oder Möglichkeiten

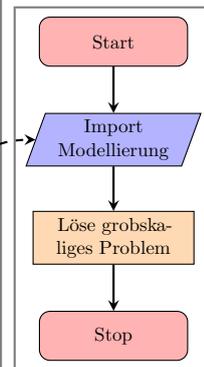
ALGORITHMUS 6.1



KAPITEL 7



Vorbereitung



Laufzeit

Abbildung 7.1: Schematische Darstellung der Grundidee zur Gewinnung eines Viskositätsmodells

zur Wiederverwendung liegen, sondern bei der Demonstration der grundsätzlichen Vorgehensweise. Insofern werden wir für jedes betrachtete Laufzeit-Problem auch separate Trainingsdaten generieren, aus denen dann ein optimales Viskositätsmodell gewonnen werden kann.

7.1 Aufwandsbetrachtung

Bevor auf die Details eingegangen wird, betrachten wir zunächst den Aufwand bzw. die erwartete Aufwandsreduktion. Zur Berechnung der Trainingsdaten müssen 2 feinskalige Probleme (Primal und Dual) in X^h sowie $2k$ grobskalige Probleme (Primal, Dual) in X^H gelöst werden. Zusätzlich muss k mal das gekoppelte Problem gelöst werden. Dabei handelt es sich um ein grobskaliges Problem in N^H . $k \in \mathbb{N}$ beschreibt die Anzahl der Iterationen, die aufgrund der Abhängigkeiten der Lösungen nötig sind.

Wenn für (V^H, Q^H) ein Taylor-Hood-Ansatz gemacht wird, haben wir

$$\dim(\hat{V}^H) = \dim(V^H) + \dim(Q^H) \approx 2d \dim(Q^H) + \dim(Q^H) = (2d + 1) \dim(Q^H).$$

Auch für die feinen Skalen (V^h, Q^h) gilt dann

$$\dim(\hat{V}^h) \approx (2d + 1) \dim(Q^h).$$

Für eine Verfeinerung um genau eine Gittergröße zwischen Q^H und Q^h ergibt sich

$$\dim(Q^h) \approx 2^d \dim(Q^H).$$

Wenn außerdem $N^H = Q^H$ gewählt ist, ergeben sich

$$\begin{aligned} \text{dof}_{\text{full}} &= 2 \dim(\hat{V}^h) + k(2 \dim(\hat{V}^H) + \dim(N^H)) \\ &\approx 2(2d + 1) \dim(Q^h) + 2k(2d + 1) \dim(Q^H) + k \dim(N^H) \\ &= (2(2d + 1)2^d + 2k(2d + 1) + k) \dim(Q^H) \\ &= (d2^{d+2} + 2^{d+1} + k(4d + 3)) \dim(Q^H) \end{aligned}$$

als zu bestimmende Freiheitsgrade für einen Datensatz der Trainingsdaten. Der Aufwand für die Abkürzung hingegen ist

$$\begin{aligned} \text{dof}_{\text{appr}} &= \dim(\hat{V}^H) + \dim(N^H) \\ &\approx (2d + 1) \dim(Q^H) + \dim(N^H) \\ &= 2(d + 1) \dim(Q^H). \end{aligned}$$

Hierbei wird das Auswerten der Approximation m mit dem Aufwand $\dim(N^H)$ berücksichtigt. Im Vergleich erhalten wir also

$$\frac{\text{dof}_{\text{full}}}{\text{dof}_{\text{appr}}} \approx \frac{(d2^{d+2} + 2^{d+1} + k(4d + 3))}{2(d + 1)} = 2^{d+1} + 2k - \frac{2^{d+1} + k}{2(d + 1)} \approx 2 \left(2^d + k \right).$$

Der Aufwand für das unstabilisierte grobskalige Verfahren hingegen ist

$$\text{dof}_{\text{unstab}} \approx \dim(\hat{V}^H) = (2d + 1) \dim(Q^H),$$

also im Vergleich

$$\frac{\text{dof}_{\text{appr}}}{\text{dof}_{\text{unstab}}} \approx \frac{2(d+1)}{2d+1} = 1 + \frac{1}{2d+1} \approx 1.$$

Der Mehraufwand des approximativen Verfahrens gegenüber dem Unstabilisierten ist also unwesentlich. Die Kosten zwischen dem vollständigen und dem approximierten Verfahren unterscheiden sich hingegen um einen Faktor von rund $2(2^d + k)$.

7.2 Modelltrainingsdaten

Die in den nächsten Abschnitten vorgestellten Modellierungsapproximationen benötigen Trainingsdaten als Grundlage. Unter Trainingsdaten verstehen wir eine Menge

$$P_{\text{train}} = \{(d_i, r_i) \in D \times R \mid i \in I_{\text{train}}\},$$

wobei I_{train} eine endliche Indexmenge bezeichnet. Jede Modellierungsapproximation liefert nun basierend auf den Trainingsdaten P_{train} eine Abbildung $\tilde{m}_i : D \rightarrow R$, die für die vorgegebenen Trainingsdaten gut geeignet ist, also

$$\tilde{m}(d_i) \approx r_i \quad \forall i \in I_{\text{train}}.$$

Um die in den Abschnitten 7.4 bis 7.6 vorgestellten Modellierungsapproximationen im Hinblick auf praktische Verwendbarkeit bewerten zu können, werden sie mit Trainingsdaten von drei beispielhaften Strömungsrechnungen numerisch erprobt. Insbesondere ist es wichtig, bei allen Approximationen dieselben Trainingsdaten zu verwenden, damit die Ergebnisse vergleichbar sind. Als zugrunde liegende Trainingsdaten sollen die bereits in Abschnitt 6.2 vorgestellten Rechnungen des inversen Problems dienen. Mit *Modelltrainingsdaten A* bezeichnen wir die Daten, die sich bei der Berechnung des inversen Problems aus Abschnitt 6.2.1 ergeben, mit *Modelltrainingsdaten B* entsprechend jene, die sich aus Abschnitt 6.2.2 ergeben, und schließlich mit *Modelltrainingsdaten C* jene, die wir in Abschnitt 6.2.3 bestimmt haben.

7.3 Räumlich globale bzw. lokale Regression

Während im zeitlichen Verlauf bereits aufgrund der Diskretisierung des primalen Problems nur eine sequentielle d. h. zeitlich lokale Vorgehensweise möglich ist, kann die Approximation räumlich betrachtet auf zwei verschiedenen Varianten erfolgen:

- Bei der räumlich *globalen* Regression $m_g : \hat{V}^H \rightarrow N^H$ liegt für jeden diskreten Zeitpunkt genau ein Datenpunkt (\hat{u}^H, ν^H) vor. Zur vereinheitlichten Notation ist dazu

$$\tilde{m}_g : D \rightarrow R$$

mit

$$\tilde{m}_g := m_g, \quad D := \hat{V}^H, \quad R := \nu^H.$$

- Bei der räumlich *lokalen* Regression $m_l : \hat{V}^H \rightarrow N^H$ werden für jeden Zeitpunkt auch verschiedene örtliche Positionen als einzelne Datenpunkte betrachtet. Dazu ist es sinnvoll, sich an den Freiheitsgraden $(x_i)_{i=1,\dots,N} \subseteq \Omega$ von N^H zu orientieren: Es wird eine Abtastfunktion $d : \hat{V}^H \times \Omega \rightarrow \mathbb{R}^k$ erklärt, die charakteristische Größen von $\hat{u}^H \in \hat{V}^H$ in $x \in \Omega$ bestimmt, etwa

$$d(\hat{u}^H, x) := \begin{pmatrix} \|u^H(x)\| \\ \|\nabla u^H(x)\| \\ p^H(x) \\ \|x\| \end{pmatrix}, \quad (7.1)$$

hier ist also $k = 4$. Die eigentliche Approximation erfolgt dann (lediglich mithilfe der zeit- und räumlich lokalen Daten) durch $\tilde{m}_l : \mathbb{R}^k \rightarrow \mathbb{R}$. In Analogie zu den Bezeichnungen bei der räumlich globalen Regression ist $D := \mathbb{R}^k$ und $R := \mathbb{R}$. Schließlich werden die zu den einzelnen Abtastpunkten $(x_i)_{i=1,\dots,N}$ gehörigen Daten wieder zu einer räumlich globalen Funktion in N^H zusammengefügt: $I_{N^H} : \mathbb{R}^N \rightarrow N^H$. Dies ist bei entsprechender Wahl der Abtastpunkte $(x_i)_{i=1,\dots,N}$ eine Interpolation bzw. kann bei passenden Ansatzfunktionen sogar der diskreten FEM-Darstellung entsprechen, so dass dazu keine weitere Rechnung nötig ist.

Insgesamt stellt sich die räumlich lokale Regression also wie folgt dar:

$$m_l(\hat{u}^H) := I_{N^H} \begin{pmatrix} \tilde{m}_l(d(\hat{u}^H, x_1)) \\ \tilde{m}_l(d(\hat{u}^H, x_2)) \\ \vdots \\ \tilde{m}_l(d(\hat{u}^H, x_N)) \end{pmatrix}.$$

Mit $m_i : \hat{V}^H \rightarrow N^H$, $i \in \{g, l\}$ haben wir zwei verschiedene Varianten der gesuchten Approximation beschrieben. Kernaufgabe des jeweiligen Verfahrens ist es, eine sinnvolle Regression $\tilde{m}_i : D \rightarrow R$, $i \in \{g, l\}$ zu finden. Dafür sind verschiedene Möglichkeiten denkbar. Sie werden in den nachfolgenden Abschnitten genauer betrachtet und verglichen.

Unabhängig von der konkreten Wahl von \tilde{m}_i , $i \in \{g, l\}$ haben die beiden vorgestellten Ansätze ihre prinzipbedingten Vor- und Nachteile.

Den räumlich globalen Algorithmen ist gemein, dass die Regression auf die Funktionsräume \hat{V}^H , N^H und damit insbesondere auf das zugrunde liegende Gitter fixiert ist. Bei adaptiven Verfahren jedoch ändern sich diese Funktionsräume. Damit anschließend die bereits gewonnene Regression noch zu verwenden ist, sind zusätzliche Anpassungen nötig.

Bei paralleler Rechnung kommt noch eine weitere Schwierigkeit hinzu: Bei den globalen Algorithmen gehen in die Regression Daten von allen Berechnungsknoten ein, folglich muss die Implementierung synchronisieren und Informationen zwischen allen Knoten austauschen. Dies kann einen zusätzlichen Flaschenhals darstellen und sollte unter Performance-Gesichtspunkten vermieden werden.

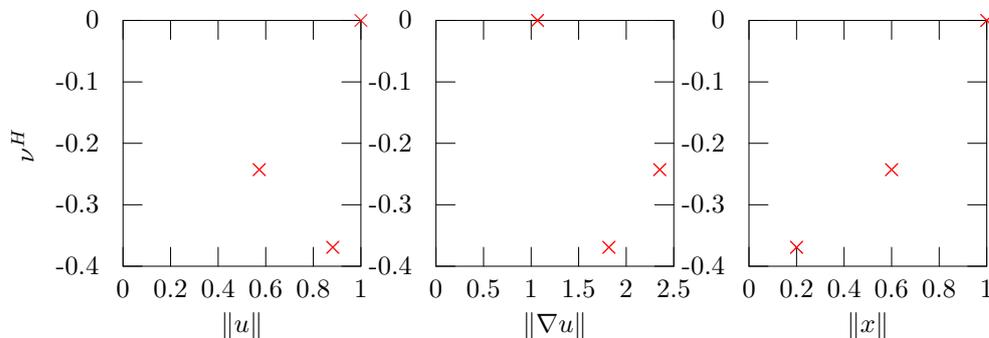
Auch die Übertragung einer einmal gewonnenen Approximation auf (hinreichend) ähnliche Probleme ist bei den globalen Regressionen nicht oder nur sehr bedingt möglich. Sie sind also

in dieser Hinsicht „starr“ und unflexibel.

Die räumlich lokale Regression hingegen hat den Nachteil, dass die in \tilde{m}_i steckende eigentliche Approximationsintelligenz „blind“ ist. Sie hat keinen Überblick, etwa über Ausprägung eines größeren Phänomens oder Entwicklungsstand des Systems als Ganzes. Sie kann nur die lokal in diesem Punkt vorherrschenden Größen verwenden, ohne Informationen von entfernten Punkten.

Je nach Absicht oder äußeren Gegebenheiten können also beide Arten der Regressionen sinnvoll sein. Es muss von Fall zu Fall entschieden werden, womit vernünftigerweise gearbeitet werden sollte.

Beispiel 7.3.1. (Fortführung von Beispiel 5.3.1, 5.3.2, 5.3.3) Bei dem Beispiel der gewöhnlichen Differentialgleichung ist die Aufgabe, eine geeignete lokale Regression zu finden, aufgrund der wenigen Freiheitsgrade relativ einfach.



Allerdings ist aus diesem Beispiel auch noch kein Indiz für einen allgemeingültigen Ansatz abzuleiten.

Eine globale Regression ist noch einfacher, macht aber weniger Sinn, da aufgrund fehlender zeitlicher Entwicklung des primalen Problems lediglich ein Datenpunkt zur Verfügung steht.

Unabhängig von der Wahl einer räumlich globalen oder lokalen Regression bleibt die Frage nach dem geeigneten Verfahren zum Finden von $\tilde{m}_i : D \rightarrow R, i \in \{g, l\}$. Dieser wollen wir in den nachfolgenden drei Abschnitten nachgehen.

7.4 Modellierung mittels reduzierter Basisfunktionen

Mithilfe der Technik der reduzierten Basen soll eine Regression $\tilde{m}_i : D \rightarrow R$ gefunden werden. Die Grundidee der reduzierten Basen ist es, eine Darstellung zu finden, bei der weniger wichtige Anteile schlicht weggelassen werden. Die Größen werden stattdessen mit weniger Informationen durch ihre wesentlichen, d.h. dominanten Anteile beschrieben.

7.4.1 Grundprinzip der reduzierten Basis

Die gewöhnliche Methode der reduzierten Basis beschreibt lediglich die Darstellung von Elementen $r_i \in R$. Es liegt dazu eine Trainingsmenge $R_{\text{train}} = \{r_i \in R \mid i \in I_{\text{train}}\}$ vor, wobei I_{train} eine endliche Indexmenge bezeichnet. Sie soll verwendet werden, um die Parameter einer

Darstellungsform

$$r \approx \sum_{j=1}^M \xi_j x_j^M(r)$$

zu bestimmen mit $\xi_j \in R$ und $x_j^M \in R^*$, konkret

$$x_j^M(r) = \sum_{k=1}^M c_{jk}^M \sigma_k(r),$$

wobei $c_{jk}^M \in R$ und $\sigma_k \in R^*$. $M \in \mathbb{N}$ ist dabei die Anzahl der neuen Basisfunktionen,

$$\Xi_M = \{\xi_m\}_{m=1}^M \subseteq R$$

die Menge der Basisfunktionen und

$$\Sigma_M = \{\sigma_m\}_{m=1}^M \subseteq R^*$$

die Menge der Funktionalen. Eingesetzt erhalten wir die Darstellung

$$r \approx \sum_{j,k=1}^M \xi_j c_{jk}^M \sigma_k(r).$$

Die Elemente $r_i \in R_{\text{train}}$ werden also gewissermaßen mit den Funktionalen $\sigma_k \in \Sigma_M$ abgetastet und mit den Basiselementen $\xi_j \in \Xi_M$ dargestellt. Dazwischen sorgen die Darstellungskoeffizienten $c_{jk}^M \in R$ für die passende Gewichtung.

Die Darstellungskoeffizienten $c_{jk}^M \in R$ für gegebene Mengen Ξ_M, Σ_M zu finden ist nicht besonders schwierig, die optimale Wahl läuft auf eine Matrixinversion der Dimension $M \times M$ hinaus:

$$\sum_{j=1}^M \sigma_i(\xi_j) c_{jk}^M = \delta_{ik} \quad \forall i, k = 1, \dots, M$$

mit

$$\delta_{ik} = \begin{cases} 1, & i = k \\ 0 & i \neq k. \end{cases}$$

Insbesondere ist zu beachten, dass c_{jk}^M unabhängig von einem Testelement $r_i \in R_{\text{train}}$ ist. Für die Wahl der eigentlichen Basiselemente bietet sich eine Greedy-Strategie an, wie in [DHO12]

$$\xi_{M+1} = r_{\max} - \sum_{j,k=1}^M \xi_j c_{jk}^M \sigma_k(r_{\max})$$

mit dem $r_{\max} \in R_{\text{train}}$, das bislang am schlechtesten dargestellt wurde und dem Funktional

$$\sigma_{M+1} = \operatorname{argsup}_{\sigma \in \Sigma} \{|\sigma(\xi_{M+1})|\},$$

das den größten Einfluss auf den Fehler hat. Hierbei ist Σ eine vorgegebene Menge von Funktionalen, die eine Basis für R^* bildet. Algorithmus 7.1 beschreibt die Kombination der

oben vorgestellten Schritte, dazugehörige Approximationsresultate sind Abbildung 7.2 zu entnehmen.

Algorithm 7.1 Klassische Reduktion der Basis von R

Input: $R_{\text{train}} \subseteq R, \varepsilon_{\text{tol}} > 0, M_{\text{max}} \in \mathbb{N}, \Sigma \subseteq R^*$

Output: $M \in \mathbb{N}_0, \{\sigma_i \mid i = 1, \dots, M\} \subseteq R^*, \{\xi_i \mid i = 1, \dots, M\} \subseteq R$

$M \leftarrow 0$

repeat

Berechne die Interpolationskoeffizienten $c_{jk}^M \in \mathbb{R}$ durch Lösung des linearen Gleichungssystems

$$\sum_{j=1}^M \sigma_j(\xi_j) c_{jk}^M = \delta_{ik} \quad \forall i, k = 1, \dots, M.$$

Finde die Approximation mit dem größten Fehler

$$r_{M+1} \leftarrow \operatorname{argsup}_{r \in R_{\text{train}}} \left\| r - \sum_{j=1}^M \xi_j c_{jk}^M \sigma_k(r) \right\|.$$

Berechne das Residuum $R_{M+1} \in R$ zwischen r_{M+1} und der derzeitigen Approximation

$$R_{M+1} \leftarrow r_{M+1} - \sum_{j=1}^M \xi_j c_{jk}^M \sigma_k(r_{M+1}).$$

Bestimme die Basisfunktion, die den größten Einfluss auf das Residuum hat

$$\sigma_{M+1} \leftarrow \operatorname{argsup}_{\sigma \in \Sigma} \{|\sigma(R_{M+1})|\}.$$

Normalisiere, um so eine zusätzliche Funktion zur Basis hinzuzufügen $\xi_{M+1} \leftarrow \frac{R_{M+1}}{\sigma_{M+1}(R_{M+1})}$.

Erhöhe die Anzahl der Approximationsterme $M \leftarrow M + 1$.

until $\|R_M\| < \varepsilon_{\text{tol}}$ **or** $M \geq M_{\text{max}}$

Zu beachten ist, dass für die so konstruierte Basis Ξ und Funktionale Σ gilt:

$$\sigma_k(\xi_j) = \delta_{jk} \quad 1 \leq k \leq j \leq M.$$

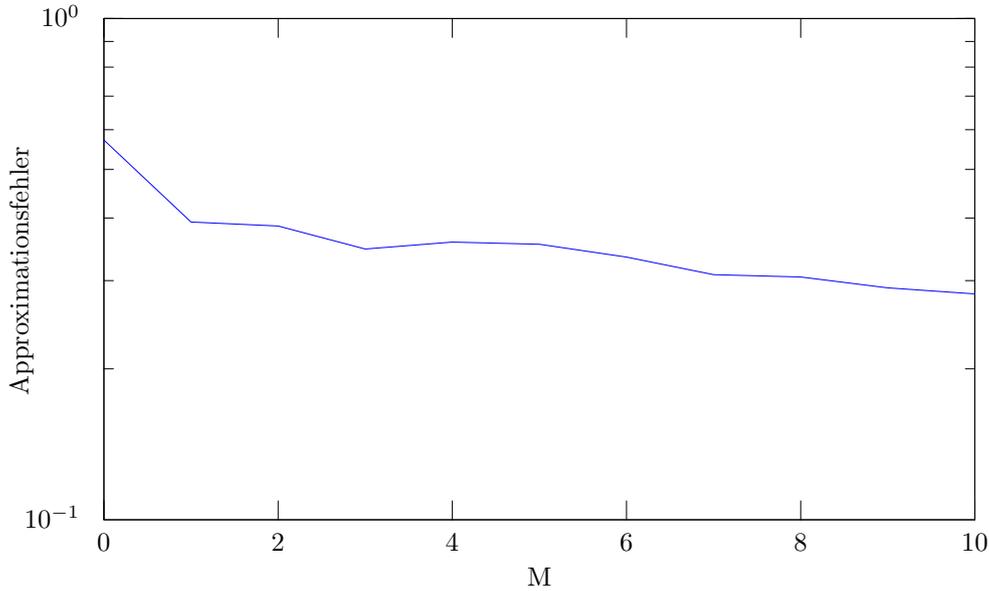


Abbildung 7.2: Approximationsfehler mit Algorithmus 7.1 (klassisch reduzierte Basis) bei synthetischen Daten, $|P_{\text{train}}| = 50$, $\dim R = 20$

7.4.2 Abbildung mittels reduzierter Basis

Das Ziel dieses Kapitels ist es, eine Approximation in der Form einer Abbildung zu finden. Das im vorangegangenen Abschnitt vorgestellte Verfahren kann zwar zur Reduktion der Basis verwendet werden, jedoch hilft es zunächst nicht, eine derartige Approximation zu finden.

Durch eine nun folgende Modifikation ist es jedoch möglich, das Grundprinzip an eine etwas geänderte Situation anzupassen:

Statt einer Darstellung $R \rightarrow R$ durch Reduktion der Basiselemente sind wir an einer Abbildung $\tilde{m}_i : D \rightarrow R$ interessiert.

Dazu liegt uns eine Trainingsmenge $P_{\text{train}} = \{(d_i, r_i) \in D \times R \mid i \in I_{\text{train}}\}$ vor. I_{train} bezeichne dabei eine endliche Indexmenge. Ziel ist es, $\xi_j \in R$, $c_{jk}^M \in \mathbb{R}$ und $\lambda_k \in D^*$ zu gewinnen, so dass die Darstellung

$$r \approx \sum_{j,k=1}^M \xi_j c_{jk} \lambda_k(d)$$

für Elemente $(d, r) = (d_i, r_i) \in P_{\text{train}}$ der Trainingsmenge gut ist, d. h.

$$\left\| \left(r_i - \sum_{j,k=1}^M \xi_j c_{jk} \lambda_k(d_i) : i \in I_{\text{train}} \right) \right\|$$

klein ist. Im Unterschied zum Grundprinzip ist statt Σ_M hier

$$\Lambda_M = \{\lambda_m\}_{m=1}^M \subseteq D^*$$

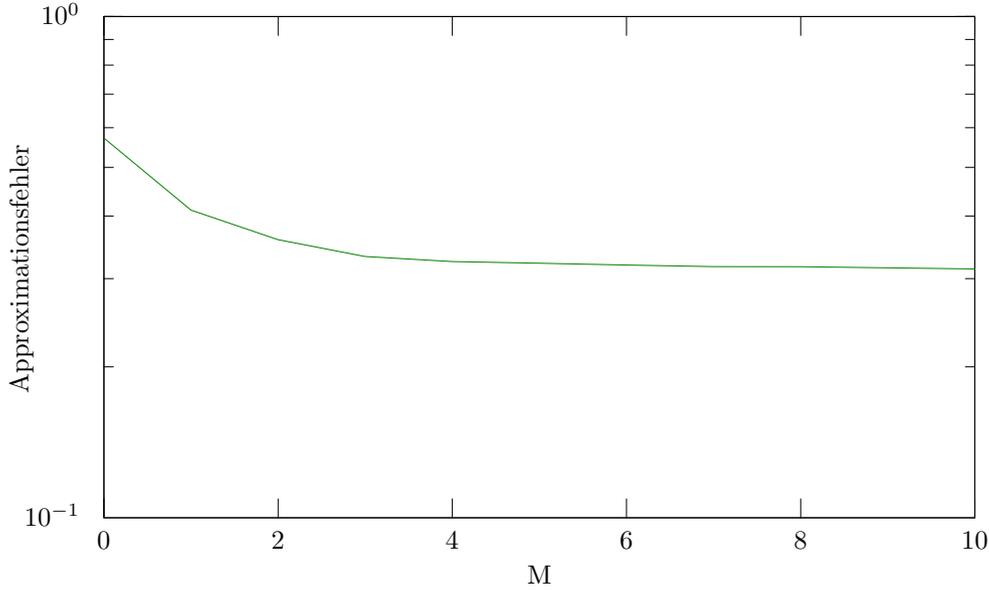


Abbildung 7.3: Approximationsfehler mit Algorithmus 7.2 (reduzierte Basis als Abbildung) bei synthetischen Daten, $|P_{\text{train}}| = 50$, $\dim D = 100$, $\dim R = 20$

zu bestimmen.

Da nun D^* und R nicht zueinander dual sind, lassen sich die Darstellungskoeffizienten c_{jk}^M nicht auf dem selben Weg bestimmen wie zuvor. Stattdessen versuchen wir,

$$\sum_{k=1}^M c_{jk}^M \lambda_k(d) = \xi_j^*(r) \quad \forall j = 1, \dots, M$$

zu lösen. Die so bestimmten c_{jk}^M hängen allerdings von $(d, r) \in P_{\text{train}}$ ab. Um eine Lösung c_{jk}^M zu finden, die alle Daten aus P_{train} gleichermaßen berücksichtigt, können wir die c_{jk}^M mittels Kleinst-Quadrat-Approximation gemäß

$$\sum_{j,k=1}^M \sum_{(d,r) \in P_{\text{train}}} \xi_i^*(\xi_j) c_{jk}^M \lambda_k(d) \lambda_m(d) = \sum_{(d,r) \in P_{\text{train}}} \xi_i^*(r) \lambda_m(d) \quad \forall i, m = 1, \dots, M$$

bestimmen.

Für die Wahl der Basiselemente $\xi_j \in R$ bzw. der Funktionale $\lambda_k \in D^*$ gehen wir grundsätzlich ähnlich vor wie zuvor gezeigt. Mit einer Greedy-Strategie wird der größte Einfluss auf den Fehler der Regression ausgewählt und als Grundlage für die neue Korrektur verwendet. Diese Übertragung wird in Algorithmus 7.2 beschrieben.

In Abbildung 7.3 sind Approximationseigenschaften bei Verwendung dieses Verfahrens auf aus Zufallszahlen generierten synthetischen Daten zu sehen. Es wurden dazu $|P_{\text{train}}| = 50$ Datensätze generiert, und es sollte eine Abbildung aus einem Raum mit Dimension $\dim D = 100$ in einen Raum mit Dimension $\dim R = 20$ gefunden werden.

Algorithm 7.2 Darstellung von r in Abhängigkeit von d mit reduzierter Basis

Input: $P_{\text{train}} \subset D \times R, \varepsilon_{\text{tol}} > 0, M_{\text{max}} \in \mathbb{N}, \Sigma \subset R^*$

Output: $M \in \mathbb{N}_0, \{\lambda_i \mid i = 1, \dots, M\} \subseteq D^*, \{\xi_i \mid i = 1, \dots, M\} \subseteq R$

$M \leftarrow 0$

repeat

Berechne die Regressionskoeffizienten $c_{jk}^M \in \mathbb{R}$ durch Lösung des linearen Gleichungssystems

$$\sum_{j,k=1}^M \sum_{(d,r) \in P_{\text{train}}} \xi_i^*(\xi_j) c_{jk}^M \lambda_k(d) \lambda_m(d) = \sum_{(d,r) \in P_{\text{train}}} \xi_i^*(r) \lambda_m(d) \quad \forall i, m = 1, \dots, M.$$

Finde die Approximation mit dem größten Fehler

$$(d_{M+1}, r_{M+1}) \leftarrow \operatorname{argsup}_{(d,r) \in P_{\text{train}}} \left\| r - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d) \right\|.$$

Berechne das Residuum $R_{M+1} \in R$ zwischen r_{M+1} und der derzeitigen Approximation

$$R_{M+1} \leftarrow r_{M+1} - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d_{M+1}).$$

Bestimme die Basisfunktion, die den größten Einfluss auf die zum größten Fehler gehörenden Daten hat

$$\lambda_{M+1} \leftarrow \operatorname{argsup}_{l \in \Sigma} \{|l(d_{M+1})|\}.$$

Normalisiere, um eine zusätzliche Funktion zur Basis hinzuzufügen $\xi_{M+1} \leftarrow \frac{R_{M+1}}{\lambda_{M+1}(d_{M+1})}$.

Erhöhe die Anzahl der Approximationsterme $M \leftarrow M + 1$.

until $\|R_M\| < \varepsilon_{\text{tol}}$ **or** $M \geq M_{\text{max}}$

Zur Charakterisierung der synthetischen Daten wurden ihre Konditionszahlen erfasst. Für die absolute Kondition ergibt sich:

$$0.23151 \leq \frac{\|r_1 - r_2\|}{\|d_1 - d_2\|} \leq 0.75261 \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}$$

und für die relative Kondition:

$$0.40947 \leq \frac{\frac{\|r_1 - r_2\|}{\|r_1\|}}{\frac{\|d_1 - d_2\|}{\|d_1\|}} \leq 1.6601 \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}.$$

Die Greedy-Strategie zur Gewinnung eines neuen Funktionals beim zuletzt vorgestellten Verfahrens ist zwar einfach, kann aber auch Probleme mit sich bringen. Die im folgenden vorgestellte Alternative bedeutet zwar mehr Aufwand, führt aber hoffentlich auch zu einer höheren Approximationsgüte.

Statt sich bei der Wahl des Funktionals nur auf eine vorgegebene Basis zu beschränken, können wir ein neues Funktional konstruieren, das etwa die bereits verwendeten Funktionalen berücksichtigt. Als Ausgangspunkt wollen wir uns dazu an dem bislang größten Approximationsfehler orientieren.

Konkret wollen wir zu den Daten d_{M+1} mit dem bislang größten Einfluss auf den Approximationsfehler das duale Element d_{M+1}^* betrachten. Dieses Funktional sollte durch Orthogonalisierung unabhängig von den bereits verwendeten Funktionalen gemacht werden. Der zentrale Schritt zur Konstruktion des neuen Funktionals λ_{M+1} unter Berücksichtigung der bereits ausgewählten Funktionalen $\lambda_k \in \Lambda_M$ sieht dann wie folgt aus:

$$\lambda_{M+1} \leftarrow d_{M+1}^* - \sum_{k=1}^M \lambda_k(d_{M+1}) \lambda_k.$$

Algorithmus 7.3 fasst das vorgestellte Verfahren zusammen, Ergebnisse sind in Abbildung 7.4 dargestellt.

Algorithm 7.3 Darstellung von r in Abhängigkeit von d mit reduzierter Basis durch Orthonormalisierung des kanonischen Duals

Input: $P_{\text{train}} \in D \times R, \varepsilon_{\text{tol}} > 0, M_{\text{max}} \in \mathbb{N}$

Output: $M \in \mathbb{N}_0, \{\lambda_i \mid i = 1, \dots, M\} \subseteq D^*, \{\xi_i \mid i = 1, \dots, M\} \subseteq R$

$M \leftarrow 0$

repeat

Berechne die Regressionskoeffizienten $c_{jk}^M \in \mathbb{R}$ durch Lösung des linearen Gleichungssystems

$$\sum_{j,k=1}^M \sum_{(d,r) \in P_{\text{train}}} \xi_j^* c_{jk}^M \lambda_k(d) \lambda_m(d) = \sum_{(d,r) \in P_{\text{train}}} \xi_j^*(r) \lambda_m(d) \quad \forall i, m = 1, \dots, M.$$

Finde die Approximation mit dem größten Fehler

$$(d_{M+1}, r_{M+1}) \leftarrow \operatorname{argsup}_{(d,r) \in P_{\text{train}}} \left\| r - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d) \right\|.$$

Berechne das Residuum $R_{M+1} \in R$ zwischen r_{M+1} und der derzeitigen Approximation

$$R_{M+1} \leftarrow r_{M+1} - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d_{M+1}).$$

Orthonormalisiere das duale Element der Daten zum größten Fehler bezüglich der bereits gewählten Basis

$$\begin{aligned} \hat{\lambda}_{M+1} &\leftarrow (d_{M+1})^* - \sum_{k=1}^M \lambda_k(d_{M+1}) \lambda_k, \\ \lambda_{M+1} &\leftarrow \frac{\tilde{\lambda}_{M+1}}{\|\hat{\lambda}_{M+1}\|}. \end{aligned}$$

Normalisiere, um so eine zusätzliche Funktion zur Basis hinzuzufügen $\xi_{M+1} \leftarrow \frac{R_{M+1}}{\lambda_{M+1}(d_{M+1})}$.

Erhöhe die Anzahl der Approximationsterme $M \leftarrow M + 1$.

until $\|R_M\| < \varepsilon_{\text{tol}}$ **or** $M \geq M_{\text{max}}$

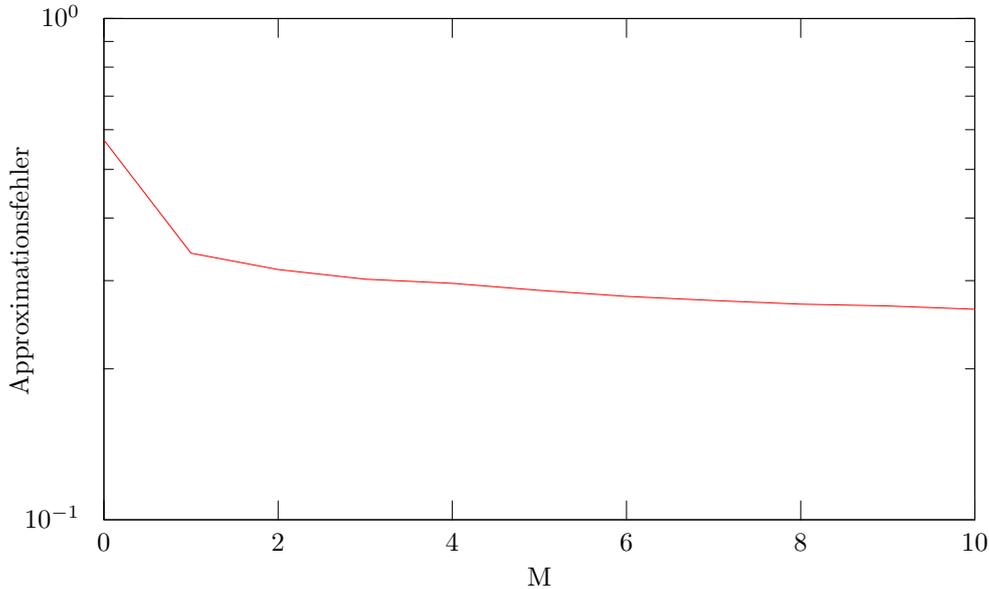


Abbildung 7.4: Approximationsfehler mit Algorithmus 7.3 (orthonormalisierte Funktionale) bei synthetischen Daten, $|P_{\text{train}}| = 50$, $\dim D = 100$, $\dim R = 20$

Es sind noch weitere Alternativen zur Gewinnung der Basiselemente sowie der Funktionale denkbar. Ein wiederum deutlich aufwendigeres Verfahren ist es, zur linearen Regression der jeweiligen Fehler in der Approximation $r - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d)$ auf die dazugehörigen Daten d , $(d, r) \in P_{\text{train}}$, die größten Einflüsse zu bestimmen.

Diese größten Einflüsse können durch eine Singulärwertzerlegung (engl. *singular value decomposition*, SVD) der zugehörigen Koeffizientenmatrix bestimmt werden. Da jedoch eine solche komplette Singulärwertzerlegung sehr aufwendig ist und wir nur an dem größten Einfluss interessiert sind, können wir hier auch eine Iteration ähnlich der Potenzmethode zur Bestimmung des größten Eigenwertes verwenden.

Sei etwa eine Matrix $C \in \mathbb{R}^{n \times m}$ sowie einen Vektor $r_1 \in \mathbb{R}^n$ vorgegeben, dann können wir durch

$$\hat{r}_{i+1} \leftarrow CC^{\top} r_i, \quad r_{i+1} \leftarrow \frac{\hat{r}_i}{\|\hat{r}_i\|} \quad \forall i = 1, 2, \dots$$

eine Iteration durchführen, bis $\|r_{i+1} - r_i\| < \varepsilon$. Setzen wir nun noch $s \leftarrow \frac{C^{\top} r_i}{\|C^{\top} r_i\|}$, dann erhalten wir die Approximation

$$C \approx r_i \sigma_1 s^{\top},$$

wobei σ_1 eine Annäherung an den größten Singulärwert von C ist.

Übertragen auf eine Abbildung C_{M+1} sowie das Funktional S_{M+1} und die Basis R_{M+1} ist diese Vorgehensweise in Algorithmus 7.4 zusammengefasst. Es ist zu beachten, dass nicht der Aufwand einer vollständigen Singulärwertzerlegung nötig ist. Allerdings muss zunächst die lineare Regression C_{M+1} (für jedes M) bestimmt werden, was bereits hinreichend aufwendig sein kann. Folglich hat der Algorithmus 7.4 bereits eine deutlich höhere Laufzeit als die zuvor vorgestellten Verfahren. Bemerkenswert ist, dass als Teilschritt insbesondere eine exakte

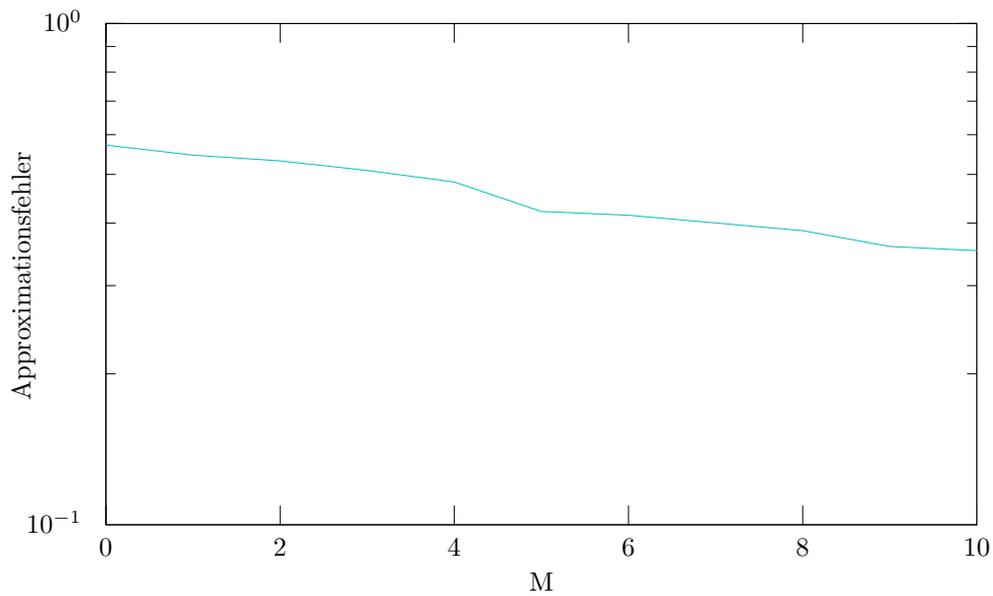


Abbildung 7.5: Approximationsfehler mit Algorithmus 7.4 (näherungsweise Singulärwertzerlegung) bei synthetischen Daten, $|P_{\text{train}}| = 50$, $\dim D = 100$, $\dim R = 20$

lineare Regression $C_1: D \rightarrow R$ bestimmt wird. Ergebnisse der Approximation können der Abbildung 7.5 entnommen werden.

Algorithm 7.4 Darstellung von r in Abhängigkeit von d mit reduzierter Basis durch näherungsweise Singulärwertbestimmung

Input: $P_{\text{train}} \subseteq D \times R, \varepsilon_{\text{tol}} > 0, M_{\text{max}} \in \mathbb{N}$

Output: $M \in \mathbb{N}_0, \{\lambda_i \mid i = 1, \dots, M\} \subseteq D^*, \{\xi_i \mid i = 1, \dots, M\} \subseteq R$

$M \leftarrow 0$

repeat

Berechne die Regressionskoeffizienten $c_{jk}^M \in \mathbb{R}$ durch Lösung des linearen Gleichungssystems

$$\sum_{j,k=1}^M \sum_{(d,r) \in P_{\text{train}}} \xi_i^*(\xi_j) c_{jk}^M \lambda_k(d) \lambda_m(d) = \sum_{(d,r) \in P_{\text{train}}} \xi_i^*(r) \lambda_m(d) \quad \forall i, m = 1, \dots, M.$$

Berechne die Regression $C_{M+1}: D \rightarrow R$ durch Lösen des linearen Gleichungssystems

$$C_{M+1}(d) = r - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d) \quad \forall (d, r) \in P_{\text{train}}.$$

Finde die Approximation mit derzeit größtem Fehler

$$(d_{M+1}, r_{M+1}) \leftarrow \operatorname{argsup}_{(d,r) \in P_{\text{train}}} \left\| r - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d) \right\|.$$

Berechne das Residuum $R_{M+1} \in R$ zwischen r_{M+1} und der derzeitigen Approximation

$$R_{M+1} \leftarrow r_{M+1} - \sum_{j,k=1}^M \xi_j c_{jk}^M \lambda_k(d_{M+1}).$$

repeat

Berechne $S_{M+1} \in D^*$ durch Anwendung des dualen Operators

$$\hat{S}_{M+1} \leftarrow C_{M+1}^*(R_{M+1}), \quad S_{M+1} \leftarrow \frac{\hat{S}_{M+1}}{\|\hat{S}_{M+1}\|}.$$

Berechne $R_{M+1} \in R$ aus $\hat{R}_{M+1} \leftarrow C_{M+1}(S_{M+1}), \quad R_{M+1} \leftarrow \frac{\hat{R}_{M+1}}{\|\hat{R}_{M+1}\|}.$

until $\|\hat{S}_{M+1} - C_{M+1}^*(R_{M+1})\| < \varepsilon_{\text{tol}}$

Verwende die so angenäherten Argumente für die Approximation

$$\lambda_{M+1} \leftarrow \hat{S}_{M+1}, \quad \xi_{M+1} \leftarrow R_{M+1}.$$

Erhöhe die Anzahl der Approximationsterme $M \leftarrow M + 1.$

until $\|R_M\| < \varepsilon_{\text{tol}}$ **or** $M \geq M_{\text{max}}$

Als Vergleich zu den vorgestellten Methoden der Basisreduktion soll nun noch eine vollständige Singulärwertzerlegung angeführt werden. Grundlage hierfür ist eine Matrix-Darstellung der linearen Abbildung $C_1: D \rightarrow R$ aus Algorithmus 7.4. Sofern die Räume D und R endlich sind, kann mithilfe einer Basis $\Lambda = \{\lambda_i \mid i = 1, \dots, \dim D\}$ von D^* sowie einer Basis $\Xi = \{\xi_i \mid i = 1, \dots, \dim R\}$ von R die Matrixdarstellung $A \in \mathbb{R}^{\dim R \times \dim D}$ von C_1 bestimmt werden, so dass

$$a_{ij} = \xi_i^*(C_1(\lambda_j^*)) \quad \forall i = 1, \dots, \dim R \quad \forall j = 1, \dots, \dim D.$$

Dann bezeichne

$$LSR^\top = A$$

die Singulärwertzerlegung von A , wobei $L \in \mathbb{R}^{\dim R \times \dim R}$ und $R \in \mathbb{R}^{\dim D \times \dim D}$ orthogonale Matrizen sowie $S \in \mathbb{R}^{\dim R \times \dim D}$ eine (nicht quadratische) Matrix mit Nichtnulleinträgen lediglich auf der Hauptdiagonalen, also $s_{ij} = 0$ für alle $i \neq j$. Insbesondere seien die Einträge von S der Größe nach sortiert, so dass $s_{11} \geq s_{22} \geq \dots s_{ii} \geq 0$.

Wir setzen nun $S^M \in \mathbb{R}^{\dim R \times \dim D}$ mit

$$s_{ij}^M = \begin{cases} s_{ij} & i \leq M \text{ und } j \leq M \\ 0 & \text{sonst} \end{cases}$$

und $M \in \mathbb{N}$, dann enthält S^M ausschließlich die ersten M Diagonaleinträge aus S und hat alle späteren Einträge auf 0 gesetzt. $LS^M R^\top$ ist folglich eine Approximation an A , wobei nur die größten M Komponenten berücksichtigt werden. Die ersten M Spalten von L beschreiben die dazugehörige Basis in R , die ersten M Spalten von R die dazugehörige Basis in D . Auch auf diese Weise kann die Abbildung $C_1: D \rightarrow R$ zerlegt werden.

Für die Praxis ist diese Zerlegung allerdings nur bei kleinen Daten brauchbar, da nicht nur die Regression C_1 , sondern auch eine vollständige Singulärwertzerlegung ihrer Matrixdarstellung berechnet werden muss. Für unsere synthetischen Daten hingegen können wir die so beschriebenen Approximationen bestimmen (siehe Abbildung 7.6) und als Vergleich hinzuziehen. Bei praktischen Betrachtungen stoßen wir mit dieser Vorgehensweise jedoch schnell auf Grenzen der Durchführbarkeit.

Beim Vergleich der verschiedenen Verfahren (Abbildung 7.7) fällt auf, dass für kleine M die Greedy-Strategie der Heuristiken die besseren Approximationen liefert, während für größere M die (angenäherte) Singulärwertzerlegung wesentlich bessere Ergebnisse erzielt. Insbesondere bei dem einfachsten Verfahren werden bei wachsendem M auch wesentliche Fehler erzeugt. Da die Singulärwertzerlegungen in der Praxis jedoch sehr aufwendig sind, startet Algorithmus 7.3 nach diesen Voruntersuchungen als Favorit für einen günstigen Kompromiss zwischen guten Ergebnissen bei akzeptablem Aufwand.

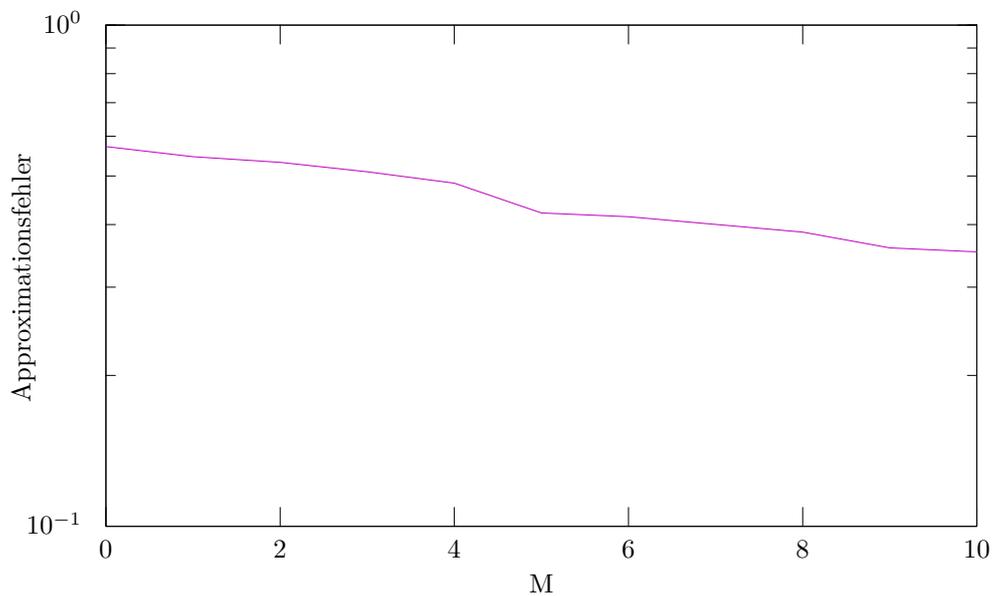


Abbildung 7.6: Approximationsfehler mit reduzierter Singulärwertzerlegung bei synthetischen Daten, $|P_{\text{train}}| = 50$, $\dim D = 100$, $\dim R = 20$

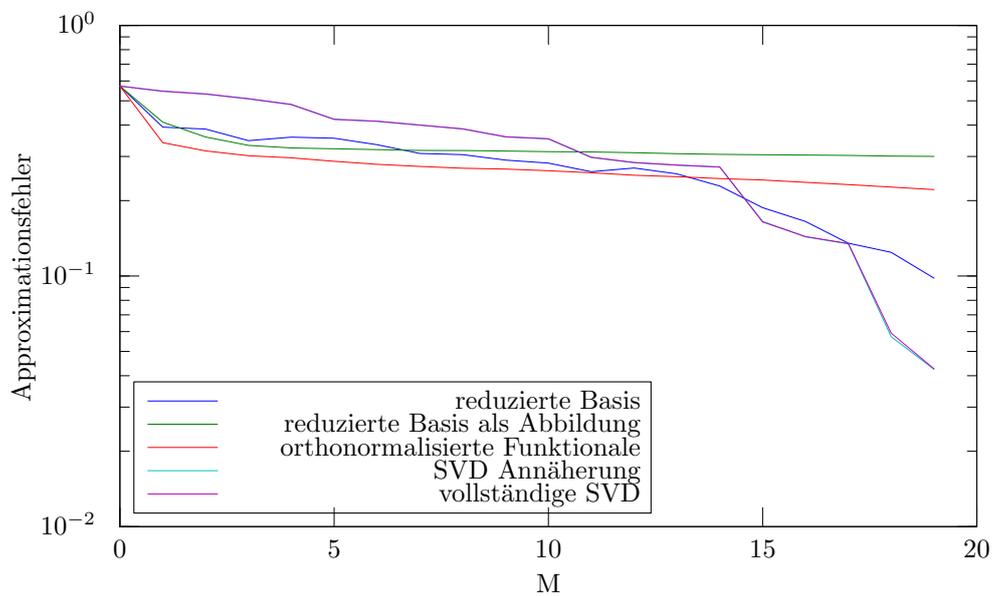


Abbildung 7.7: Vergleich der Approximationsfehler mit den verschiedenen Verfahren bei synthetischen Daten, $|P_{\text{train}}| = 50$, $\dim D = 100$, $\dim R = 20$ („SVD Annäherung“ und „vollständige SVD“ sind nahezu deckungsgleich)

7.4.3 Anwendung auf die Modelltrainingsdaten

Modelltrainingsdaten A

Auf die in Abschnitt 7.2 definierten Modelltrainingsdaten angewendet ergeben sich die beiden Möglichkeiten:

- Räumlich *lokale* Regression: Jeder Abtastpunkt wird als unabhängige Information zur Regression verwendet. Hierbei ist $\dim R = 1$, $\dim D = 7$ und wir haben Trainingsdaten im Umfang von $|P_{\text{train}}| = 127512$.
- Räumlich *globale* Regression: Alle Daten zu einem festen Zeitschritt werden als gebündelte Information zur Regression verwendet. Hierbei ist $\dim R = 792$, $\dim D = 5544$ und wir haben Trainingsdaten im Umfang von $|P_{\text{train}}| = 161$.

In Abbildung 7.8 sind die Ergebnisse dieser beiden Möglichkeiten dargestellt. Bei der globalen Regression (Abbildung 7.8(a)) schneidet wie erwartet das Verfahren der Singulärwert-Annäherung am besten ab. Sowohl des Verfahren der orthonormalisierten Funktionale als auch die reduzierte Basis als Abbildung scheinen nach einem brauchbaren Start keine wesentliche Verbesserung mehr durch Hinzunahme neuer Basisfunktionen erzielen zu können: Hier stagniert der Approximationsfehler trotz steigender Anzahl an Basiselementen. Lediglich die beiden außer Konkurrenz gestarteten Verfahren (reduzierte Basis sowie vollständige SVD) schneiden sehr gut ab.

Bei der lokalen Regression liefern die beiden Verfahren mit (angenäherter) Singulärwertzerlegung ein mäßige Approximation, die beiden Verfahren der orthonormalisierten Funktionale sowie die reduzierte Basis als Abbildung scheinen überhaupt keine Reduktion des Fehlers zu liefern. Lediglich das Verfahren der klassischen Basisreduktion ohne Abbildung erreicht bei $M = 1$ eine zufriedenstellende Approximation.

Insgesamt stellen wir fest, dass die Algorithmen auf den Modelltrainingsdaten A nicht ganz so gute Ergebnisse liefern, wie bei den synthetischen Daten.

Die Daten wurden daher zusätzlich auf Konditionszahlen untersucht. Bei räumlich globalen Daten ergibt sich für die absolute Kondition:

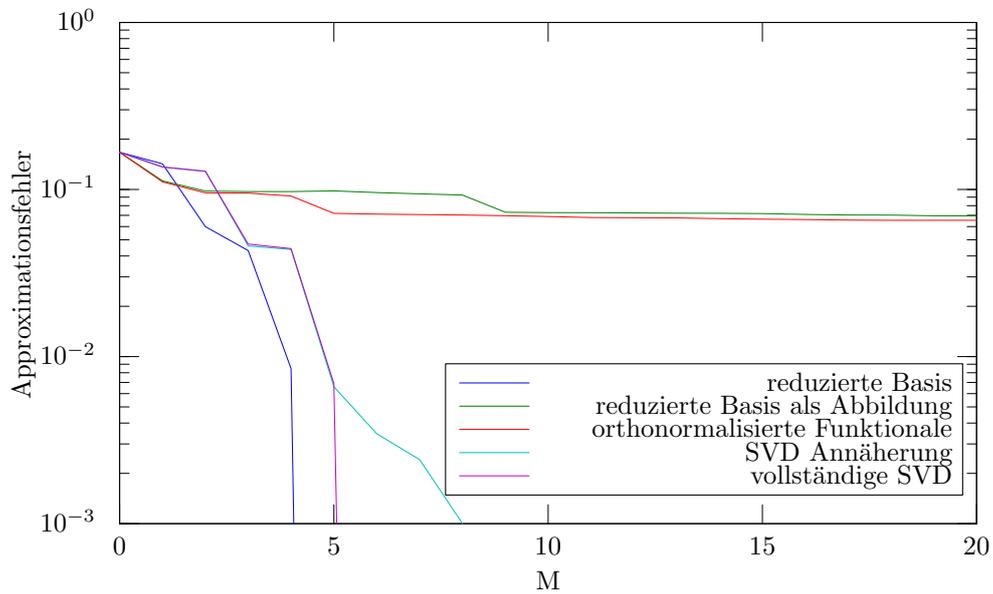
$$4.2066 \cdot 10^{-3} \leq \frac{\|r_1 - r_2\|}{\|d_1 - d_2\|} \leq 5.6022 \cdot 10^{-3} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}$$

und für die relative Kondition:

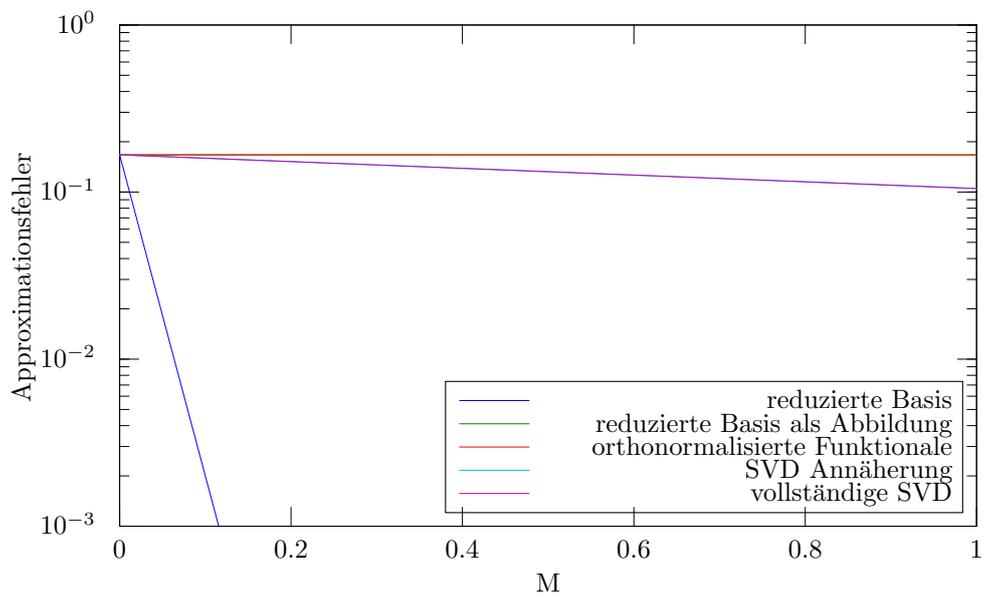
$$0.5101 \leq \frac{\frac{\|r_1 - r_2\|}{\|r_1\|}}{\frac{\|d_1 - d_2\|}{\|d_1\|}} \leq 0.70459 \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}.$$

Bei räumlich lokalen Daten ergibt sich für die absolute Kondition:

$$0 \leq \frac{\|r_1 - r_2\|}{\|d_1 - d_2\|} \leq 3.4018 \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}$$



(a)



(b)

Abbildung 7.8: Vergleich der Approximationsfehler bei (a) globaler bzw. (b) lokaler Regression der Modelltrainingsdaten A mit den verschiedenen Verfahren. Bei (b) liegen die Kurven „reduzierte Basis als Abbildung“ und „orthonormalisierte Funktionale“ sowie „SVD Annäherung“ und „vollständige SVD“ jeweils übereinander.

und für die relative Kondition:

$$0 \leq \frac{\frac{\|r_1 - r_2\|}{\|r_1\|}}{\frac{\|d_1 - d_2\|}{\|d_1\|}} \leq 9.955 \cdot 10^{50} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}.$$

Modelltrainingsdaten B

Auch bei den Modelltrainingsdaten B ergeben sich zwei Möglichkeiten:

- Räumlich *lokale* Regression: Jeder Abtastpunkt wird als unabhängige Information zur Regression verwendet. Hierbei ist $\dim R = 1$, $\dim D = 7$ und wir haben Trainingsdaten im Umfang von $|P_{\text{train}}| = 8814592$.
- Räumlich *globale* Regression: Alle Daten zu einem festen Zeitschritt werden als gebündelte Information zur Regression verwendet. Hierbei ist $\dim R = 32768$, $\dim D = 229376$ und wir haben Trainingsdaten im Umfang von $|P_{\text{train}}| = 269$.

Aufgrund der hohen Dimension von D bei der räumlich globalen Regression wird für die Abbildung C_1 bzw. ihre Matrixdarstellung A sehr viel Speicherplatz benötigt (rund 784 GB), so dass diese Analyse auf derzeit üblichen Rechnern nicht durchführbar war. Die beiden Verfahren, die auf der Singulärwertzerlegung dieser Matrix beruhen, wurden daher nur für die räumlich lokale Regression verwendet.

In Abbildung 7.9 sind die Ergebnisse dieser beiden Möglichkeiten dargestellt. Bei der globalen Regression liefert nur das Verfahren zur klassischen Basisreduktion gute Approximationseigenschaften, während die anderen beiden Verfahren ungefähr auf einem Approximationslevel bleiben. Auffällig ist, dass das Verfahren der orthonormalisierten Funktionale vereinzelt Approximationen liefert, die schlechter sind, als die Ausgangsdaten. Bei näherer Untersuchung der Ergebnisse stellt sich heraus, dass dies auf die Verwendung von (nahezu) linear abhängigen Basisfunktionen zurückzuführen ist. Dadurch ist die Wahl der Koeffizienten ein schlecht gestelltes Problem und es kommt gelegentlich zu solchen Ausreißern.

Auch bei der räumlich lokalen Regression werden keine guten Ergebnisse erzielt: Die Verfahren verhalten sich im Wesentlichen wie schon bei den Modelltrainingsdaten A gezeigt.

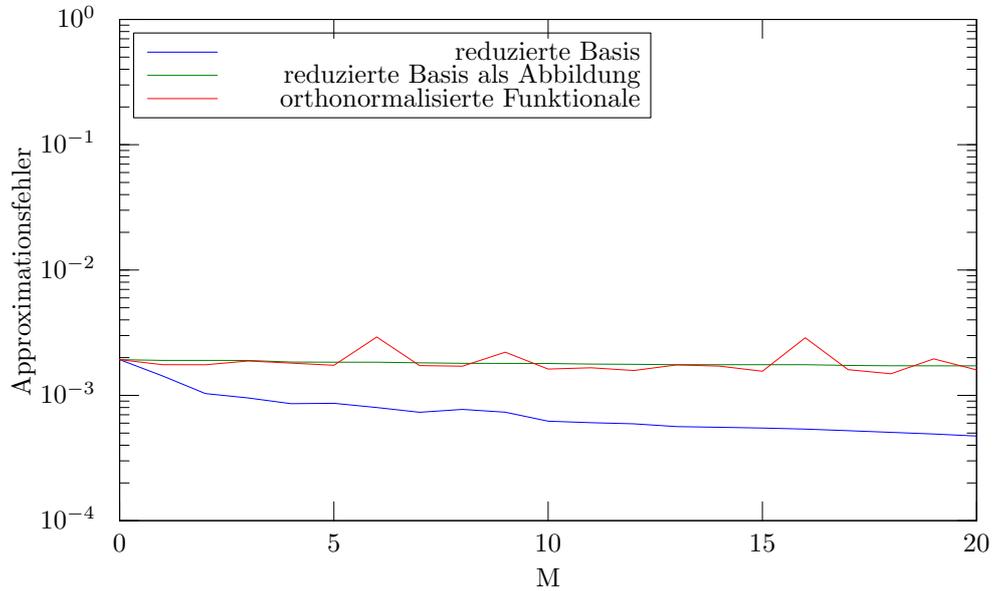
Auch hier müssen wir zusammenfassend festhalten, dass die Ergebnisse der Algorithmen unter den Erwartungen von den synthetischen Daten blieben.

Bei räumlich globalen Daten ergibt sich für die absolute Kondition

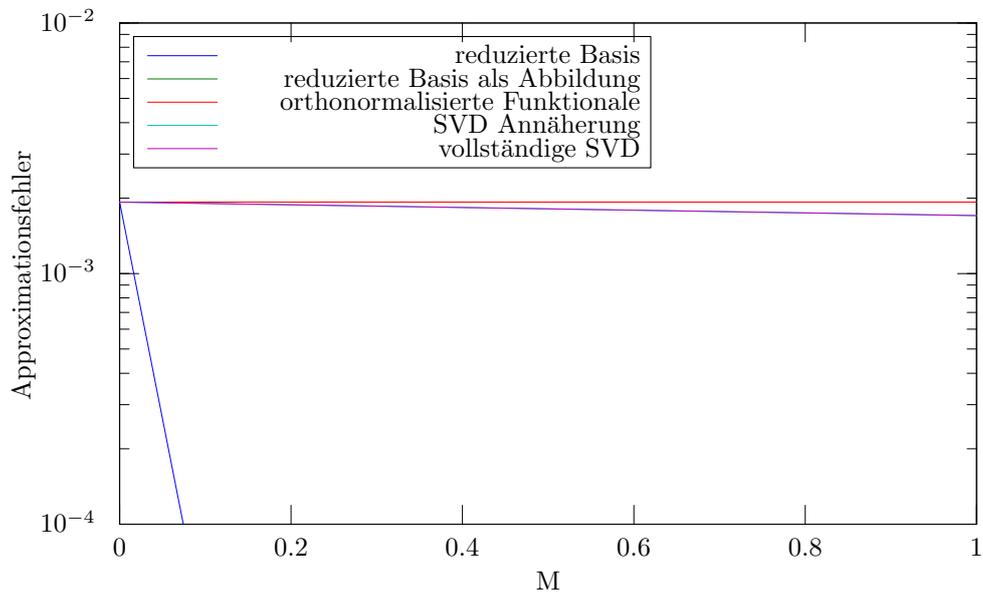
$$0 \leq \frac{\|r_1 - r_2\|}{\|d_1 - d_2\|} \leq 3.1937 \cdot 10^{-5} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}$$

und für die relative Kondition

$$0 \leq \frac{\frac{\|r_1 - r_2\|}{\|r_1\|}}{\frac{\|d_1 - d_2\|}{\|d_1\|}} \leq 1.154 \cdot 10^{49} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}.$$



(a)



(b)

Abbildung 7.9: Vergleich der Approximationsfehler bei (a) globaler bzw. (b) lokaler Regression der Modelltrainingsdaten B mit den verschiedenen Verfahren. Bei (b) liegen die Kurven „reduzierte Basis als Abbildung“ und „orthonormalisierte Funktionale“ sowie „SVD Annäherung“ und „vollständige SVD“ jeweils übereinander.

Bei räumlich lokalen Daten ergibt sich für die absolute Kondition

$$0 \leq \frac{\|r_1 - r_2\|}{\|d_1 - d_2\|} \leq 0.92848 \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}$$

und für die relative Kondition

$$0 \leq \frac{\frac{\|r_1 - r_2\|}{\|r_1\|}}{\frac{\|d_1 - d_2\|}{\|d_1\|}} \leq 6.791 \cdot 10^{49} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}.$$

Modelltrainingsdaten C

Auch bei den Modelltrainingsdaten C ergeben sich die beiden Möglichkeiten:

- Räumlich *lokale* Regression: Jeder Abtastpunkt wird als unabhängige Information zur Regression verwendet. Hierbei ist $\dim R = 1$, $\dim D = 13$ und wir haben Trainingsdaten im Umfang von $|P_{\text{train}}| = 1208064$.
- Räumlich *globale* Regression: Alle Daten zu einem festen Zeitschritt werden als gebündelte Information zur Regression verwendet. Hierbei ist $\dim R = 2816$, $\dim D = 36608$ und wir haben Trainingsdaten im Umfang von $|P_{\text{train}}| = 429$.

In Abbildung 7.10 sind die Ergebnisse beider Möglichkeiten dargestellt. Bei der räumlich lokalen Regression und diesen Daten zeigen die Algorithmen besser, dass sie zu gebrauchen sind: Wie auch schon bei den Modelltrainingsdaten A liefern die beiden Algorithmen der reduzierten Basis als Abbildung sowie später auch der orthonormalisierten Funktionale nach einigen (wenigen) Basisfunktionen keine Verbesserung der Approximation. Bis zu diesem Punkt sind ihre Approximationseigenschaften aber durchaus konkurrenzfähig im Vergleich zu den anderen Verfahren. Die Verfahren, die auf der Singulärwertzerlegung basieren, sowie die klassische Basisreduktion liefern hier eine gleichmäßige Reduktion des Approximationsfehlers und liefern vergleichbare Ergebnisse. Folglich liefert das Verfahren zur näherungsweise Singulärwertzerlegung in diesem Fall eine Lösung nah am Optimum (für die vorliegenden Daten).

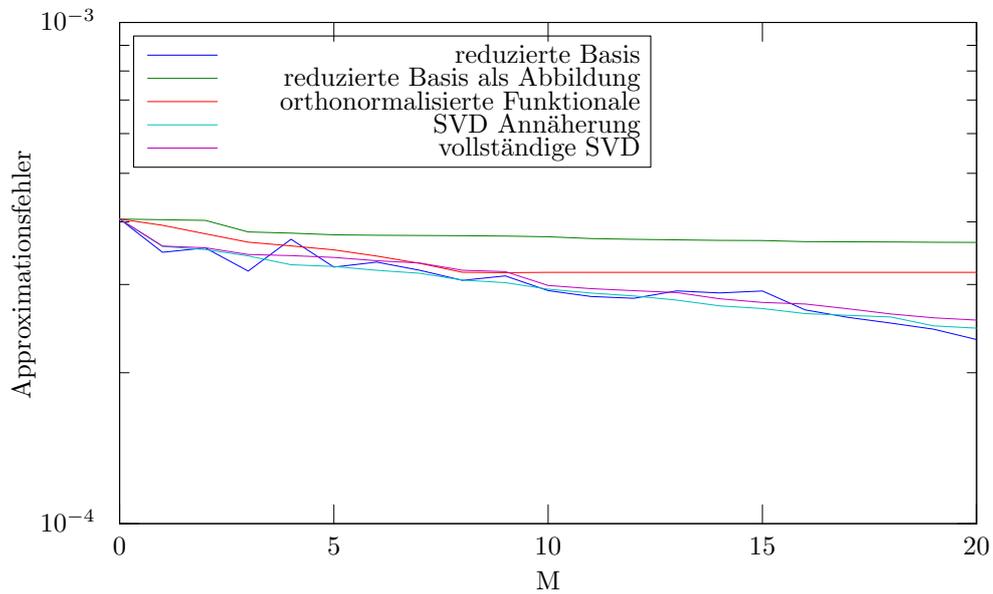
Für die räumlich lokale Regression ergibt sich wieder ein ähnliches Bild wie bereits bei den Modelltrainingsdaten A und B: Außer der klassischen Basisreduktion werden hier keine guten Ergebnisse erzielt.

Bei räumlich globalen Daten ergibt sich für die absolute Kondition

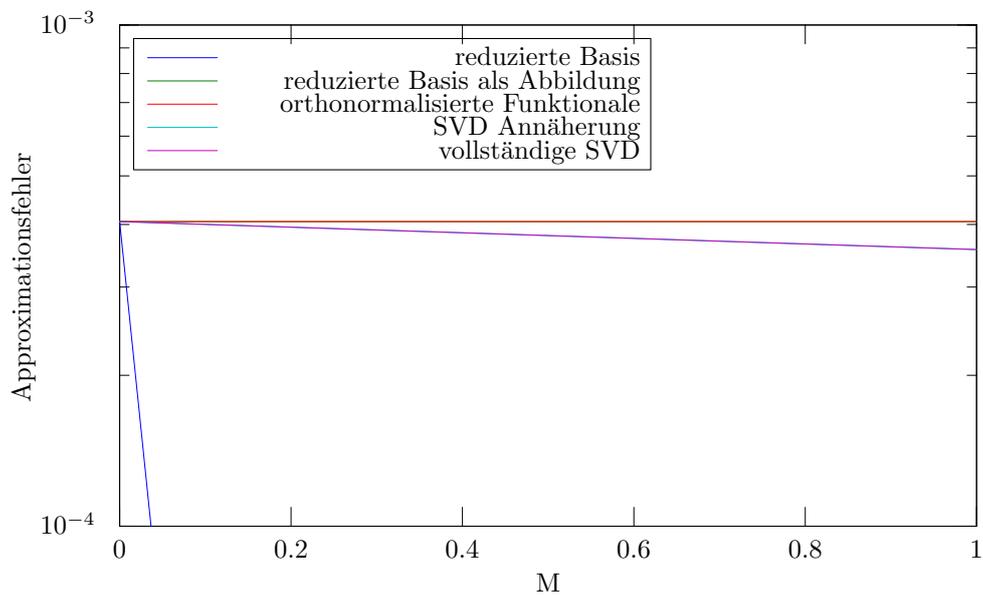
$$0 \leq \frac{\|r_1 - r_2\|}{\|d_1 - d_2\|} \leq 1.0092 \cdot 10^{-5} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}$$

und für die relative Kondition

$$0 \leq \frac{\frac{\|r_1 - r_2\|}{\|r_1\|}}{\frac{\|d_1 - d_2\|}{\|d_1\|}} \leq 4.35838 \cdot 10^{48} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}.$$



(a)



(b)

Abbildung 7.10: Vergleich der Approximationsfehler bei (a) globaler bzw. (b) lokaler Regression der Modelltrainingsdaten C mit den verschiedenen Verfahren. Bei (b) liegen die Kurven „reduzierte Basis als Abbildung“ und „orthonormalisierte Funktionale“ sowie „SVD Annäherung“ und „vollständige SVD“ jeweils übereinander.

Bei räumlich lokalen Daten ergibt sich für die absolute Kondition

$$0 \leq \frac{\|r_1 - r_2\|}{\|d_1 - d_2\|} \leq 1.3715 \cdot 10^{-3} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}$$

und für die relative Kondition

$$0 \leq \frac{\frac{\|r_1 - r_2\|}{\|r_1\|}}{\frac{\|d_1 - d_2\|}{\|d_1\|}} \leq 1.8309 \cdot 10^{48} \quad \forall (d_1, r_1), (d_2, r_2) \in P_{\text{train}}.$$

7.4.4 Zusammenfassung

Die vorgestellten Verfahren konnten die synthetischen Daten relativ gut in der gewünschten Form approximieren. Insbesondere bei kleinem M ergab sich neben dem Aufwandsvorteil eine höhere Approximationsgüte als bei der (angenäherten) Singulärwertzerlegung. Bei den praktischen Daten aus Strömungen der Modellprobleme zeigte sich ein etwas anderes Verhalten: Hier waren die Verfahren der reduzierten Basis als Abbildung oder der orthonormalisierten Funktionale nie (auch nicht für kleine M) besser als die anderen Verfahren. Bei Modelltrainingsdaten A und C konnten sie immerhin anfangs noch mithalten, dann fielen sie jedoch ab. Lediglich das Verfahren der angenäherten Singulärwertzerlegung lieferte durchweg brauchbare Ergebnisse, meist in der Nähe der vollständigen Singulärwertzerlegung. Allerdings ist es auch sehr aufwendig – die Modelltrainingsdaten B waren bereits eine zu große Herausforderung für diesen Algorithmus.

Die Idee, mithilfe einer räumlich globalen Regression eine höhere Approximationsgüte zu erzielen, war mit den vorgestellten Algorithmen bedingt erfolgreich. Zur räumlich lokalen Regression eigneten sie sich eher nicht. In dem Fall hingegen stehen weitere Werkzeuge zur Verfügung. Die bislang vorgestellten Verfahren sind bestenfalls eine lineare (nicht affin-lineare) Regression. Im folgenden Abschnitt sollen daher Verfahren der nichtlinearen lokalen Regression untersucht werden.

7.5 Modellierung mittels vorgegebener Ansatzfunktionen

Grundidee der Verfahren zur räumlich lokalen Regression dieses Abschnittes ist, dass eine von Parametern $b_0, \dots, b_k \in \mathbb{R}$ abhängige Abbildung $\tilde{m}_{l,b} : \hat{\mathbb{R}}^k \rightarrow \mathbb{R}$ vorgegeben wird. Anhand der Trainingsdaten werden anschließend die Parameter günstig eingestellt.

Wir beginnen mit einer klassischen linearen Regression, um anhand der auftretenden Residuen auf eine angepasste Regressionsgleichung zu gelangen. Zunächst wollen wir die *Approximation 1* mit

$$\tilde{m}(d) = b_0 + \sum_{i=1}^8 b_i d_i$$

testen, wobei

$$d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7 \\ d_8 \end{pmatrix} = \begin{pmatrix} u_1(t, x) \\ u_2(t, x) \\ \partial_1 u_1(t, x) \\ \partial_2 u_1(t, x) \\ \partial_1 u_2(t, x) \\ \partial_2 u_2(t, x) \\ p(t, x) \\ \text{dist}(\Gamma_{\text{obs}}, x) \end{pmatrix}.$$

Hierbei bezeichnet $\text{dist}(\Gamma_{\text{obs}}, x)$ den euklidischen Abstand zur Wand des Hindernisses, beim Modellproblem A ist also $\text{dist}(\Gamma_{\text{obs}}, x) = \max(0, \|x\| - 0.2)$. Bei Modellproblem B wird hier die Wand des Rohrs verwendet, also $\text{dist}(\Gamma_{\text{obs}}, x) = 0.5 - |x_2|$.

Offensichtlich bezieht sich die obige Definition lediglich auf zwei Raumdimensionen. Dies ist für die Modellprobleme A sowie B genau richtig. Bei Modellproblem C ist $\Omega \subset \mathbb{R}^3$, der Ansatz für *Approximation 1* ist entsprechend:

$$\tilde{m}(d) = b_0 + \sum_{i=1}^{14} b_i d_i,$$

wobei

$$d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7 \\ d_8 \\ d_9 \\ d_{10} \\ d_{11} \\ d_{12} \\ d_{13} \\ d_{14} \end{pmatrix} = \begin{pmatrix} u_1(t, x) \\ u_2(t, x) \\ u_3(t, x) \\ \partial_1 u_1(t, x) \\ \partial_2 u_1(t, x) \\ \partial_3 u_1(t, x) \\ \partial_1 u_2(t, x) \\ \partial_2 u_2(t, x) \\ \partial_3 u_2(t, x) \\ \partial_1 u_3(t, x) \\ \partial_2 u_3(t, x) \\ \partial_3 u_3(t, x) \\ p(t, x) \\ \text{dist}(\Gamma_{\text{obs}}, x) \end{pmatrix}.$$

Die Approximationsergebnisse sind in den Abbildungen 7.11 bis 7.13 dargestellt.

Die Approximation 1 ist relativ kanonisch, weil sie sämtliche zur Verfügung stehenden Daten in ihrer Rohform verwendet. Gleichzeitig ist sie sehr eingeschränkt, da diese Rohdaten eine wesentliche Charakteristik der Strömung enthalten. Insbesondere hatten wir bereits das Problem, dass sie in Abhängigkeit vom Modellproblem (bzw. der Dimension von Ω) definiert werden musste. Um etwa unabhängig von der Hauptströmungsrichtung zu werden und keinen problemabhängigen Modellierungsansatz verwenden zu müssen, soll nun statt der Komponenten u_1, u_2 (und im dreidimensionalen Fall u_3) die Norm $\|u\|$ verwendet werden, sowie $\|\nabla^s u\|$ statt $\partial_1 u_1, \partial_2 u_1, \partial_1 u_2, \partial_2 u_2$ (und entsprechend $\partial_1 u_3, \partial_2 u_3, \partial_3 u_1, \partial_3 u_2, \partial_3 u_3$). Wir suchen daher

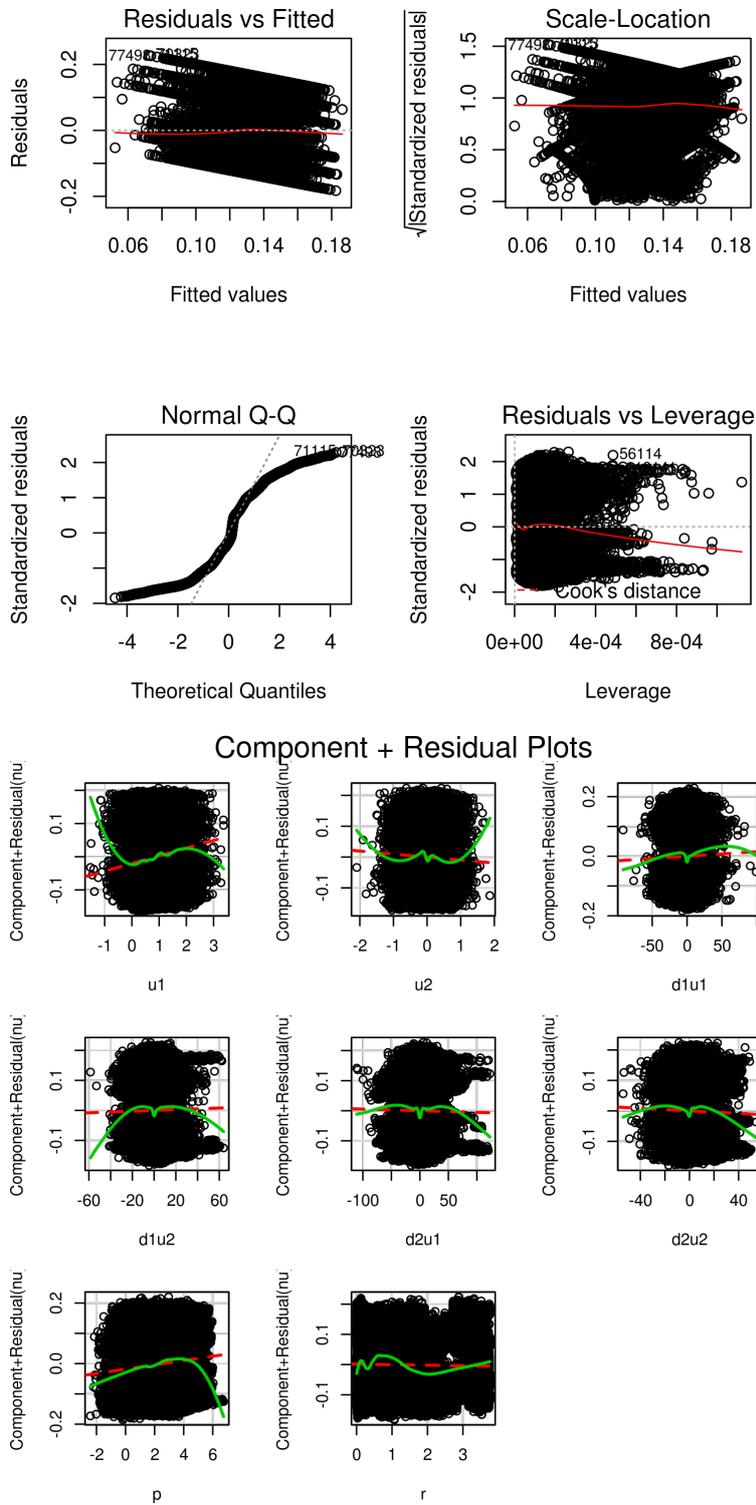


Abbildung 7.11: Approximation 1 der Modelltrainingsdaten A

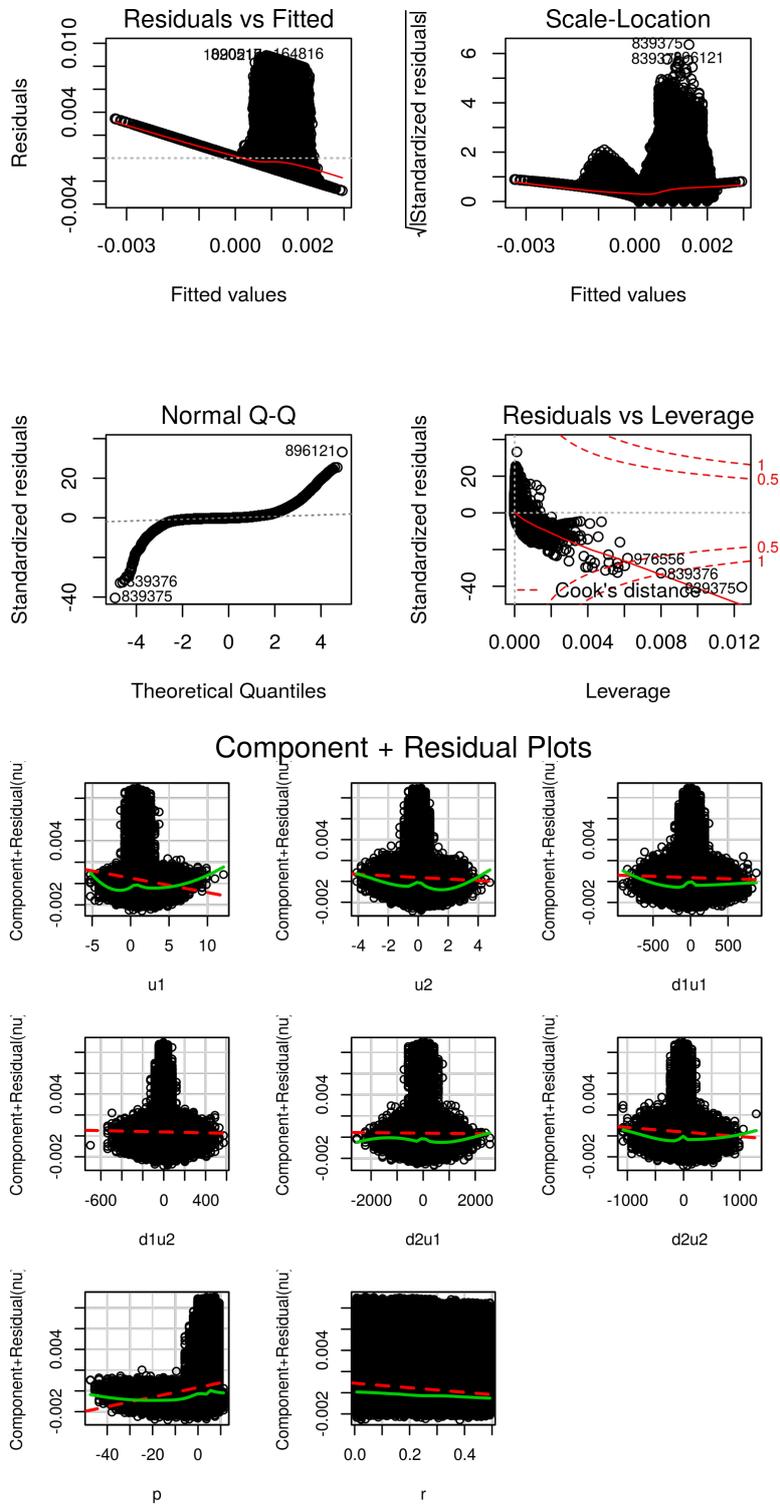


Abbildung 7.12: Approximation 1 der Modelltrainingsdaten B

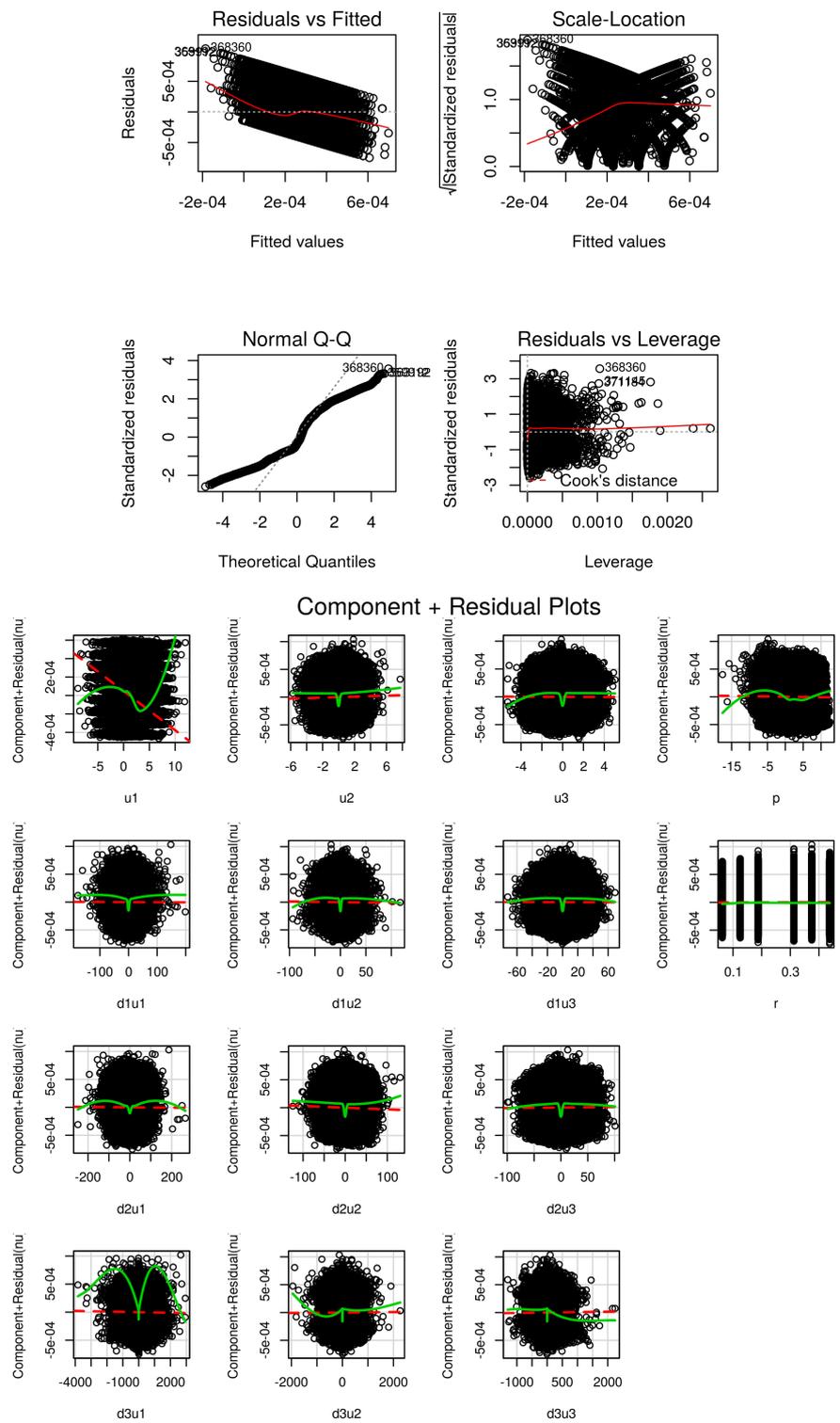


Abbildung 7.13: Approximation 1 der Modelltrainingsdaten C

eine *Approximation 2* mit

$$\tilde{m}(d) = b_0 + \sum_{i=1}^4 b_i d_i,$$

wobei

$$d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix} = \begin{pmatrix} \|u(t, x)\| \\ \|\nabla^s u(t, x)\| \\ p(t, x) \\ \text{dist}(\Gamma_{\text{obs}}, x) \end{pmatrix}.$$

Die Approximationsergebnisse sind in den Abbildungen 7.14 bis 7.16 dargestellt.

Die Approximationsgüte bei Approximation 2 verschlechtert sich nicht wesentlich gegenüber der Approximation 1. Im Gegenteil haben jetzt alle Koeffizienten b_i tatsächlich Relevanz und können aus den Daten sicherer ermittelt werden. Die zur Orientierung angegebene Lowess-Linie in den Plots (grün in den Abbildungen 7.14 bis 7.16) weicht von der Geraden der Regression meist nur wenig ab. Die einzige naheliegende Auffälligkeit hierbei ist das zunehmende Steigen bei größeren $\|\nabla^s u\|$. Um dieses Verhalten besser abbilden zu können suchen wir eine *Approximation 3* mit

$$\tilde{m}(d) = b_0 + \sum_{i=1}^5 b_i d_i,$$

wobei

$$d = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \end{pmatrix} = \begin{pmatrix} \|u(t, x)\| \\ \|\nabla^s u(t, x)\| \\ \|\nabla^s u(t, x)\|^2 \\ p(t, x) \\ \text{dist}(\Gamma_{\text{obs}}, x) \end{pmatrix}.$$

Die Approximationsergebnisse sind in den Abbildungen 7.17 bis 7.19 dargestellt.

Bemerkung 7.5.1. Die Approximation 3 ist aufgrund der Linearität in den Koeffizienten immer noch eine lineare Regression, aber aufgrund des polynomiellen Ansatzes in $\|\nabla^s u\|$ ist sie nichtlinear bezüglich der unabhängigen Größen. Dies führt dazu, dass die Komponenten-Plots von „norm_du“ und „norm_du.square“ in den Abbildungen 7.17 bis 7.19 nur bedingt isoliert betrachtet werden können, sondern in Kombination verstanden werden müssen.

Als Vergleich zu diesen eher datenorientierten Approximationsansätzen soll nun noch ein relativ etablierter Ansatz für Turbulenz-Modellierung verwendet werden: das Smagorinsky-Modell. Es ist mit einer einzigen Ansatzfunktion gegenüber den anderen Approximationen vergleichsweise einfach gehalten. Als *Approximation 4* oder *Smagorinsky-Modell* wurde der Ansatz

$$\tilde{m}(d) = b_1 d_1$$

gewählt, wobei

$$d = (d_1) = (\|\nabla^s u(t, x)\|^2)$$

und $b_1 \in \mathbb{R}$ wie bei den Approximationen 1-3 anhand der Modelltrainingsdaten optimal bestimmt wird.

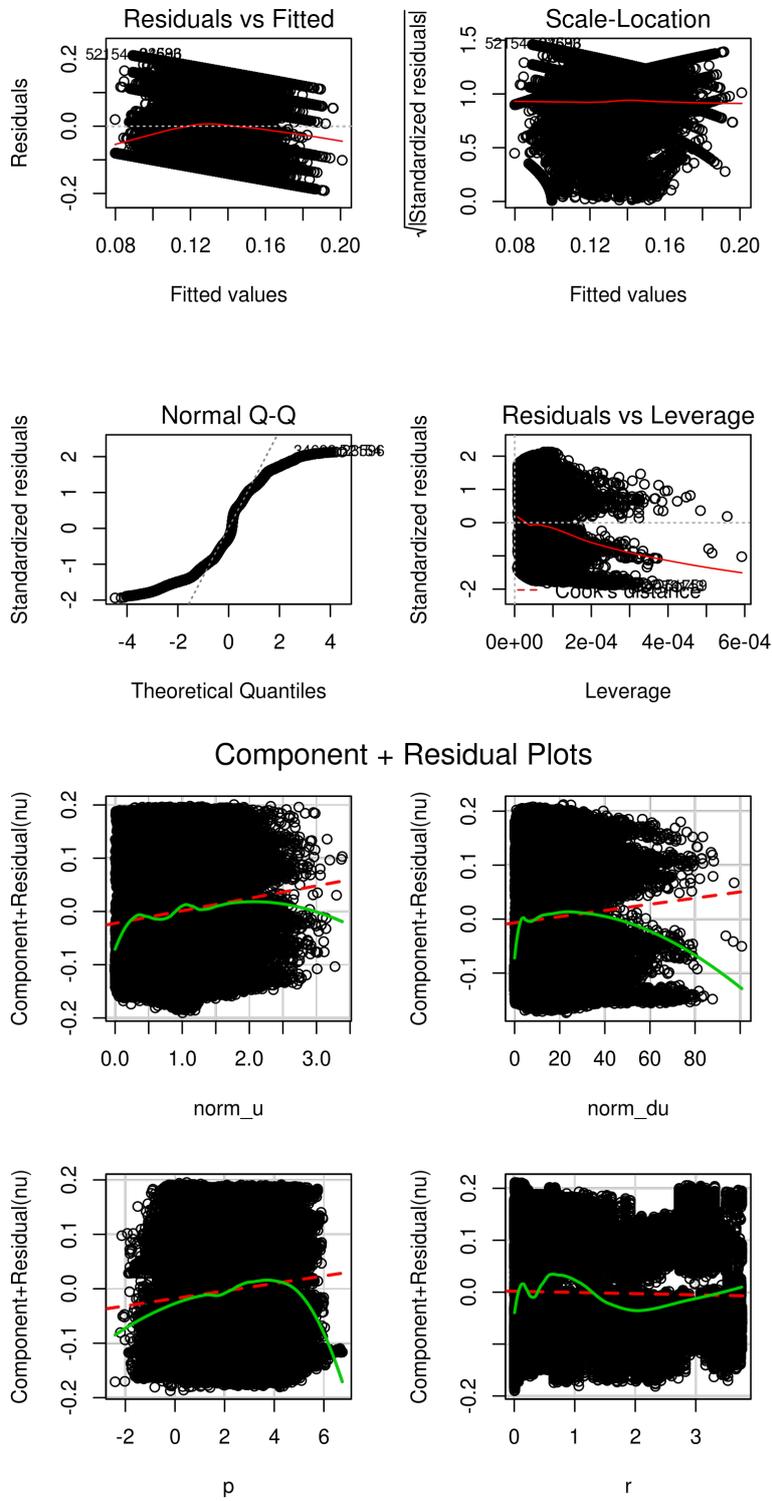


Abbildung 7.14: Approximation 2 der Modelltrainingsdaten A

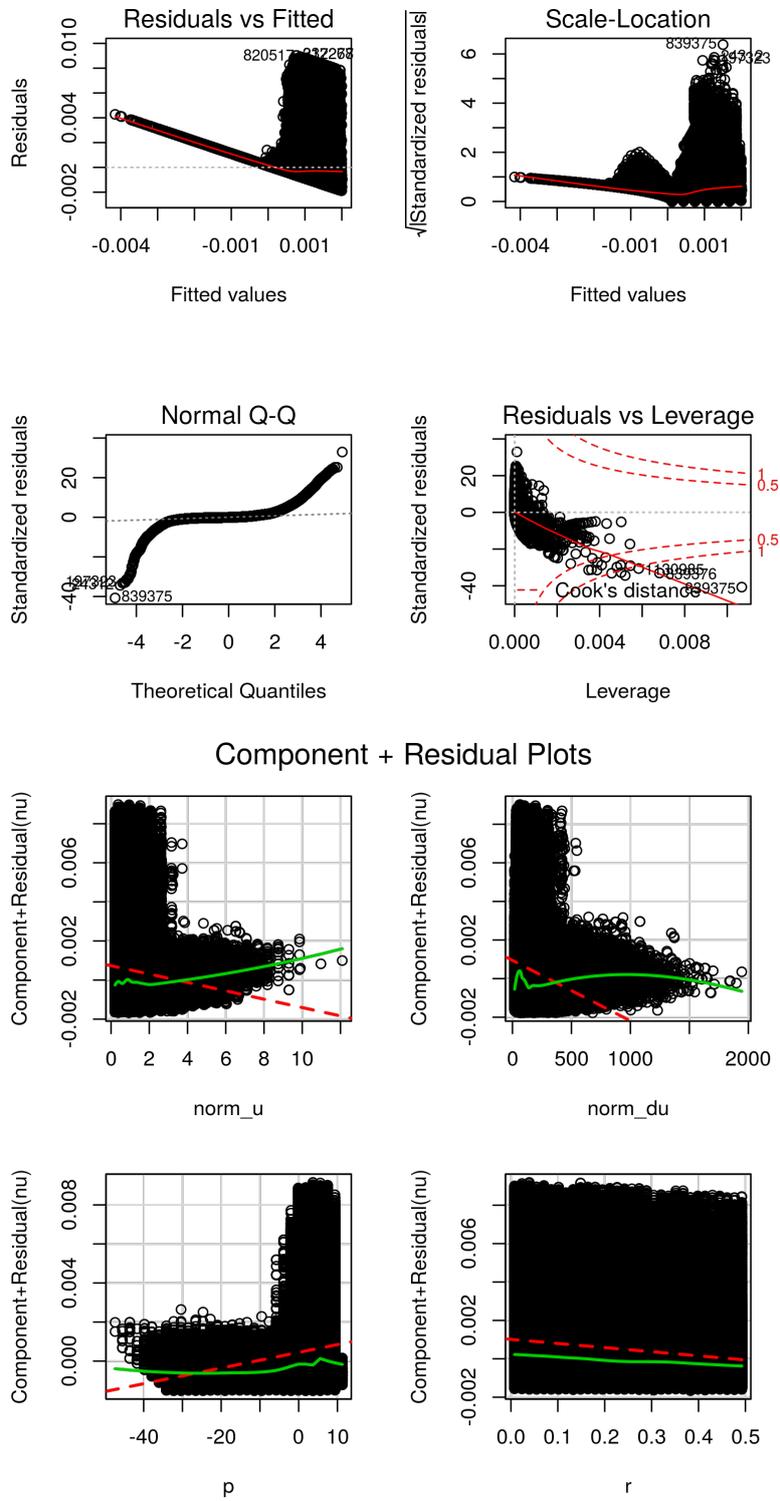


Abbildung 7.15: Approximation 2 der Modelltrainingsdaten B

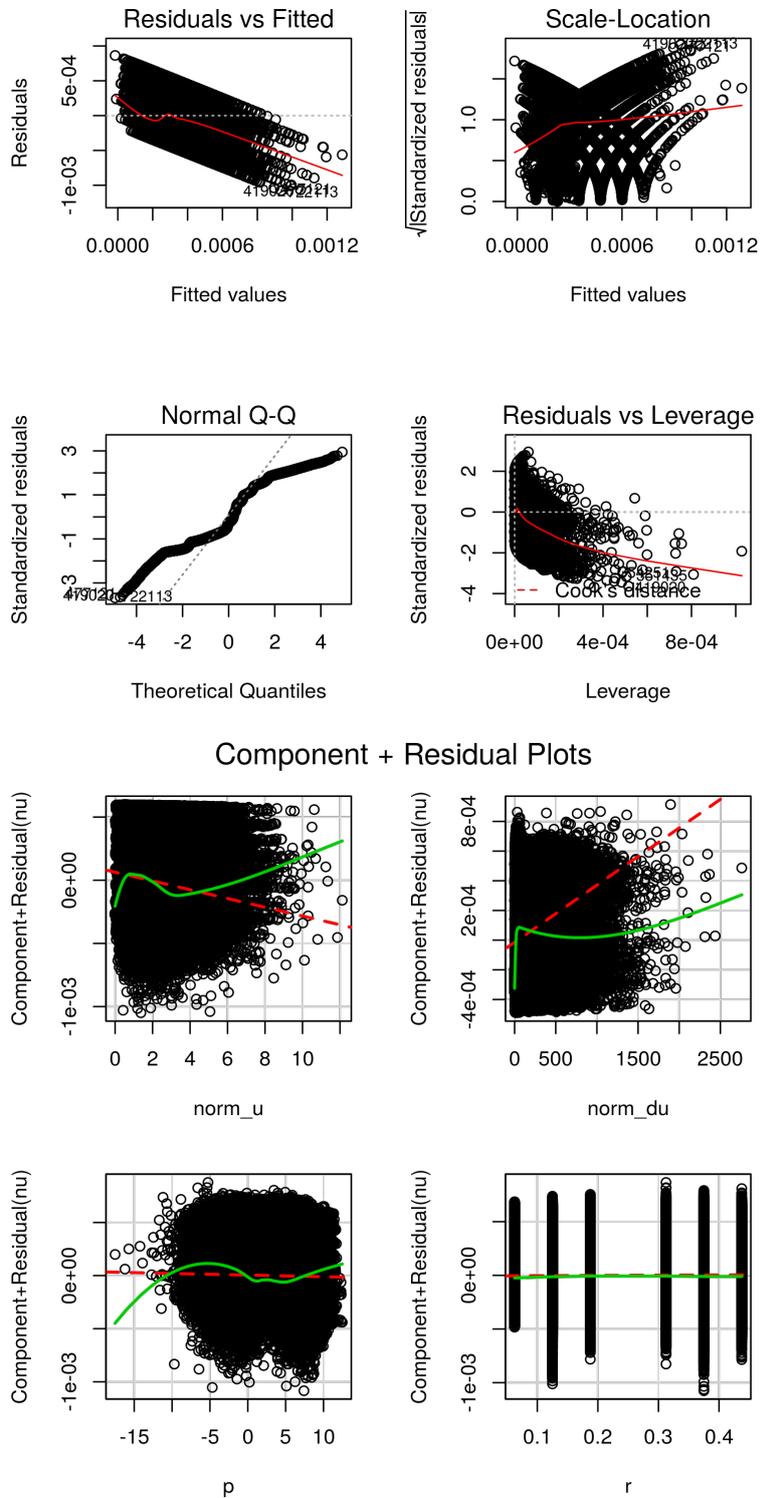


Abbildung 7.16: Approximation 2 der Modelltrainingsdaten C

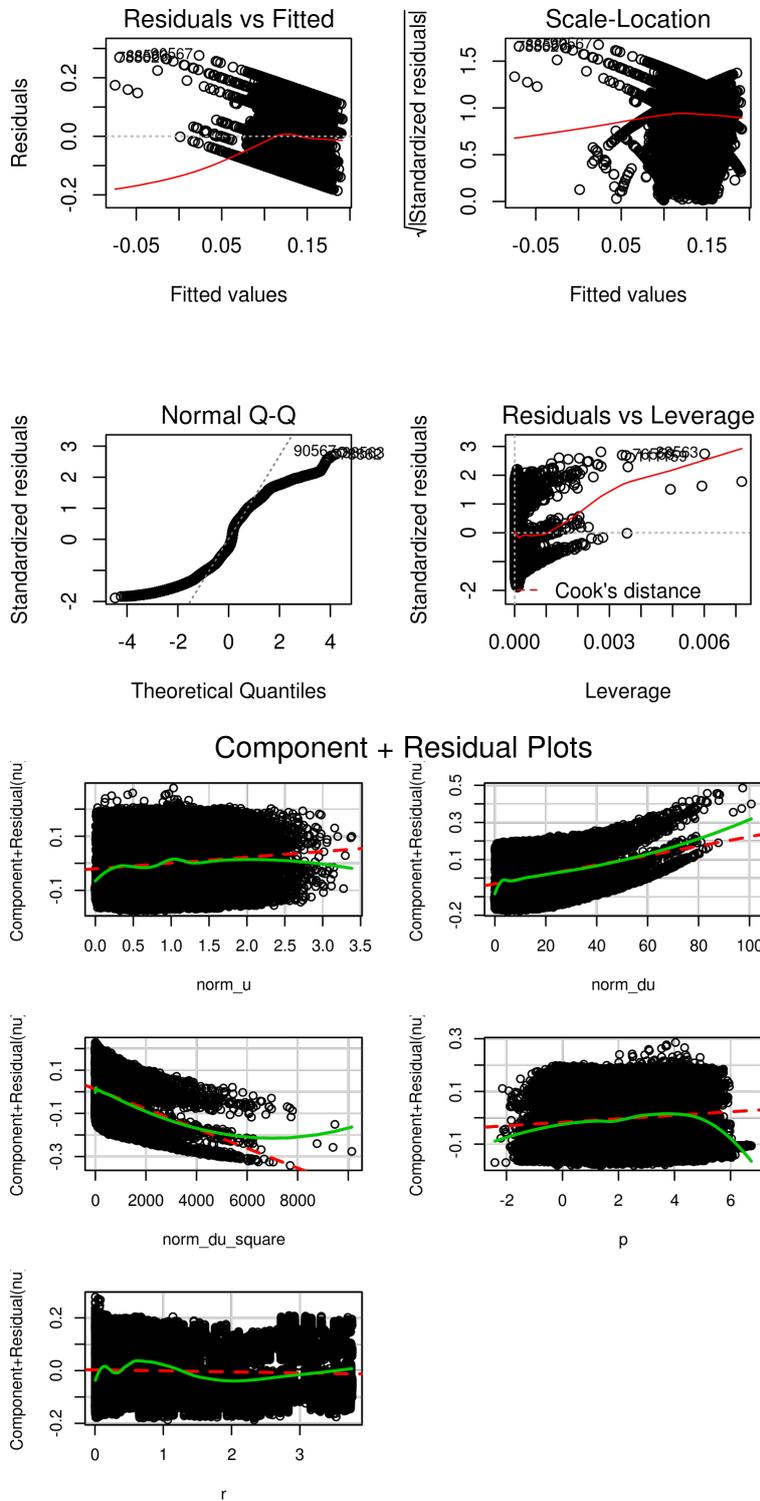


Abbildung 7.17: Approximation 3 der Modelltrainingsdaten A

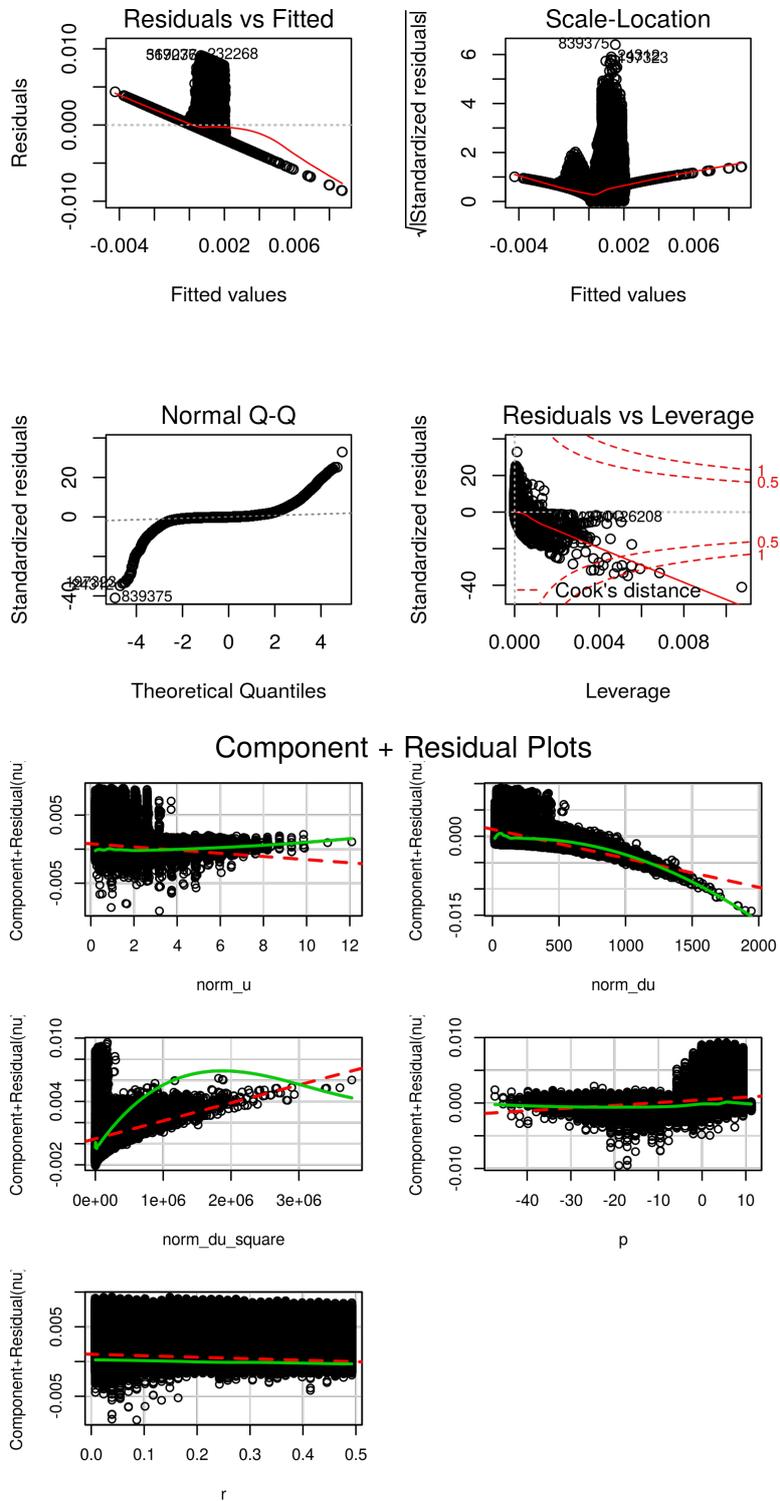


Abbildung 7.18: Approximation 3 der Modelltrainingsdaten B

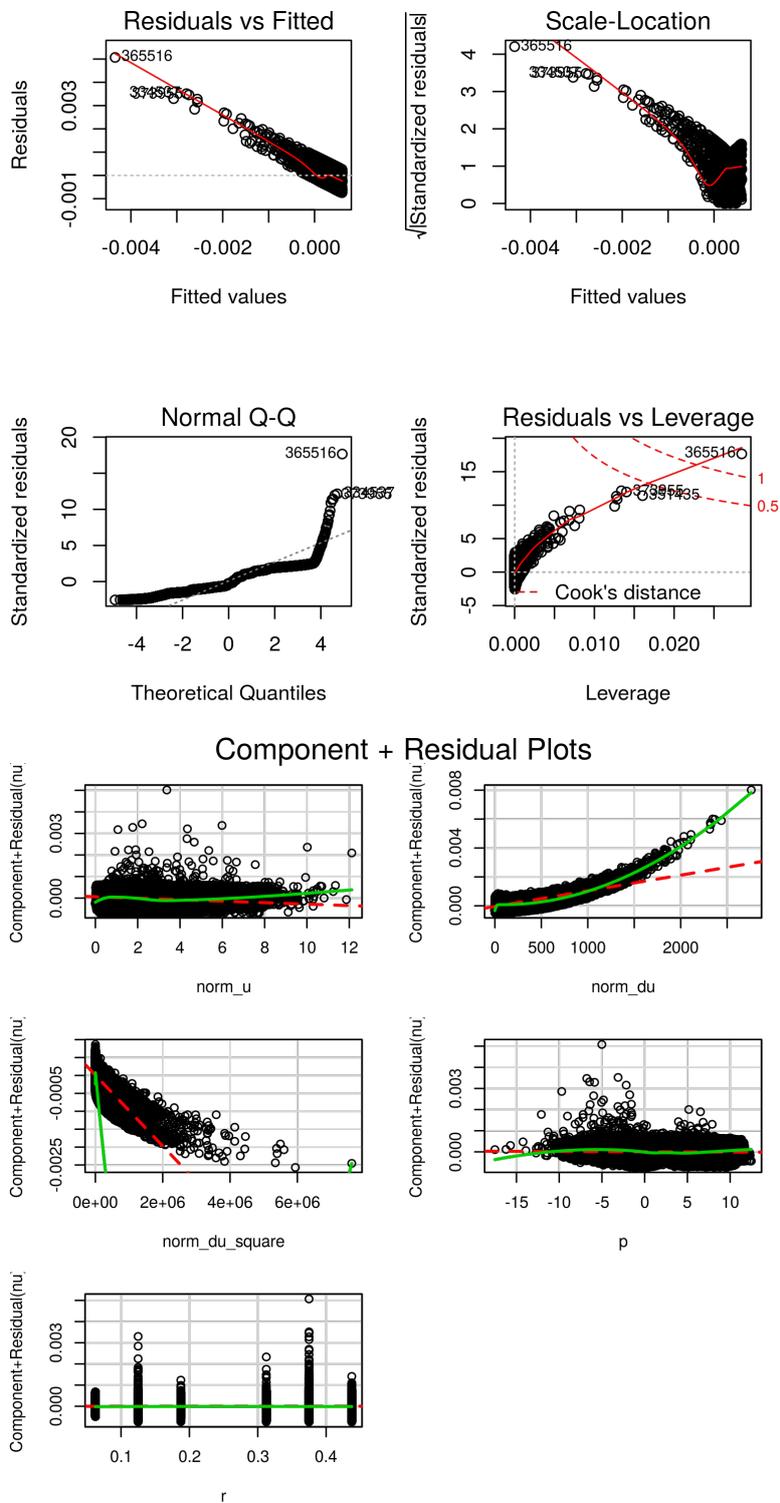


Abbildung 7.19: Approximation 3 der Modelltrainingsdaten C

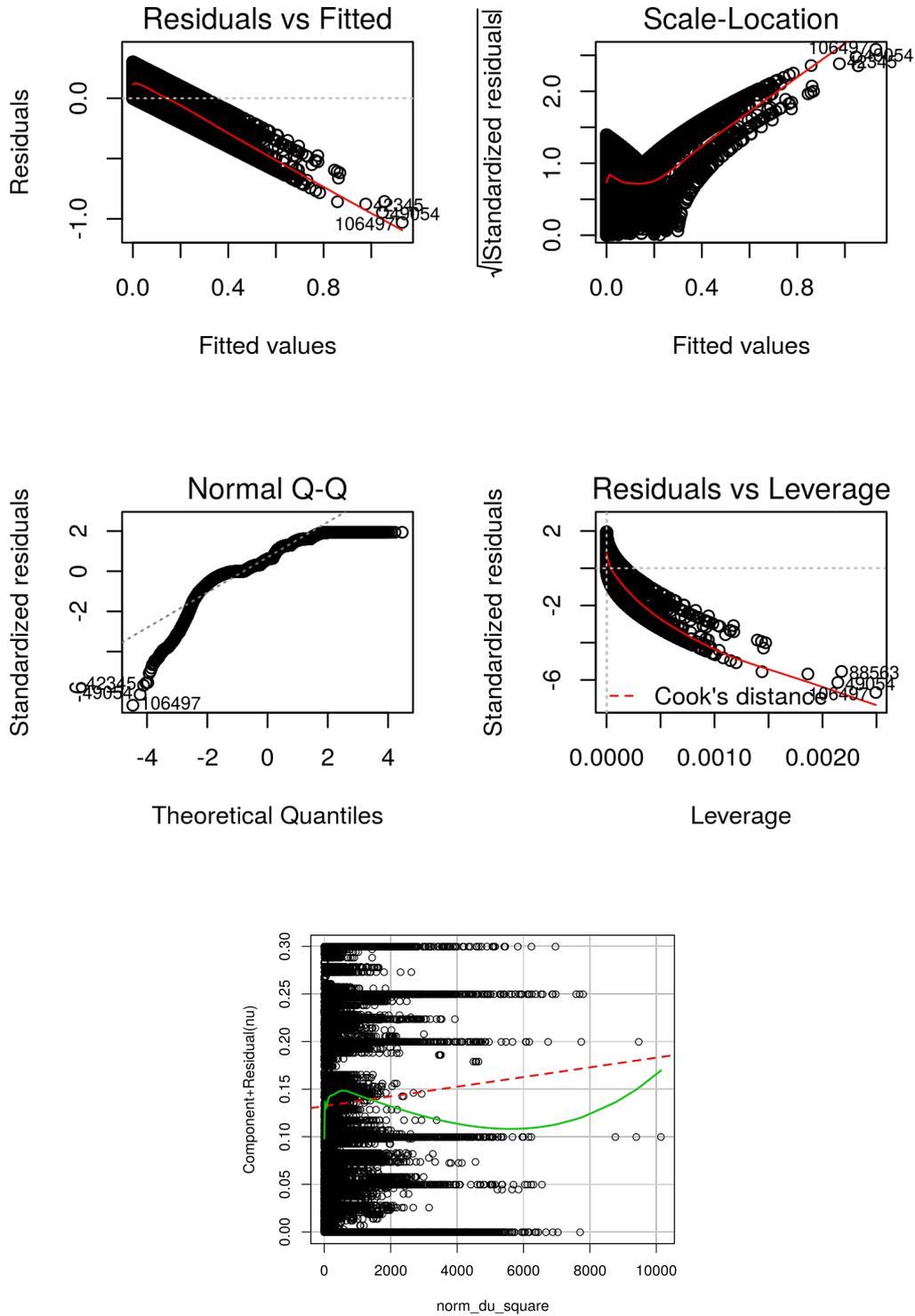


Abbildung 7.20: Approximation 4 (Smagorinsky) der Modelltrainingsdaten A

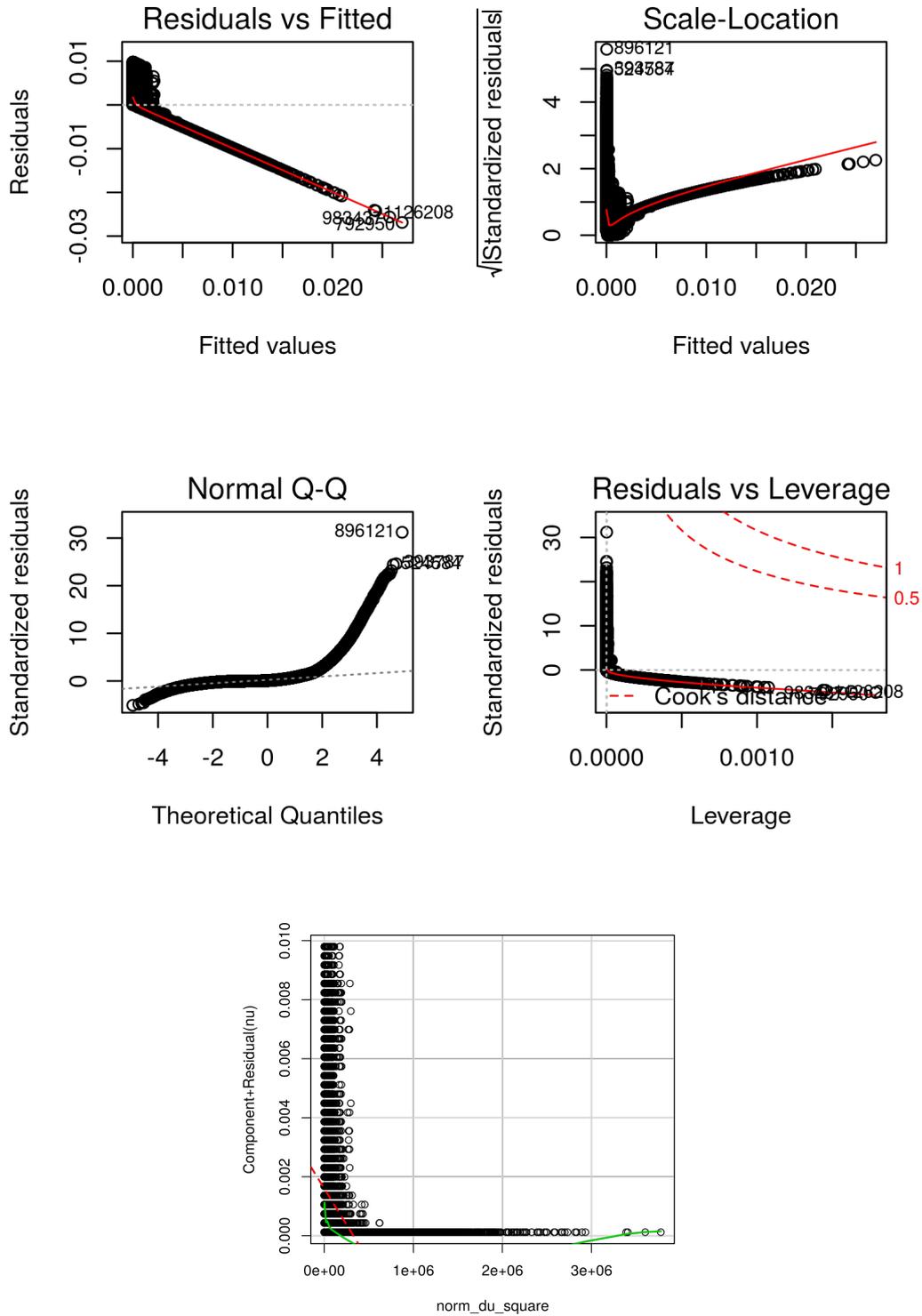


Abbildung 7.21: Approximation 4 (Smagorinsky) der Modelltrainingsdaten B

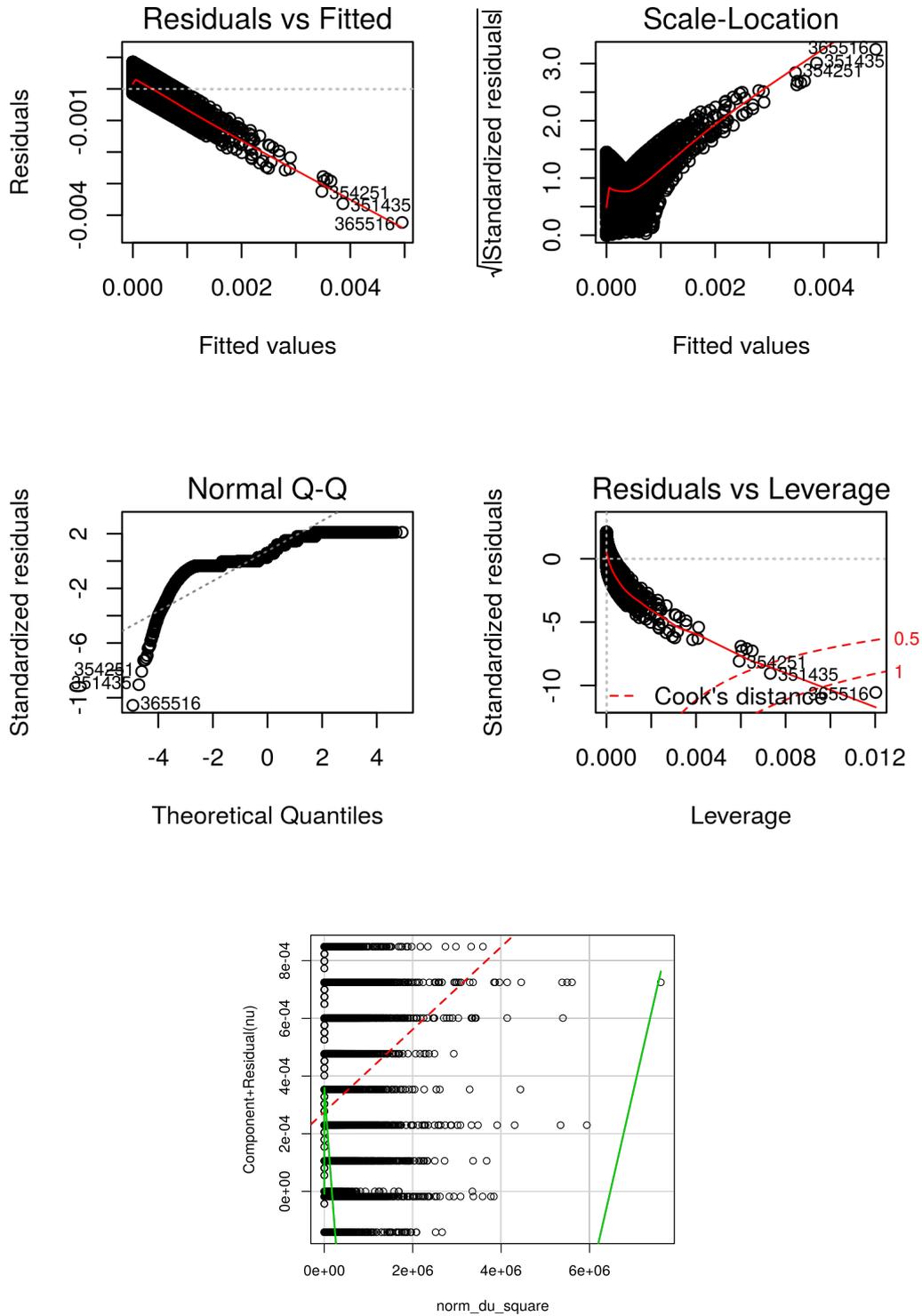


Abbildung 7.22: Approximation 4 (Smagorinsky) der Modelltrainingsdaten C

	Modellproblem A	Modellproblem B
b_0	$9.662\text{e-}02 \pm 9.096\text{e-}04$	$1.429\text{e-}03 \pm 1.488\text{e-}06$
b_1	$2.309\text{e-}02 \pm 5.457\text{e-}04$	$-2.344\text{e-}04 \pm 1.035\text{e-}06$
b_2	$-9.729\text{e-}03 \pm 8.978\text{e-}04$	$-1.568\text{e-}04 \pm 3.279\text{e-}06$
b_3	$1.525\text{e-}04 \pm 3.073\text{e-}05$	$-8.649\text{e-}08 \pm 1.712\text{e-}08$
b_4	$1.318\text{e-}04 \pm 3.804\text{e-}05$	$-4.147\text{e-}08 \pm 3.199\text{e-}08$
b_5	$-5.425\text{e-}05 \pm 1.967\text{e-}05$	$2.035\text{e-}08 \pm 6.517\text{e-}09$
b_6	$-2.079\text{e-}04 \pm 4.040\text{e-}05$	$-5.365\text{e-}07 \pm 1.574\text{e-}08$
b_7	$7.090\text{e-}03 \pm 2.103\text{e-}04$	$4.883\text{e-}05 \pm 9.084\text{e-}08$
b_8	$-1.663\text{e-}03 \pm 3.508\text{e-}04$	$-1.275\text{e-}03 \pm 3.758\text{e-}06$

Tabelle 7.1: Koeffizienten sowie deren Genauigkeit bei Approximation 1 der Modelltrainingsdaten A und B

Die Approximationsergebnisse sind in den Abbildungen 7.20 bis 7.22 dargestellt.

Bemerkung 7.5.2. Bei der Regression der Daten der Modellprobleme gab es die Schwierigkeit, dass die eingesetzte Software „R“ bei der Berechnung und Darstellung der Lowess-Kurve (grün) für Modellproblem B an ihre Grenzen stieß. Um die Abbildungen 7.12, 7.15, 7.18 und 7.21 zu erzeugen, wurde die Daten vorgefiltert: Nahe beieinander liegende Datenpunkte wurden zu einem Repräsentanten zusammengefasst. Außerdem wurde bei der Approximation eine Gewichtung vorgenommen, bei der sich das Gewicht des Repräsentanten durch die Anzahl der durch ihn zusammengefassten Punkte ergibt. Diese Clusterbildung ist zum Teil auch in den Abbildungen 7.12, 7.15, 7.18 und 7.21 zu erkennen. Sie ist jedoch nur für die grafische Darstellung nötig – die eigentliche Berechnung der Regression wurde anhand der exakten Daten durchgeführt.

Die Ergebnisse der Regressionen können den Tabellen 7.2, 7.3, 7.4 und 7.5 entnommen werden. Insbesondere anhand der Varianz der Koeffizienten bei den Approximationen 2 und 3 ist erkennbar, dass der polynomielle Ansatz in $\|\nabla^s u\|$ besser zur Approximation der Daten passt.

Die numerischen Fehler bei den Approximationen sind in Tabelle 7.6 sowie Abbildung 7.23 zu finden. Beim Vergleich dieser Daten mit denen der reduzierten Basis (Kapitel 7.4) ist festzuhalten, dass die räumlich globalen Verfahren bei Modellproblem A bereits mit wenigen Basisfunktionen bessere Approximationen liefern. Jedoch sind die hier vorgestellten Approximationen 1-3 alle genauer als die räumlich lokale Approximation in Kapitel 7.4 – und auch besser als die einfacheren (reduzierte Basis als Abbildung sowie orthonormalisierte Funktionale) räumlich globalen Approximationen bei Modellproblem C. Bei den Modelltrainingsdaten B ist mit keinem der Verfahren ein echter Gewinn zu verzeichnen. Auffällig ist, dass Approximation 4 die schlechteste Approximationsgüte liefert.

Modellproblem C	
b_0	$3.222e-04 \pm 6.553e-07$
b_1	$-4.244e-05 \pm 1.813e-07$
b_2	$4.468e-06 \pm 3.713e-07$
b_3	$-6.101e-07 \pm 3.668e-07$
b_4	$-1.958e-08 \pm 2.186e-08$
b_5	$-1.221e-07 \pm 3.358e-08$
b_6	$-2.741e-08 \pm 3.455e-08$
b_7	$-3.372e-08 \pm 1.280e-08$
b_8	$-3.047e-07 \pm 2.074e-08$
b_9	$9.937e-08 \pm 2.138e-08$
b_{10}	$-5.651e-09 \pm 3.095e-09$
b_{11}	$4.013e-09 \pm 5.327e-09$
b_{12}	$7.486e-09 \pm 6.211e-09$
b_{13}	$-9.515e-07 \pm 7.498e-08$
b_{14}	$1.978e-05 \pm 1.964e-06$

Tabelle 7.2: Koeffizienten sowie deren Genauigkeit bei Approximation 1 der Modelltrainingsdaten C

	Modellproblem A	Modellproblem B	Modellproblem C
b_0	$8.866e-02 \pm 9.609e-04$	$1.459e-03 \pm 1.543e-06$	$3.255e-04 \pm 7.885e-07$
b_1	$2.342e-02 \pm 5.805e-04$	$-2.718e-04 \pm 1.071e-06$	$-3.468e-05 \pm 2.646e-07$
b_2	$5.720e-04 \pm 2.608e-05$	$2.701e-07 \pm 8.802e-09$	$3.851e-07 \pm 3.456e-09$
b_3	$6.876e-03 \pm 2.073e-04$	$4.767e-05 \pm 9.127e-08$	$-1.527e-06 \pm 7.617e-08$
b_4	$-2.315e-03 \pm 3.495e-04$	$-1.272e-03 \pm 3.758e-06$	$2.170e-05 \pm 2.001e-06$

Tabelle 7.3: Koeffizienten sowie deren Genauigkeit bei Approximation 2

	Modellproblem A	Modellproblem B	Modellproblem C
b_0	$8.109e-02 \pm 9.813e-04$	$1.427e-03 \pm 1.588e-06$	$3.096e-04 \pm 7.959e-07$
b_1	$2.122e-02 \pm 5.813e-04$	$-2.660e-04 \pm 1.073e-06$	$-3.421e-05 \pm 2.631e-07$
b_2	$2.532e-03 \pm 6.245e-05$	$1.222e-06 \pm 1.441e-08$	$1.081e-06 \pm 6.882e-09$
b_3	$-4.561e-05 \pm 1.322e-06$	$-2.648e-09 \pm 3.173e-11$	$-9.889e-10 \pm 8.472e-12$
b_4	$6.642e-03 \pm 2.064e-04$	$4.635e-05 \pm 9.259e-08$	$-1.349e-06 \pm 7.576e-08$
b_5	$-3.963e-03 \pm 3.512e-04$	$-1.265e-03 \pm 3.758e-06$	$6.680e-06 \pm 1.994e-06$

Tabelle 7.4: Koeffizienten sowie deren Genauigkeit bei Approximation 3

	Modellproblem A	Modellproblem B	Modellproblem C
b_1	$1.114e-04 \pm 7.611e-07$	$4.110e-09 \pm 2.294e-11$	$6.505e-10 \pm 5.805e-12$

Tabelle 7.5: Koeffizienten sowie deren Genauigkeit bei Approximation 4 (Smagorinsky)

Approximation	keine	1	2	3	4 (Smagorinsky)
Modellproblem A	1.6702e-01	9.898e-02	9.884e-02	9.838e-02	1.545e-01
Modellproblem B	1.9966e-03	1.636e-03	1.634e-03	1.634e-03	1.993e-03
Modellproblem C	4.0525e-04	2.897e-04	2.928e-04	2.912e-04	4.032e-04

Tabelle 7.6: Vergleich der Approximationsfehler bei den verschiedenen Approximationen

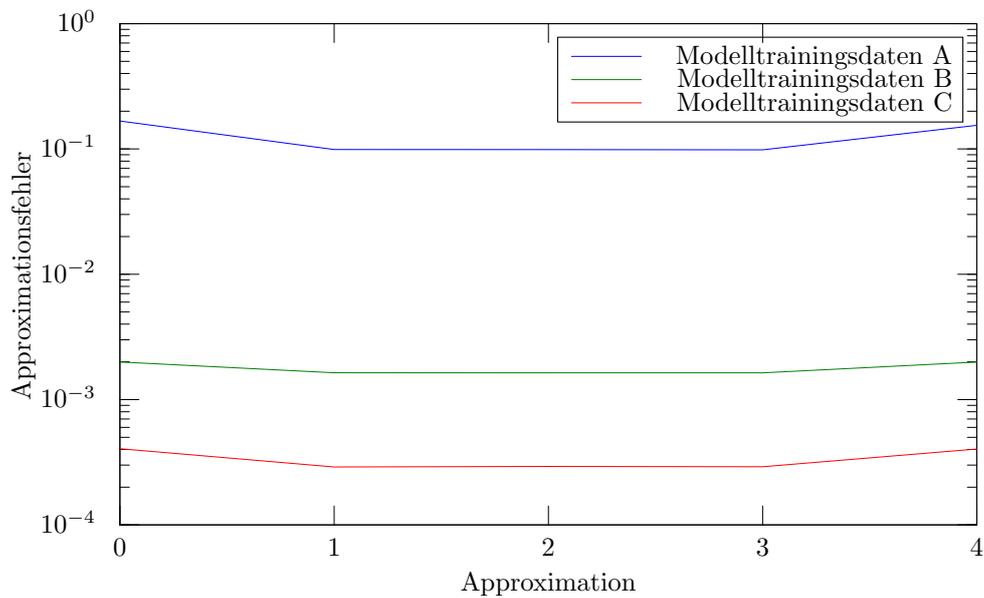


Abbildung 7.23: Vergleich der Approximationsfehler bei lokaler Regression der Modelltrainingsdaten A, B und C nach Approximationen 1, 2, 3 und Smagorinsky-Modell (4) sowie keiner Modellierung (0)

7.5.1 Verwendung der Abbildungen als Viskositätsmodell

Die so bestimmten Modelle wurden nun statt des inversen Problems zur Wahl der künstlichen Viskosität verwendet und die Werte des Funktionals \hat{j} ermittelt. Dabei wurden lediglich die Approximationen 2-4 untersucht, da die Approximation 1 speziellere Anpassung benötigt hätte (Dimensionsabhängigkeit). Da ihre Approximationsgüte sich nicht signifikant von denen der Approximationen 2 und 3 unterscheidet, war hier kein positiver Mehrwert zu erwarten. Die Ergebnisse sind in Abbildungen 7.24, 7.25 bzw. 7.26 zu finden.

Obwohl mit den vorgestellten Approximationen durchaus gute Ergebnisse bezüglich der Approximationsgüte erzielt wurden, ergeben sich nicht ganz so schöne Ergebnisse in der praktischen Anwendung als künstliche Viskosität.

Bei Modellproblem A stellt sich für die Approximationen 2 und 3 nach einer kurzen Aufbauphase der Strömung ein sehr stationäres, laminares Verhalten ein. Dies ist zwar eine Stabilisierung gegenüber der grobskaligen Rechnung ohne Modellierung, jedoch geht sie zu weit und entspricht nicht der Referenzrechnung. Insbesondere ist auch der zeitliche Mittelwert des Funktionals wesentlich zu niedrig. Bei Verwendung der Approximation 4 ist der zeitliche Mittelwert der Rechnung wesentlich besser, allerdings oszilliert die Rechnung hier strukturiert.

Bei Modellproblem B erkennen wir ab $t \approx 2$ deutliche Abweichungen der einzelnen Rechnungen. Nach einem guten Anfang „übersteuert“ das Viskositätsmodell bei den Approximationen 2 und 3 und entfernt sich zunehmend von der gewünschten Lösung. Lediglich das Funktional der Lösung mit Smagorinsky-Modell folgt jener von u_V^H aus dem inversen Problem gut.

Interessant ist, dass bei allen drei Approximationen (sowie den unstabilierten Rechnungen) nach dem Aufbau der laminaren Strömung ein Einbruch der Strömung stattfindet, bevor sich die turbulente Strömung einstellt. Dieser Einbruch ist bei den Approximationen 2 und 3 sehr viel später ($t \approx 4.5$) als bei den unstabilierten Rechnungen sowie der mit Smagorinsky-Modell ($t \approx 2.5$). Allerdings ist die Geschwindigkeit, mit der dieser Strömungsabfall eintritt, eher bei den unstabilierten Rechnungen und den Approximationen 2 und 3 übereinstimmend. Beim Smagorinsky-Modell hingegen erfolgt dieser Übergang deutlich gedämpfter.

Bei Modellproblem C liefern alle drei Approximationen ähnliche Ergebnisse: Bei allen grobskaligen Rechnungen (mit sowie ohne Modellierung) erfolgt der laminare Strömungsaufbau etwas langsamer und die laminare Phase dauert länger als bei der feinskaligen Rechnung. Der anschließende Einbruch des Funktionalwertes wird vom Smagorinsky-Modell am besten nachgebildet, die Approximationen 2 und 3 laufen hier etwas langsamer nach. Ab etwa $t = 6$ schwanken die Funktionale aller Rechnungen in ähnlicher Weise.

Die Strömungsprofile der Rechnung der Modellproblems B können in der Abbildung 7.27 genauer betrachtet werden. Dort wurde über ausgewählte Zeiträume jeweils die mittlere Strömung in Hauptströmungsrichtung über dem Rohrquerschnitt aufgetragen. Daran sind die einzelnen Phasen der Strömung (laminar, Übergang, turbulent) gut zu erkennen. Bemerkenswert ist, dass in der Übergangsphase der feinskaligen unstabilierten Rechnung in den Randbereichen sogar eine Rückströmung auftritt (negative Werte in Abbildung 7.27(a)). Dies wird von der Viskosität aus dem inversen Problem nachgebildet, jedoch ist ein solches Verhalten in keinem der Approximationsmodelle wiederzufinden. Natürlich ist zu beachten, dass dieses Merkmal überhaupt nicht durch das Funktional beschrieben wurde. Es ist also auch nicht weiter verwunderlich, dass hier keine Übereinstimmung erzielt wird. Die Abbildungen der Strömungsprofile

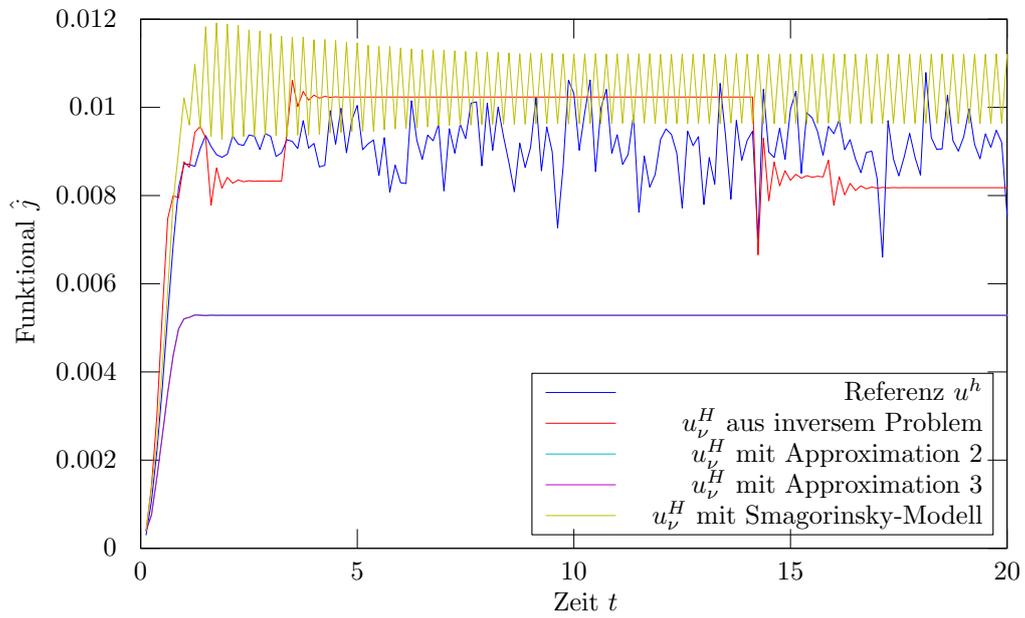


Abbildung 7.24: Entwicklung des Funktional über die Zeit bei Modellproblem A und lokaler Regression nach Approximationen 2, 3 sowie Smagorinsky-Modell

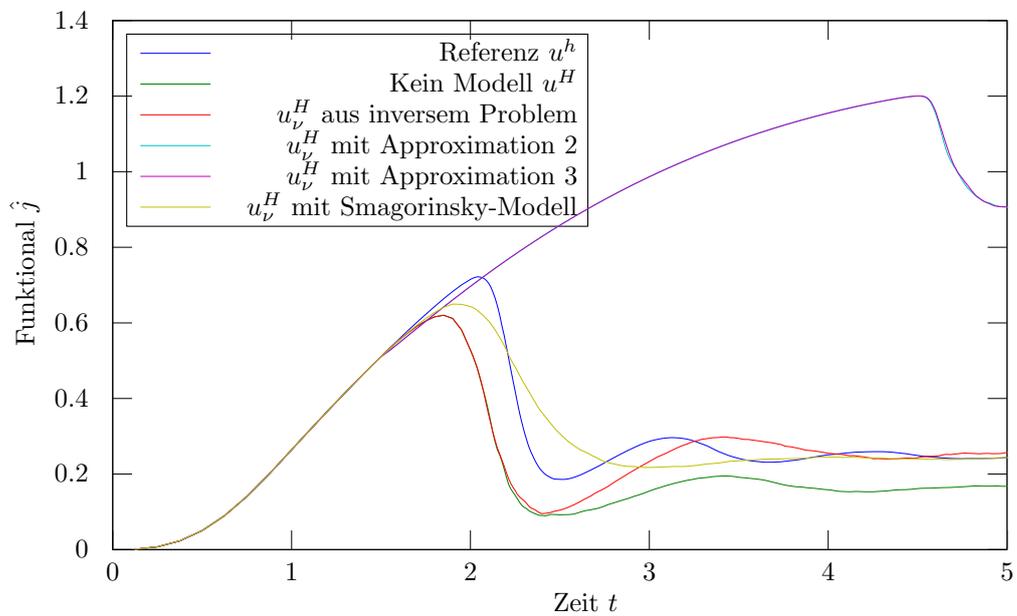


Abbildung 7.25: Entwicklung des Funktional über die Zeit bei Modellproblem B und lokaler Regression nach Approximationen 2, 3 sowie Smagorinsky-Modell

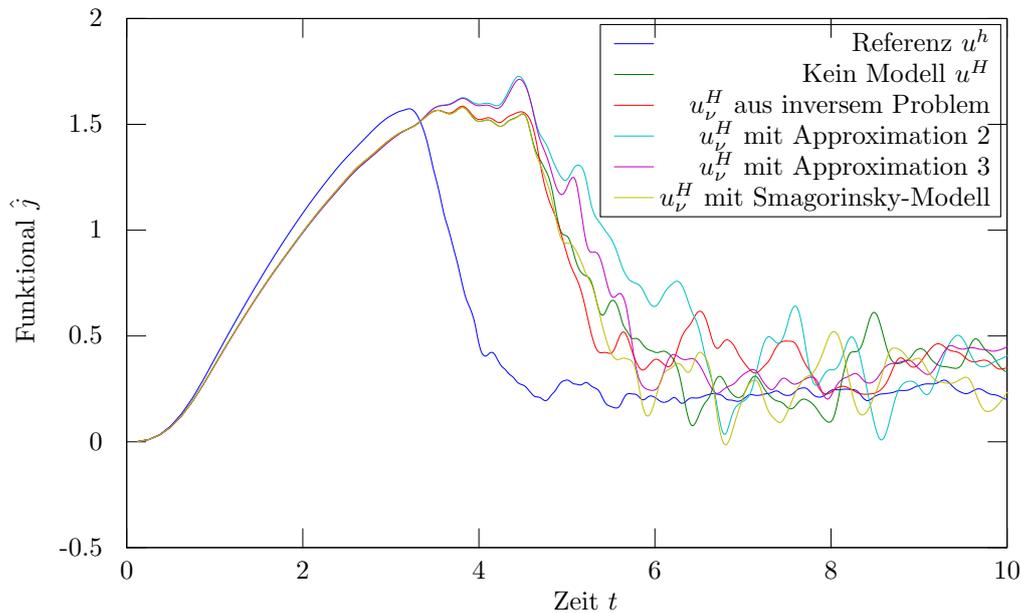
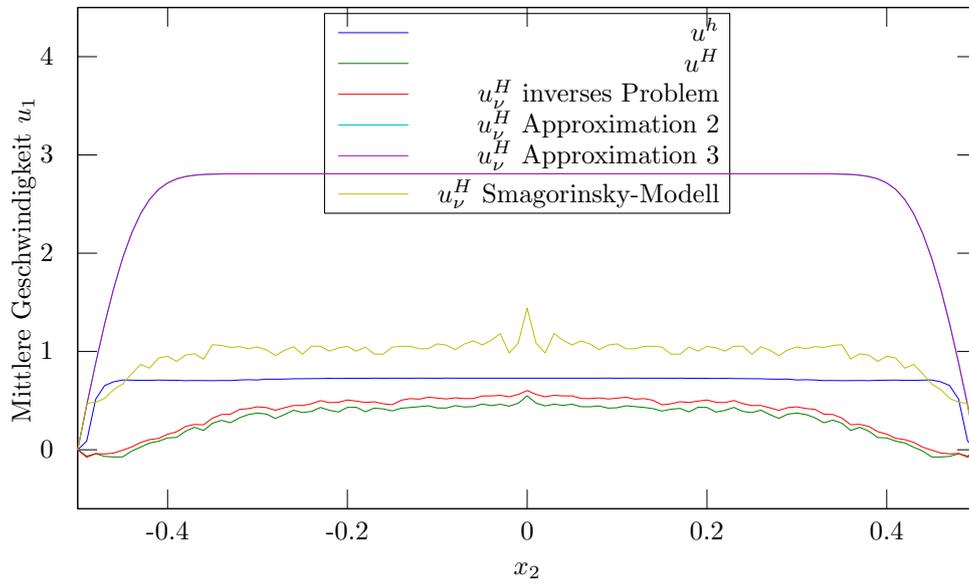


Abbildung 7.26: Entwicklung des Funktionals über die Zeit bei Modellproblem C und lokaler Regression nach Approximationen 2, 3 sowie Smagorinsky-Modell

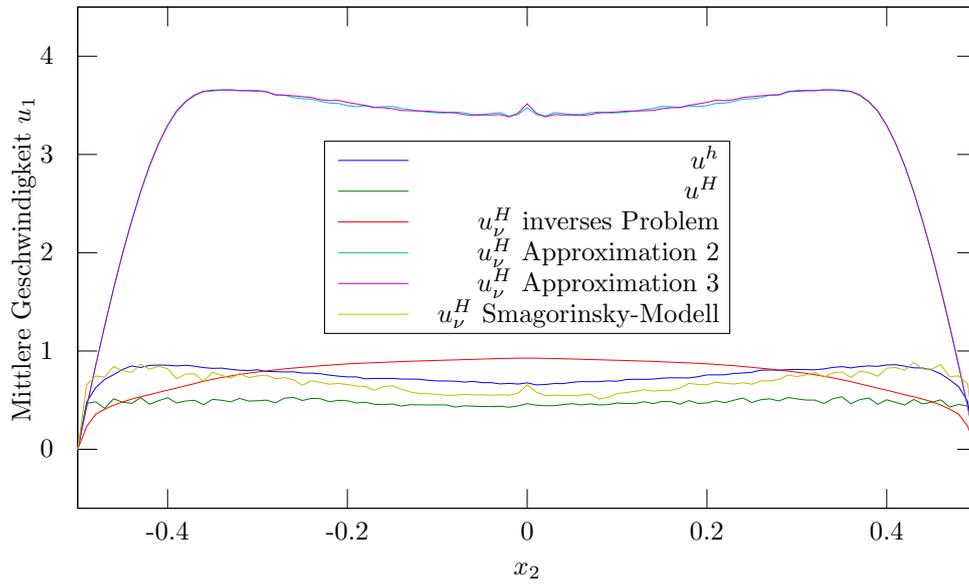
sollen vor allem einen besseren Überblick über die berechneten Strömungsergebnisse geben.

In den Strömungsquerschnitten von Modellproblem C (Abbildung 7.28) ist die Struktur der beiden den verschiedenen Rechnungen zugrunde liegenden Gittern deutlich sichtbar. Im ersten Zeitabschnitt ist erkennbar, dass das Smagorinsky-Modell der Vorgabe durch die feinskalige Rechnung bzw. den Daten aus dem inversem Problem am besten folgt. Die Approximationen 2 und 3 laufen an dieser Stelle deutlich hinterher. Im zweiten Zeitabschnitt hingegen sind kaum noch wesentliche Unterscheidungen der grobskaligen Rechnungen zu erkennen. Die verschiedenen Modelle scheinen hier keinen großen Einfluss zu haben.

Im Folgenden wird noch ein weiterer Approximationsansatz vorgestellt. Bei diesem wird die Wahl der Ansatzfunktion vereinfacht und das Verhalten für untrainierte Eingangsdaten verbessert.

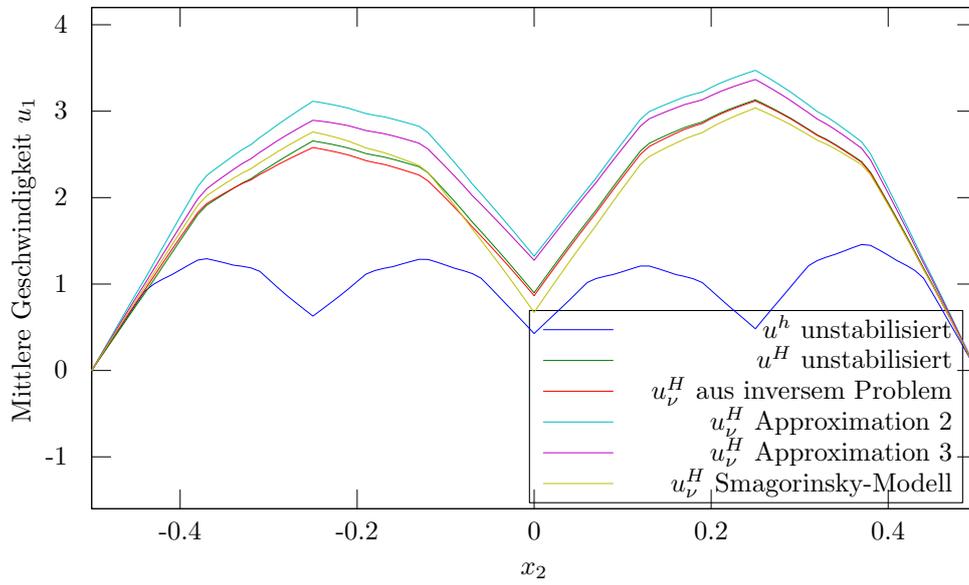


(a)

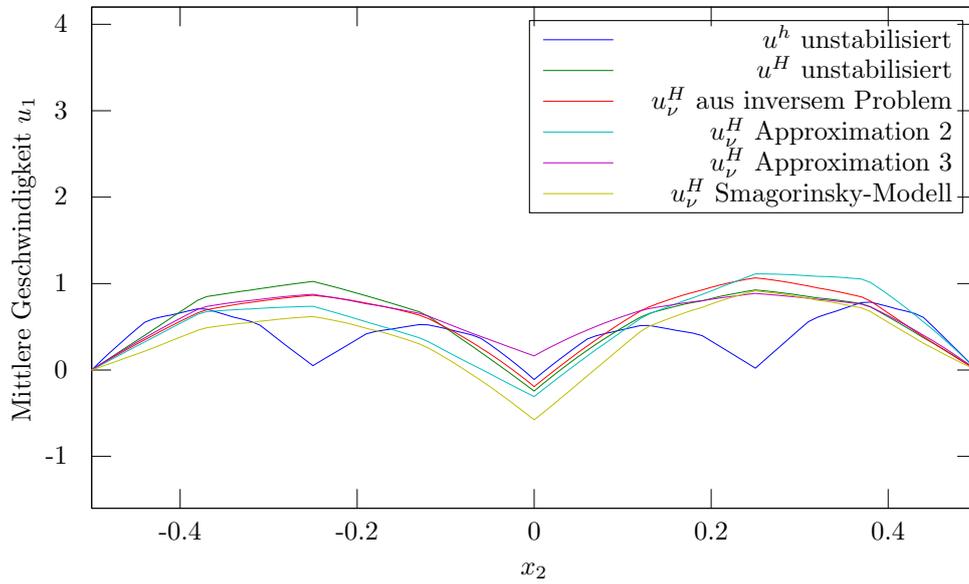


(b)

Abbildung 7.27: Geschwindigkeitsquerschnitte bei Modellproblem B mit den lokalen Approximationen über den Zeitraum (a) $t \in (2.25, 2.75)$ sowie (b) $t \in (2.75, 5)$



(a)



(b)

Abbildung 7.28: Geschwindigkeitsquerschnitte bei Modellproblem C mit den lokalen Approximationen über den Zeitraum (a) $t \in (3, 6)$ sowie (b) $t \in (8, 10)$

7.6 Modellierung mittels radialer Basisfunktionen

Eine radiale Basisfunktion ist eine Funktion $\varphi : \mathbb{R}^k \rightarrow \mathbb{R}$, deren Funktionswert nur vom Abstand zum Ursprung abhängt, also

$$\varphi(x) = \phi(\|x\|), \quad \phi : \mathbb{R}_0^+ \rightarrow \mathbb{R}.$$

Wir erhalten Variationen dieser Basisfunktion, wenn wir den Bezugspunkt zur Abstandsmessung ändern:

$$\tilde{\varphi}(x) = \phi(\|x - c\|)$$

$\tilde{\varphi}$ nimmt also für alle Punkte mit Abstand r von $c \in \mathbb{R}^k$ den Wert $\phi(r)$ an, $r > 0$. Es handelt sich also bei $\tilde{\varphi}$ gegenüber φ um eine um c zentrierte Funktion.

Es sind verschiedene Zentren c denkbar, und wir können verschiedene derart abgewandelte Funktionen $\tilde{\varphi}$ als Linearkombination zusammenfassen:

$$f(x) = \sum_{i=1}^M a_i \phi(\|x - c_i\|).$$

$M \in \mathbb{N}$ ist die Anzahl der verwendeten Basisfunktionen und $c_i \in \mathbb{R}^k$ sind die jeweiligen Zentren. Funktionen $f : \mathbb{R}^k \rightarrow \mathbb{R}$ dieser Bauart werden gern zur Approximation von multivariaten Funktionen mit weit gestreuten Daten verwendet.

Grundsätzlich sind an dieser Stelle verschiedene Wahlen von Funktionen ϕ denkbar. Für Approximationen relativ erprobt sind etwa die Gaußschen radialen Basisfunktionen. Dazu ist

$$\phi(r) = \exp(-\lambda r^2).$$

Die Bezeichnung ist als Anlehnung an die aus der Stochastik bekannte Gaußsche „Glockenkurve“, der Dichte der Normalverteilung, zu verstehen. Der zusätzliche „Streckungsparameter“ λ kann variiert werden.

Insgesamt wollen wir im Folgenden also Gaußsche radiale Basisfunktionen betrachten:

$$\tilde{m}_i(d) = \sum_{i=1}^M a_i \exp(-\lambda_i \|d - c_i\|^2).$$

Dabei sind $\lambda_i > 0$ die Streckungsparameter, $c_i \in \mathbb{R}^k$ die Zentren in \mathbb{R}^k und $a_i \in \mathbb{R}$ die Gewichtungen der Linearkombination.

Für tiefer gehende Betrachtungen zu radialen Basisfunktionen sei an dieser Stelle auf [Buh03] verwiesen.

Als einfache Strategie zur nichtlinearen Regression wird wie folgt vorgegangen:

- λ_i und c_i werden „greedy“, d.h. mit einer gierigen Strategie gewählt. Dazu werden als Zentrum c_i jene Trainingsdaten ausgewählt, die die derzeit größte Abweichung aufweisen. Für die Wahl der λ_i werden jene Trainingsdaten in der Nähe des neuen Zentrums ausgewählt, deren Fehler im Vorzeichen mit dem im Zentrum übereinstimmt. Aus ihren Abständen zu c_i und der jeweils gewünschten Korrektur kann dann der am besten

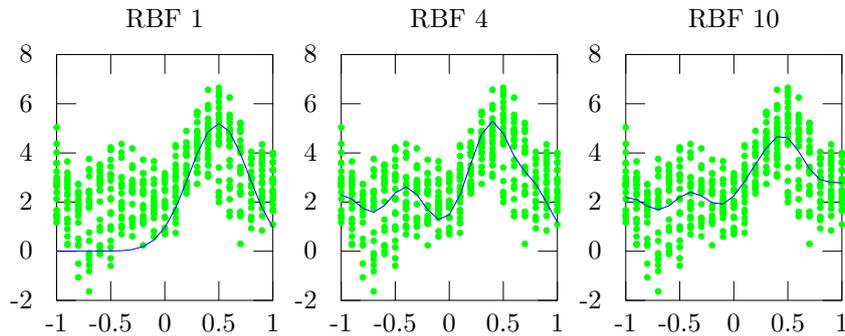


Abbildung 7.29: Eindimensionale Beispieldaten (mit Rauschen) zur Approximation mit radialen Basisfunktionen mit $M \in \{1, 4, 10\}$

passende Streckungsparameter λ_i bestimmt werden.

- Die Koeffizienten a_i werden anschließend mittels linearer Kleinste-Quadrate-Regression berechnet.

Der konkrete Ablauf zur Bestimmung einer Approximation durch radiale Basisfunktionen (RBF) ist in Algorithmus 7.5 erklärt.

Beispiel 7.6.1. Es werden eindimensionale Beispieldaten (mit Rauschen) mit $M \in \{1, 4, 10\}$ radialen Basisfunktionen nach Algorithmus 7.5 approximiert. Betrachte Abbildung 7.29.

Die Approximationen, die sich aus Algorithmus 7.5 für die Modelltrainingsdaten A ergeben, sind in Tabelle 7.7 bzw. Abbildung 7.30 ersichtlich. Entsprechende Daten sind für die Modelltrainingsdaten B in Tabelle 7.8 bzw. Abbildung 7.31, für die Modelltrainingsdaten C in Tabelle 7.9 bzw. Abbildung 7.32 zu finden.

Die Güte der Approximation mithilfe der radialen Basisfunktionen ist in der Abbildung 7.33 sowie der Tabelle 7.10 dargestellt. Es ist zu erkennen, dass der wesentliche Beitrag zur Verbesserung der Approximationsgüte schon für kleine M auftritt: Bei Modellproblem A verbessert sich die Approximation nach den ersten fünf Basisfunktionen nur geringfügig. Bei Modellproblem C scheint nach ungefähr acht Basisfunktionen die Approximationsgüte ausgereizt, die Verbesserungen danach fallen wesentlich geringer aus. Lediglich bei Modellproblem B ist kein derartig rapider Abfall im Approximationsfehler bei den ersten Basisfunktionen zu erkennen. Im Gegenteil scheint sich die Approximationsgüte bei Modellproblem B wesentlich gleichmäßiger zu verbessern. Allerdings ist die Geschwindigkeit dieser Verbesserung mit $3,6 \cdot 10^{-4}$ relativer Fehlerreduktion pro Basisfunktion recht gering.

Bemerkung 7.6.1. Der Einfluss der Basisfunktion auf die Verbesserung der Approximation ist auch in ihren Koeffizienten ablesbar: Für Modellproblem A ist zu erkennen, dass die Koeffizienten ab der vierten Basisfunktion bei RBF-50 ein deutlich höheres Niveau haben als davor bzw. bei RBF-10 oder RBF-2. Dies ist bei Modellproblem C noch deutlicher (Abbildung 7.32(a)): Ab der siebten Basisfunktion werden hier fast nur noch Koeffizienten angenommen, die größer als 1 sind. Die Koeffizienten der Basisfunktionen davor und auch die von RBF-10 sowie RBF-2 waren allesamt kleiner als 1. Die kleineren Werte sind deutlich besser nachzuvollziehen, schließlich variieren die zu approximierenden Werte der künstlichen Viskosität nur zwischen $-9 \cdot 10^{-5}$ und $9 \cdot 10^{-4}$ (vgl. etwa Tabelle 6.1).

Algorithm 7.5 Approximation mittels radialer Basisfunktionen

Input: $P_{\text{train}} \subseteq D \times R, \varepsilon_{\text{tol}} > 0, M_{\text{max}} \in \mathbb{N}, w : D \times R \rightarrow \mathbb{R}^+$ **Output:** $M \in \mathbb{N}_0, \{a_i \mid i = 1, \dots, M\} \subseteq \mathbb{R}, \{\lambda_i \mid i = 1, \dots, M\} \subseteq \mathbb{R}^+, \{c_i \mid i = 1, \dots, M\} \subseteq \mathbb{R}^k$ $M \leftarrow 0$ **repeat****for** $(d, r) \in P_{\text{train}}$ **do**

Berechne die bisherige Approximation

$$\tilde{r}_d \leftarrow \sum_{k=1}^M a_k \exp(-\lambda_k |d - c_k|^2)$$

end for

Finde die Approximation mit dem größten Fehler

$$(d_{M+1}, r_{M+1}) \leftarrow \operatorname{argsup}_{(d,r) \in P_{\text{train}}} w(d, r) \|r - \tilde{r}_d\|$$

Setze neuen Mittelpunkt $c_{M+1} \leftarrow d_{M+1}$.Setze Gewichtung der neuen Basisfunktion $a_{M+1} \leftarrow r_{M+1} - \tilde{r}_{d_{M+1}}$.

Selektiere die relevanten Daten zur Bestimmung des neuen Radius

$$P_{M+1} \leftarrow \left\{ (d, r) \in P_{\text{train}} \mid \frac{r - \tilde{r}_d}{a_{M+1}} > 0 \wedge d_i \neq c_k \forall k = 1, \dots, M \right\}$$

Berechne den neuen Radius

$$\lambda_{M+1} \leftarrow \frac{-\sum_{(d,r) \in P_{M+1}} w(d, r) \frac{\ln \frac{r - \tilde{r}_d}{a_{M+1}}}{\|d - c_{M+1}\|^2}}{\sum_{(d,r) \in P_{M+1}} w(d, r)}$$

Definiere die Matrix zur linearen Regression

$$h_{ij} \leftarrow \exp(-\lambda_j \|d - c_j\|^2) \quad \forall (d, r) \in P_{\text{train}}, j \in 1, \dots, M + 1$$

Löse

$$H^\top W H a = H^\top W y$$

nach a mit $H = (h_{ij})$ und $W = \operatorname{diag}(w(d, r))$.Erhöhe die Anzahl der Approximationsterme M

$$M \leftarrow M + 1$$

until $\|r_M\| < \varepsilon_{\text{tol}}$ **or** $M \geq M_{\text{max}}$

a_i (RBF-2)	a_i (RBF-10)	a_i (RBF-50)	λ_i	$(\ u\)_i$	$(\ \nabla^s u\)_i$	$(\text{dist}(\Gamma_{\text{obs}}, x))_i$
2.5228e-01	1.4939e-03	1.2083e-02	1.1068e-01	1.7499e+00	1.7439e+01	5.7540e-02
2.2556e-01	1.1084e-02	2.3373e+00	1.0071e-02	1.3071e+00	4.4010e+01	3.4691e-01
	-1.5275e-02	4.6006e-01	1.6083e-01	1.0206e+00	8.9482e-02	6.6572e-01
	-1.7193e+00	-2.0926e+04	3.6743e-04	1.0312e+00	8.3440e+01	6.4184e-03
	5.1347e-01	6.2346e+03	6.4865e-04	1.0427e+00	8.3866e+01	6.4184e-03
	-3.2066e-02	-3.0973e+01	8.4068e-03	9.3229e-01	7.8990e+01	6.4184e-03
	3.1750e+00	-4.6362e+02	2.3309e-04	4.9151e-01	9.7323e+01	4.9650e-02
	-1.8981e+00	-2.3326e+04	2.8789e-04	8.3777e-01	1.0070e+02	1.2247e-02
	7.8585e-02	4.8250e+03	9.2781e-04	1.2302e+00	7.3815e+01	7.4030e-03
	-4.8573e-02	3.6215e+00	1.3881e-02	9.8237e-01	6.6328e+01	6.4184e-03
		1.3875e+00	3.4975e-02	1.0637e+00	6.7640e+01	7.4030e-03
		1.2985e+01	1.5223e-02	1.0652e+00	6.7775e+01	7.4030e-03
		-1.3974e+00	5.7674e-02	1.1146e+00	7.1331e+01	7.4030e-03
		-3.6891e+02	2.9400e-03	6.4347e-01	7.1168e+01	6.4898e-02
		2.8807e+00	4.5970e-02	1.0512e+00	7.1462e+01	6.4184e-03
		1.1225e+03	2.2896e-03	9.4757e-01	6.9970e+01	6.4184e-03
		1.2025e+03	1.6818e-03	1.1000e+00	7.1732e+01	6.4184e-03
		-4.9700e+01	1.2837e-02	1.1286e+00	6.8801e+01	7.4030e-03
		3.2101e+01	9.9137e-03	1.0488e+00	6.8228e+01	7.4030e-03
		-4.1316e+04	4.5599e-04	1.0045e+00	8.7772e+01	6.4184e-03
		-2.4344e+02	8.2645e-04	1.5636e+00	6.9190e+01	3.2999e+00
		5.0991e+01	1.5821e-03	2.0049e+00	5.5060e+01	3.3851e+00
		-3.4239e+00	8.1841e-03	8.7272e-01	4.8184e+01	3.3132e+00
		1.0347e+00	5.2521e-02	9.6696e-01	6.9106e-01	3.3132e+00
		1.2438e+00	9.5609e-03	1.3706e+00	3.5146e+01	3.5151e+00
		-3.2098e+00	1.4279e-02	2.6025e-01	3.0571e+01	3.3132e+00
		1.9109e+00	1.8257e-02	5.0450e-01	3.0487e+01	3.3132e+00
		-7.1692e-02	2.8932e-02	7.4472e-01	3.0950e+01	3.3132e+00
		-4.7369e+00	6.3235e-03	1.2914e+00	3.7938e+01	3.5730e+00
		-9.0125e+02	1.1568e-03	1.3060e-01	7.3619e+01	6.4898e-02
		1.3001e+02	1.3743e-03	2.0518e+00	5.9677e+01	3.2154e+00
		-7.0483e+02	1.8568e-03	1.8246e+00	6.7049e+01	5.6183e-02
		-9.8532e+02	1.7478e-03	2.4143e-01	6.8927e+01	5.0566e-02
		2.0299e+04	3.7298e-04	1.4675e+00	9.6936e+01	1.5802e-02
		8.4980e+03	4.5686e-04	3.0038e-02	8.7976e+01	2.4375e-02
		-3.3591e+03	1.0389e-03	1.7130e+00	7.6927e+01	5.5755e-02
		4.9973e+04	3.5174e-04	9.6890e-01	8.8277e+01	6.4184e-03
		-3.2769e+01	2.8317e-04	5.3344e-01	9.3632e+01	1.2247e-02
		-1.6857e+00	1.7577e-03	4.5970e-01	5.6652e+01	1.1416e+00
		-4.2680e+01	7.1947e-03	1.1076e+00	6.4553e+01	1.6174e-01
		-1.2841e+00	1.5497e-01	7.9964e-01	4.9341e-01	3.7230e+00
		3.6448e+00	1.2798e-01	8.2442e-01	5.7839e-01	3.7230e+00
		1.9302e+00	7.8211e-02	3.4346e-01	1.2305e+00	3.0825e+00
		-2.8486e+00	8.9874e-02	6.2598e-01	1.5238e+00	3.3132e+00
		-1.2950e-01	1.7434e-01	8.1185e-01	6.6075e-01	3.7230e+00
		4.9613e-01	3.4102e-01	1.3750e-01	8.0190e-01	4.8480e-02
		-5.9688e-01	3.5174e-01	1.8885e-01	9.4119e-01	4.8480e-02
		-1.4888e+00	1.2310e-01	6.7698e-01	7.9133e-01	3.3589e+00
		-9.2532e-01	1.6396e-01	9.0857e-01	5.8227e-01	2.7273e+00
		7.5706e-01	1.0533e-01	1.1078e+00	1.2203e+00	2.8798e+00

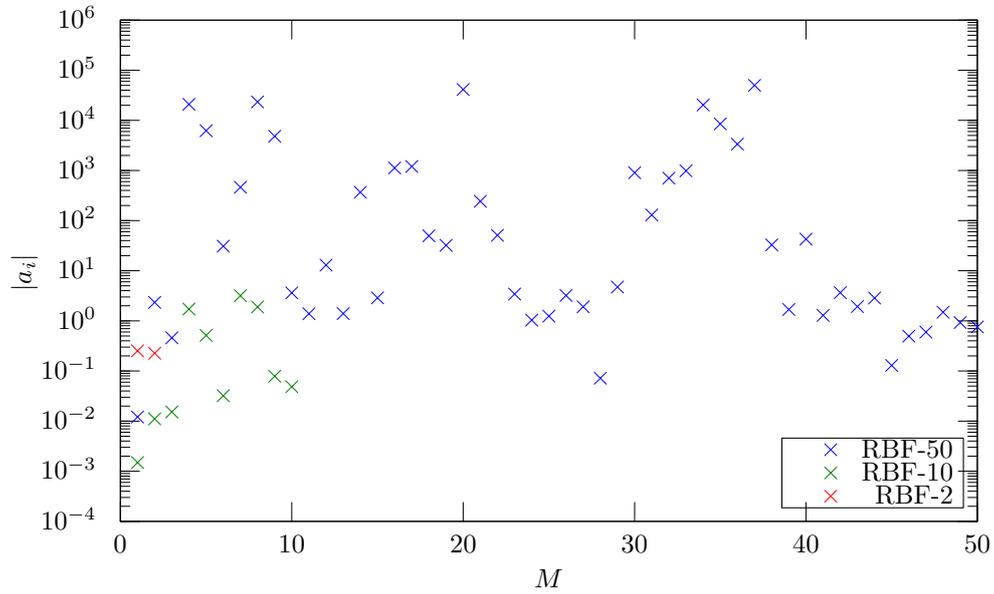
Tabelle 7.7: Daten der radialen Basisfunktionen zur Approximation der Modelltrainingsdaten A

a_i (RBF-2)	a_i (RBF-10)	a_i (RBF-50)	λ_i	$(\ u\)_i$	$(\ \nabla^s u\)_i$	$(\text{dist}(\Gamma_{\text{obs}}, x))_i$
1.9810e-03	1.9810e-03	1.8126e-03	2.8817e-01	6.5276e-01	5.1234e+01	1.3021e-01
2.3418e-03	2.3418e-03	9.8584e-04	3.0626e-01	3.9296e-01	2.7726e+01	1.2760e-01
	1.7231e-03	1.4974e-03	6.5059e-02	1.8265e+00	7.4069e+01	1.0417e-02
	1.1217e-03	1.0310e-03	2.1627e-02	1.4522e+00	1.6164e+02	1.3021e-02
	1.4891e-03	1.2585e-03	7.6232e-02	2.7695e-01	9.7294e+01	1.0417e-02
	2.1589e-03	1.9656e-03	3.5935e-01	3.7919e-01	1.0906e+01	4.2448e-01
	2.1870e-03	1.4991e-03	4.3337e-01	9.1018e-01	3.8949e+01	1.3542e-01
	4.4625e-04	8.9572e-05	4.7104e-03	1.2991e+00	2.7861e+02	4.1667e-02
	1.2273e-04	1.0051e-07	1.6704e-03	1.4652e+00	3.6841e+02	3.6458e-02
	1.3283e-03	1.0162e-03	6.9810e-02	5.0488e-01	1.1998e+02	6.5104e-02
		5.3526e-04	1.0198e-02	6.3373e-01	2.0189e+02	1.3021e-02
		1.1957e-03	1.0599e+00	5.8036e-01	2.5420e+00	4.5573e-01
		2.0844e-03	4.5542e-01	1.2014e+00	1.8854e+01	9.6354e-02
		8.8511e-04	4.2878e-02	1.7320e+00	1.3508e+02	1.4323e-01
		1.5897e-03	4.0116e-01	8.7073e-01	5.8814e+01	1.8229e-01
		1.6073e-03	6.7544e-01	8.9872e-01	4.4433e+01	1.0417e-01
		1.1936e-03	6.7839e-01	8.6009e-01	3.3725e+01	1.5365e-01
		3.9826e-04	6.1103e-03	1.8336e+00	2.3427e+02	4.1667e-02
		1.3355e-03	4.5079e-01	3.1171e-01	5.7017e+00	5.7292e-02
		1.1234e-03	2.6968e-01	8.5975e-01	8.6150e+01	1.7448e-01
		9.5641e-04	1.2094e-01	4.3493e-01	1.0881e+02	1.2760e-01
		1.5898e-03	1.5521e-01	1.1534e-01	6.3543e+01	5.2080e-03
		6.6396e-04	7.5469e-01	6.8425e-01	2.2831e+01	4.7135e-01
		7.4818e-04	2.6808e-02	9.0545e-01	1.7943e+02	4.4271e-02
		2.3567e-03	6.0815e-01	4.5479e-01	1.5189e+01	4.7917e-01
		-4.2104e-04	9.1022e-04	9.7868e-01	3.3050e+02	7.5521e-02
		1.7756e-03	4.1646e-01	1.1012e+00	4.7209e+01	2.5260e-01
		1.7149e-03	4.3920e-01	6.2202e-01	5.5438e+01	1.6927e-01
		7.4460e-04	5.1380e-02	6.3226e-01	1.4936e+02	1.7448e-01
		1.3972e-03	2.1427e-01	7.4954e-01	8.2040e+01	7.2917e-02
		1.2487e-03	1.5722e-01	1.3074e+00	9.0300e+01	1.3802e-01
		1.6925e-03	4.5329e-01	5.2236e-01	4.1859e+01	2.6042e-01
		1.7748e-03	3.0609e-01	3.1924e-01	3.1315e+01	5.2080e-03
		3.2878e-04	4.8506e-03	1.4134e+00	2.5858e+02	4.4271e-02
		1.6344e-03	2.9941e-01	1.3325e+00	3.6182e+01	1.7448e-01
		1.6110e-03	6.9758e-02	2.0944e+00	2.4566e+01	2.7604e-01
		9.8766e-04	1.0703e-01	8.9318e-01	1.0367e+02	5.2080e-03
		9.1578e-04	7.1352e-02	9.5780e-01	1.2714e+02	6.7708e-02
		2.7427e-04	5.6637e-04	1.9288e+00	3.0877e+02	2.8646e-02
		9.1767e-04	1.1550e-01	7.8529e-01	1.1381e+02	1.5885e-01
		1.2895e-03	7.8420e+00	7.5668e-01	7.2206e-01	1.6406e-01
		6.7335e-04	3.9071e-02	6.9237e-01	1.4281e+02	1.0417e-02
		1.1333e-03	5.8136e-01	1.0247e+00	8.0358e+00	9.8958e-02
		1.1987e-03	1.7581e-01	3.5301e-01	6.8184e+01	2.6040e-03
		4.4669e-04	6.4830e-03	1.1266e+00	2.1947e+02	2.8646e-02
		3.4606e-05	2.3237e-03	1.4180e+00	4.6451e+02	1.4583e-01
		4.6284e-04	9.3656e-03	1.5259e+00	1.8944e+02	1.0417e-02
		-1.8886e-05	1.1527e-03	1.1736e+00	2.8799e+02	5.2080e-03
		4.3456e-04	5.2510e-04	1.6254e+00	3.2593e+02	3.6458e-02
		-1.6533e-04	1.0269e-03	1.4026e+00	3.0336e+02	1.8229e-02

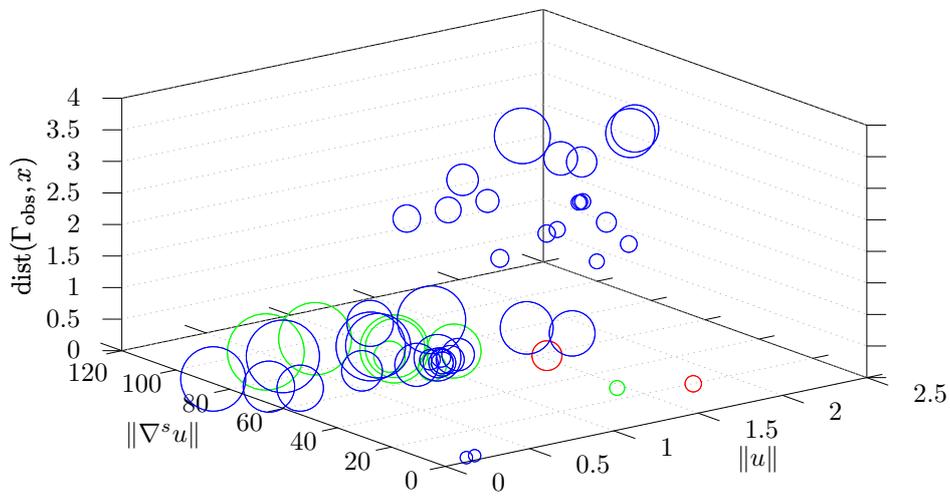
Tabelle 7.8: Daten der radialen Basisfunktionen zur Approximation der Modelltrainingsdaten B

a_i (RBF-2)	a_i (RBF-10)	a_i (RBF-50)	λ_i	$(\ u\)_i$	$(\ \nabla^s u\)_i$	$(\text{dist}(\Gamma_{\text{obs}}, x))_i$
5.4248e-04	1.3769e-04	4.7093e-05	2.7677e-01	1.0497e+00	2.2085e+01	1.2500e-01
4.0116e-04	-1.9496e-05	9.9932e-05	1.5598e-01	4.6618e-01	9.5754e+00	6.2500e-02
	1.3777e-04	-4.9517e-05	1.1238e-01	1.4783e+00	4.0772e+01	1.8750e-01
	1.1301e-04	-2.3085e-05	1.5827e-02	1.9187e+00	6.6121e+01	3.7500e-01
	-1.9093e-04	-3.0008e-04	4.9089e-04	1.7537e+00	3.5395e+02	4.3750e-01
	-2.8906e-04	-4.5944e-03	7.3739e-05	9.9929e-01	6.8894e+02	3.7500e-01
	-3.9866e-02	-7.3581e+01	1.0932e-06	3.8795e+00	1.8229e+03	3.7500e-01
	2.3529e-02	-1.2550e+02	1.1670e-06	3.0121e+00	1.8365e+03	1.2500e-01
	2.6195e-02	3.2927e+02	6.5463e-07	1.2119e+01	1.8946e+03	3.7500e-01
	-8.8592e-03	-8.0064e+02	5.6071e-07	6.6666e+00	1.7255e+03	1.2500e-01
		-4.1274e+02	4.1660e-07	3.3769e+00	2.7597e+03	3.7500e-01
		4.7602e+02	3.9523e-07	4.3493e+00	2.4374e+03	1.2500e-01
		6.8088e+01	6.2078e-07	4.6727e+00	2.1109e+03	3.7500e-01
		-1.4342e+02	5.9147e-07	4.3305e+00	2.3121e+03	1.2500e-01
		-7.5635e+01	1.0571e-06	1.3760e+00	1.9590e+03	3.7500e-01
		-8.6223e+01	4.9777e-07	5.9870e+00	2.3442e+03	3.7500e-01
		-3.8841e+01	6.8867e-07	1.9301e+00	2.0755e+03	3.7500e-01
		2.9197e+02	6.9654e-07	5.3617e+00	2.1067e+03	3.1250e-01
		-1.5879e+02	1.0246e-06	1.0019e+01	2.0346e+03	3.7500e-01
		-1.4241e+02	5.6359e-07	1.0654e+00	2.3237e+03	3.7500e-01
		-1.8247e+02	7.0050e-07	1.7648e+00	2.3198e+03	3.7500e-01
		2.0675e+00	1.3459e-06	1.7070e+00	1.8142e+03	1.2500e-01
		4.0722e+02	1.5710e-06	1.4493e+00	1.8246e+03	1.2500e-01
		-6.2918e+01	6.3826e-07	5.0634e+00	1.9418e+03	3.7500e-01
		1.9110e+02	5.7986e-07	3.3704e+00	1.9168e+03	3.7500e-01
		3.9371e+01	1.2838e-06	9.9673e-01	1.8307e+03	3.7500e-01
		2.1876e+01	4.2364e-06	1.1865e+01	1.0538e+03	3.7500e-01
		-3.8210e+01	2.3403e-06	8.4635e+00	1.2998e+03	1.8750e-01
		-4.1408e+01	3.4607e-06	1.0809e+01	1.1140e+03	3.7500e-01
		-2.1866e+02	1.0897e-06	7.5585e+00	1.7322e+03	3.1250e-01
		1.8459e+02	3.8102e-07	2.2100e+00	2.3670e+03	3.7500e-01
		-4.7668e+01	1.0653e-06	6.3443e+00	1.6316e+03	3.1250e-01
		-3.3478e+02	6.2288e-07	9.2594e-01	1.9663e+03	3.7500e-01
		-2.5213e+02	1.5578e-06	4.5503e+00	1.6553e+03	3.7500e-01
		4.6123e+02	6.0598e-07	6.2376e+00	1.8522e+03	3.7500e-01
		-1.5968e+02	3.6889e-06	1.1346e+01	1.1920e+03	1.8750e-01
		1.0525e+01	6.6219e-06	1.0700e+01	1.0437e+03	3.1250e-01
		6.3344e+01	2.3319e-06	9.9365e+00	1.3627e+03	1.8750e-01
		-4.8005e+00	6.4541e-06	7.7194e-01	1.5878e+03	3.7500e-01
		7.6146e+01	3.0605e-06	1.0104e+01	1.1926e+03	4.3750e-01
		2.6022e+02	5.9681e-07	1.5630e+00	1.9908e+03	3.7500e-01
		2.1570e+02	5.6176e-07	2.5332e+00	1.9598e+03	3.7500e-01
		1.6346e+02	2.9036e-06	8.7240e+00	1.2236e+03	1.8750e-01
		6.6968e+00	9.1095e-06	1.1242e+01	9.0244e+02	3.7500e-01
		1.0717e+02	2.6831e-06	6.2493e+00	1.3752e+03	3.1250e-01
		-1.1500e+02	2.6821e-06	1.0188e+00	1.4741e+03	3.7500e-01
		7.4717e+01	2.0730e-06	4.6980e+00	1.7120e+03	3.7500e-01
		1.2410e+00	1.2264e-05	9.1146e+00	7.0294e+02	3.1250e-01
		-2.0308e+01	6.5557e-06	1.0475e+01	9.1033e+02	3.7500e-01
		-2.5079e-03	7.6995e-05	1.0296e+01	2.9793e+02	3.1250e-01

Tabelle 7.9: Daten der radialen Basisfunktionen zur Approximation der Modelltrainingsdaten C

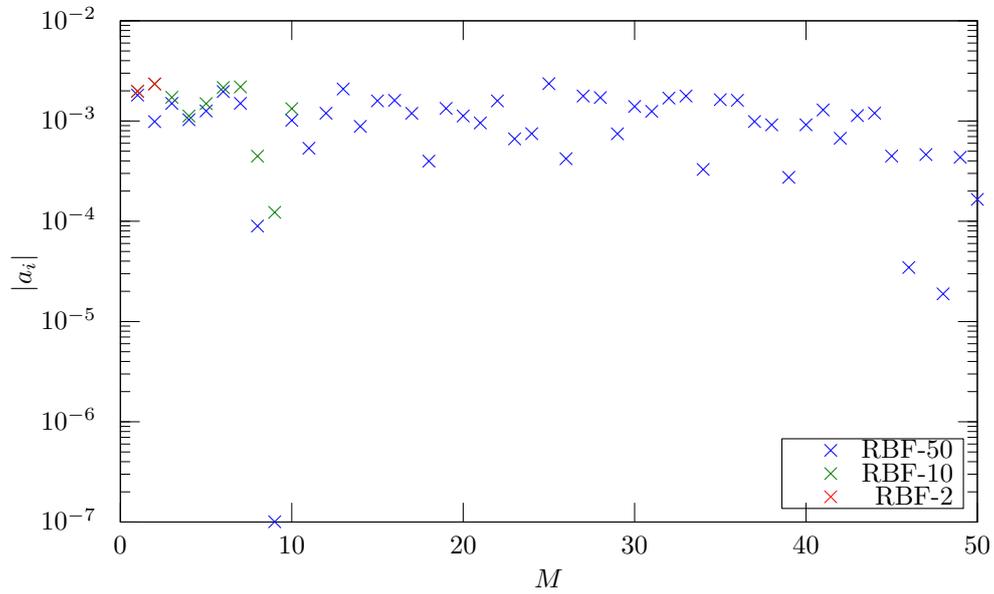


(a)

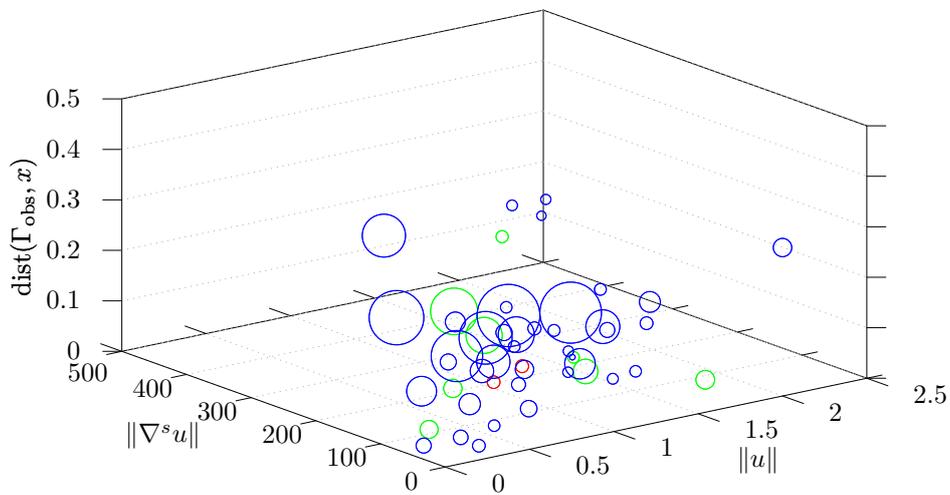


(b)

Abbildung 7.30: Approximation durch radiale Basisfunktionen der Modelltrainingsdaten A: (a) Koeffizienten der Ansatzfunktionen und (b) Wahl der Ansatzfunktionen. Die Fläche der Kreise in (b) sind proportional zu $\frac{1}{\sqrt{\lambda_i}}$.

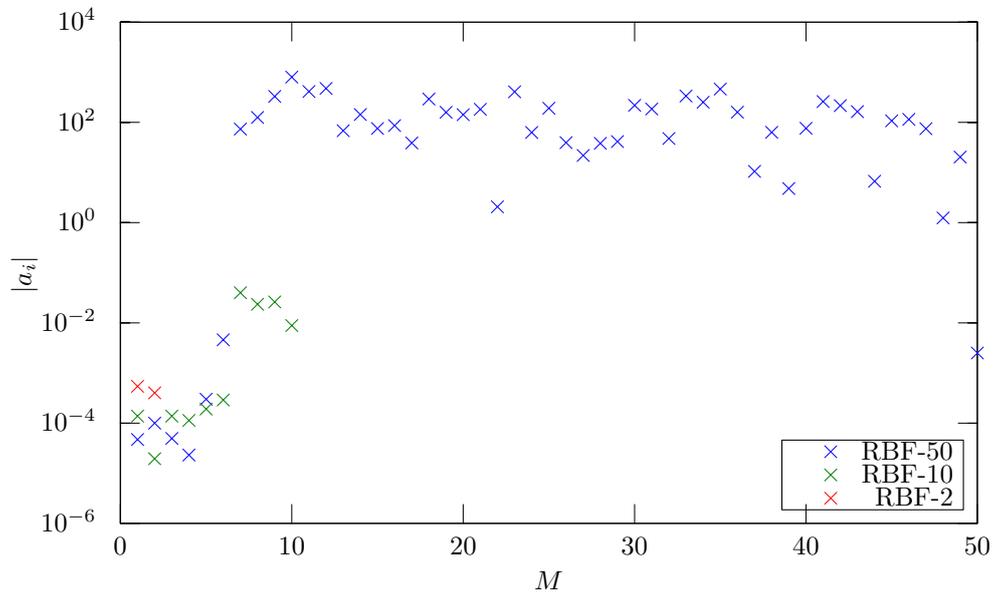


(a)

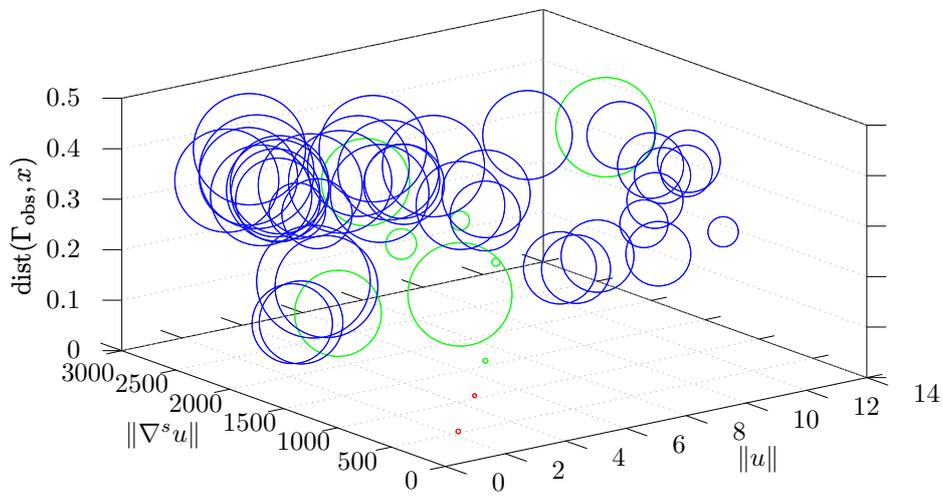


(b)

Abbildung 7.31: Approximation durch radiale Basisfunktionen der Modelltrainingsdaten B: (a) Koeffizienten der Ansatzfunktionen und (b) Wahl der Ansatzfunktionen. Die Fläche der Kreise in (b) sind proportional zu $\frac{1}{\sqrt{\lambda_i}}$. Die ersten beiden a_i von RBF-10 sind in (a) nahezu deckungsgleich zu denen von RBF-2.



(a)



(b)

Abbildung 7.32: Approximation durch radiale Basisfunktionen der Modelltrainingsdaten C: (a) Koeffizienten der Ansatzfunktionen und (b) Wahl der Ansatzfunktionen. Die Fläche der Kreise in (b) sind proportional zu $\frac{1}{\sqrt{\lambda_i}}$.

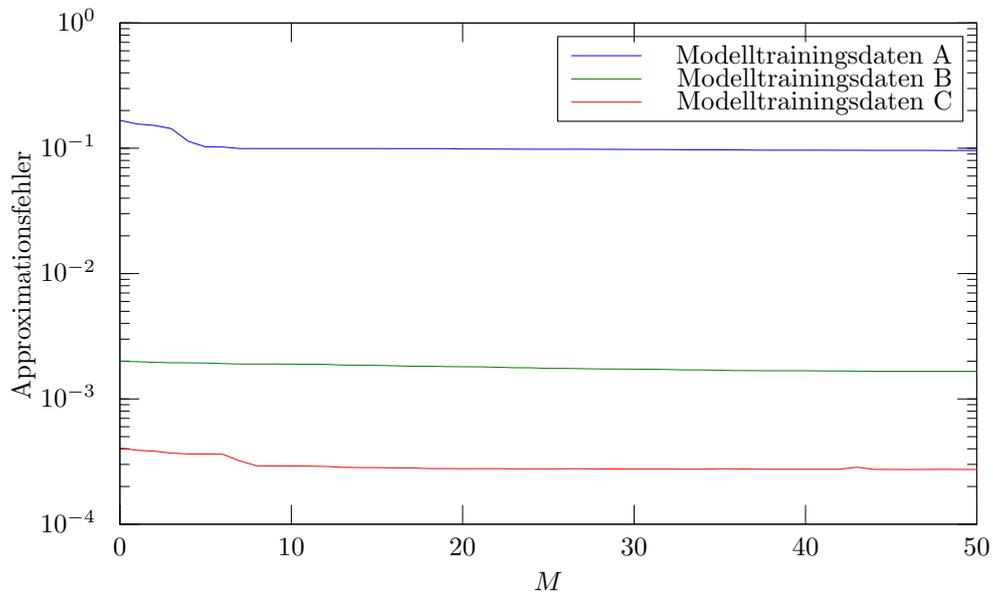


Abbildung 7.33: Fehler bei Approximation der Modelltrainingsdaten A, B und C mit radialen Basisfunktionen

Approximation	keine	RBF-2	RBF-10	RBF-50
Modellproblem A	1.6702e-01	1.524e-01	9.946e-02	9.588e-02
Modellproblem B	1.9966e-03	1.958e-03	1.894e-03	1.661e-03
Modellproblem C	4.0525e-04	3.831e-04	2.913e-04	2.731e-04

Tabelle 7.10: Fehler bei Approximation der Modelltrainingsdaten A, B und C mit radialen Basisfunktionen

Es ist also festzuhalten, dass die RBF-50 Approximationen für die Modellprobleme A und C zwar formell näher an den vorgegebenen Daten des inversen Problems liegen als etwa die RBF-10 Approximationen (Tabelle 7.10). Die Koeffizienten erscheinen jedoch fragwürdig.

Vergleichen wir dazu die Ergebnisse der Approximation mittels vorgegebenen Ansatzfunktionen (Abbildung 7.23 bzw. Tabelle 7.6), ist festzustellen, dass mit 10 bzw. 50 Basisfunktionen ähnlich gute Annäherungen erzielt werden, wie mit den Approximationen 2 oder 3. Je nach Modellproblem sind hier mal die festen Ansatzfunktionen und mal die radialen Basisfunktionen ein wenig besser. Die Näherungen RBF-2 sind ungefähr auf dem Level der Approximation 4 (Smagorinsky-Modellierung).

Im Vergleich zu der Modellierung durch Reduzierung der Basis ist auch hier festzuhalten, dass es sich bei RBF-50 (zum Teil sogar RBF-10 und RBF-2) um bessere als die dort vorgestellten räumlich lokalen Approximationen handelt. Gegenüber den räumlich globalen Approximationen ist hier keine eindeutige Klassifizierung möglich, jedoch schafft es beispielsweise RBF-50 bei den Modelltrainingsdaten C, auf eine Approximationsgüte wie bei der Singulärwertzerlegung mit 20 Basisfunktionen zu gelangen. Die Modelltrainingsdaten B konnten auch hier nicht wirklich brauchbar approximiert werden.

7.6.1 Verwendung der Abbildungen als Viskositätsmodell

Wenden wir die mithilfe radialer Basisfunktionen berechneten Modelle als künstliche Viskosität der Strömung an, so ergeben sich für das beobachtete Funktional \hat{j} Werte, die den Abbildungen 7.34, 7.35 bzw. 7.36 entnommen werden können. Interessanterweise schneidet hier bereits die Approximation mit nur zwei Basiselementen vergleichsweise gut ab. Werden mehr Basiselemente hinzugenommen, ähneln die Verläufe denen der Approximation 2 und 3 aus Abschnitt 7.5.

Bei Modellproblem A (vgl. Abbildung 7.34) liefert RBF-2 ein sehr brauchbares Verhalten. Wenngleich der Funktionalwert mehr oszilliert als bei der feinskaligen Rechnung, so sind die Schwankungen doch deutlich moderater als bei der unstabilisierten grobskaligen Rechnung. Auch ergibt sich hier im zeitlichen Mittel eine ziemlich gute Übereinstimmung mit dem Ziel. Das Funktional der Rechnung mit Modellierung durch RBF-10 liefert hier ein extrem stabilisiertes Bild: Nach dem Aufbau der Strömung ergibt sich ein konstanter Funktionalwert, die Strömung ist offenbar laminar und stationär. Bei RBF-50 zeigt sich zunächst ein ähnliches Verhalten. Allerdings gibt es hier kleine Variationen, die sich über den zeitlichen Verlauf weiter verstärken und ein „Wackeln“ hervorrufen. Selbst im zeitlichen Mittel gibt es hier noch deutlichere Abweichungen des Funktionalwertes von der Referenz als bei RBF-10.

Bei Modellproblem B (vgl. Abbildung 7.35) bewirken alle RBF-Modellierungen, dass der laminare Strömungsaufbau länger vollzogen wird als bei den unstabilisierten Rechnungen. RBF-2 bewirkt noch einen vergleichsweise guten Strömungseinbruch, lässt die Strömung aber dann ein weiteres Mal genauso wie zuvor aufbauen. Bei RBF-10 erfolgt der Abfall der Strömung nicht nur später, sondern auch langsamer und nicht so stark. Anschließend baut sie sich auf einem deutlich zu hohen Niveau erneut auf. Die Modellierung mit RBF-50 scheint erst ganz am zeitlichen Ende der Berechnung mit dem Aufbau der Strömung fertig zu sein, der Übergang wäre vermutlich ab $t = 5$ erfolgt. So jedoch sind die Funktionalwerte viel zu hoch.

Bei Modellproblem C (vgl. Abbildung 7.36) schneiden sowohl RBF-2 als auch RBF-10 recht

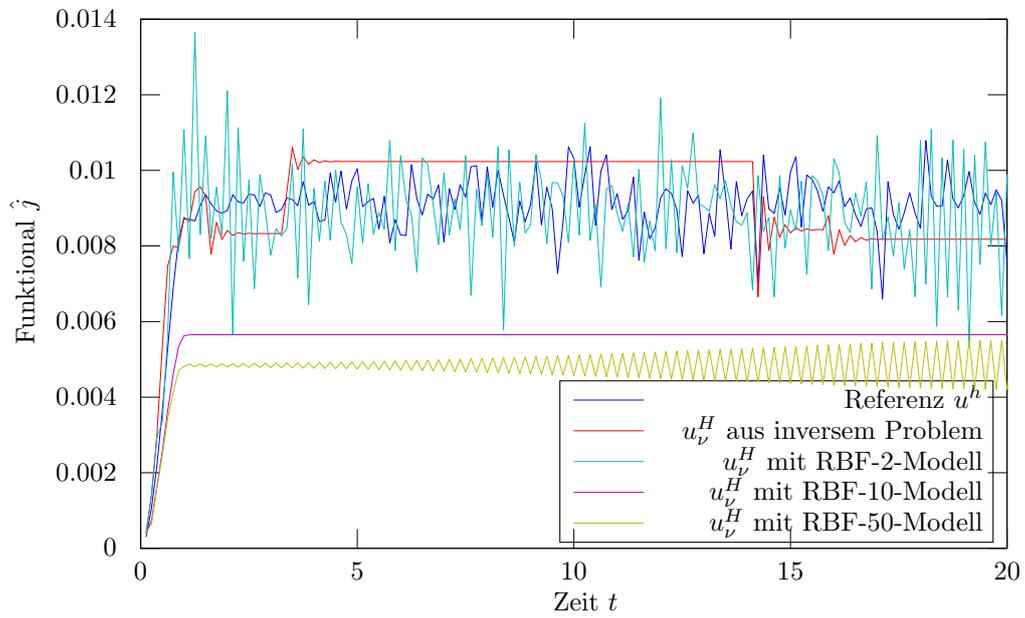


Abbildung 7.34: Entwicklung des Funktional über die Zeit bei Modellproblem A und lokaler Regression mit 2, 10 bzw. 50 radialen Basisfunktionen

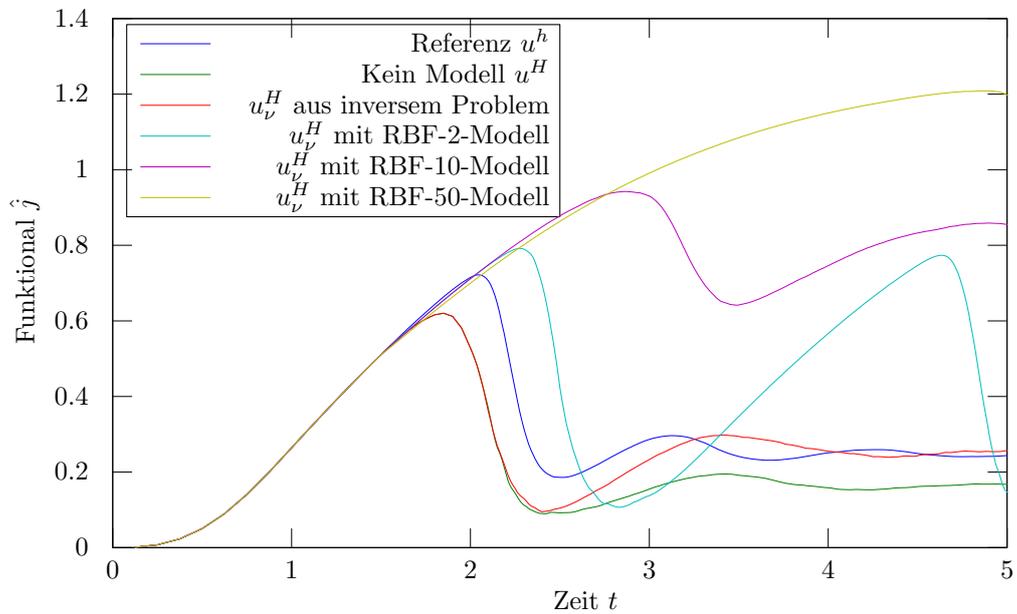


Abbildung 7.35: Entwicklung des Funktional über die Zeit bei Modellproblem B und lokaler Regression mit 2, 10 bzw. 50 radialen Basisfunktionen

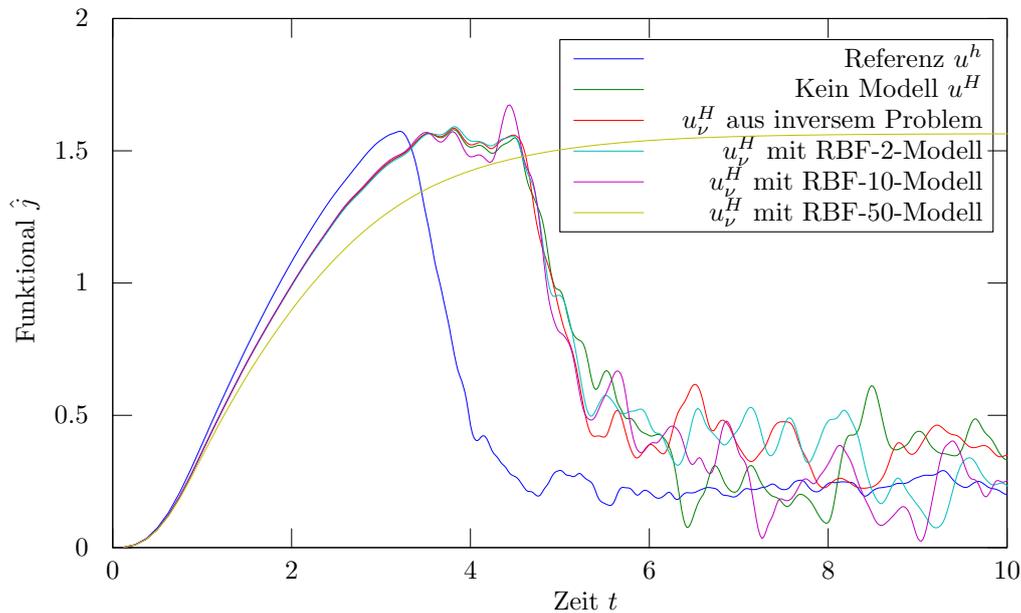
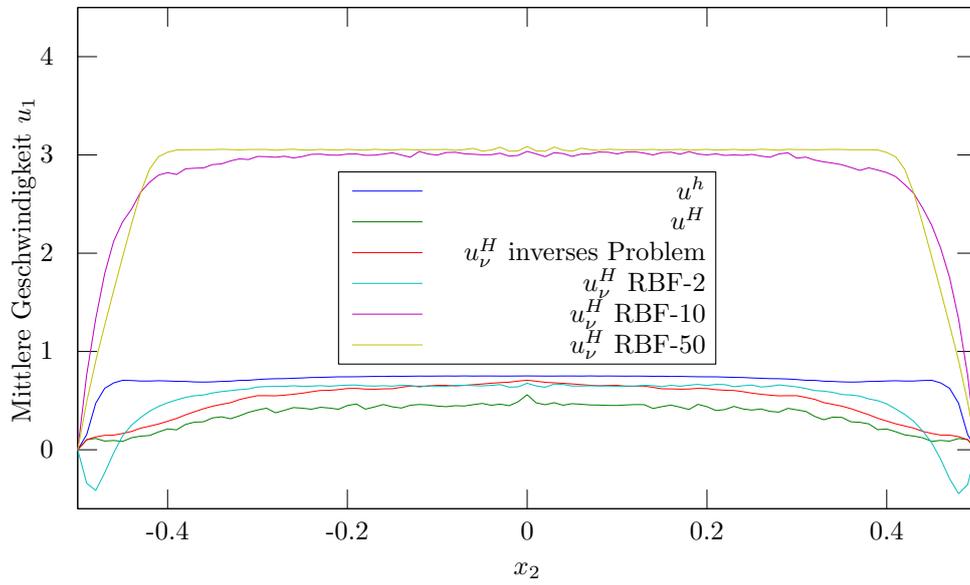


Abbildung 7.36: Entwicklung des Funktionals über die Zeit bei Modellproblem C und lokaler Regression mit 2, 10 bzw. 50 radialen Basisfunktionen

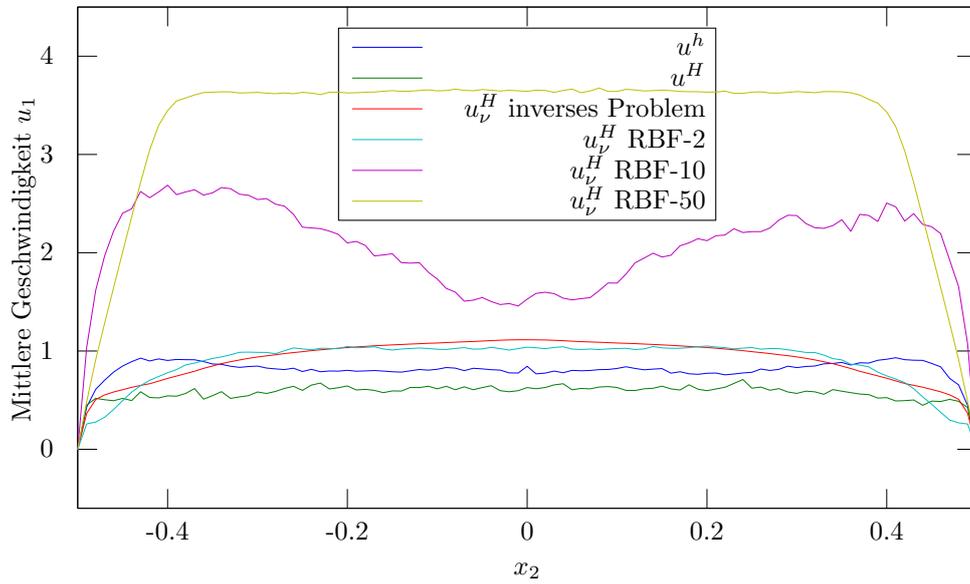
gut ab. Insbesondere ist zur grobskaligen Rechnung ohne Modellierung kein wesentlicher Unterschied auszumachen. Dies war auch schon für die Rechnung des inversen Problems der Fall, also nicht weiter verwunderlich. Auffällig ist die Rechnung mit der Modellierung nach RBF-50, hier erfolgt der Strömungsaufbau langsamer und es kommt gar nicht erst zu einem Einknicken der Strömung. Dieses ungewöhnliche Verhalten könnte mit der Beobachtung von Bemerkung 7.6.1 (Seite 118) in Zusammenhang stehen.

Für das Modellproblem B lässt sich wieder zusätzlich der Strömungsverlauf im Röhrenquerschnitt betrachten, siehe Abbildung 7.37. Wie bereits in Abschnitt 7.5 festgestellt, zeigt die Berechnung mit Viskosität aus dem inversen Problem eine negative Rückströmung im Randbereich. Als einzige Approximation mit diesem Verhalten (wenn auch etwas später) ist nur RBF-2 zu verzeichnen, wie in Abbildung 7.37(a) zu sehen ist. Selbst bei dem insgesamt besser abschneidenden Smagorinsky-Modell ist hier kein derartiges Verhalten erkennbar.

In den Strömungsquerschnitten von Modellproblem C (Abbildung 7.38) ist deutlich zu erkennen, dass die Rechnung mit 50 Basisfunktionen weit abgeschlagen von den anderen Profilen ist. Die Rechnung mit 2 bzw. 10 radialen Basisfunktionen hingegen kann der Vorgabe durch die Modelltrainingsdaten gut folgen und liefert Strömungsprofile, die sogar knapp besser sind, als jenes der Rechnung mit inversem Problem. Im Unterschied zu Modellproblem B ist hier bei keiner der Rechnungen eine Rückströmung in der Nähe des Randes zu erkennen, allerdings wirkt sich hier die Gitterstruktur noch deutlich auf die Strömungsprofile aus. Der kleinste Einfluss des Gitters auf das Ergebnis ist bei der Modellierung RBF-2 zu erkennen. Dort fällt der Abfall der Geschwindigkeit in der Mitte des Strömungsrohrs am geringsten aus.

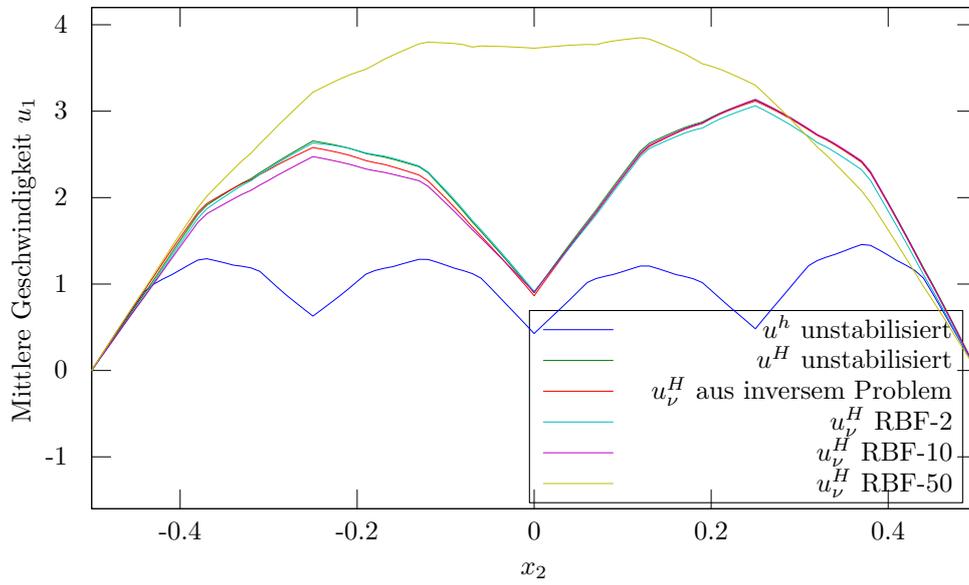


(a)

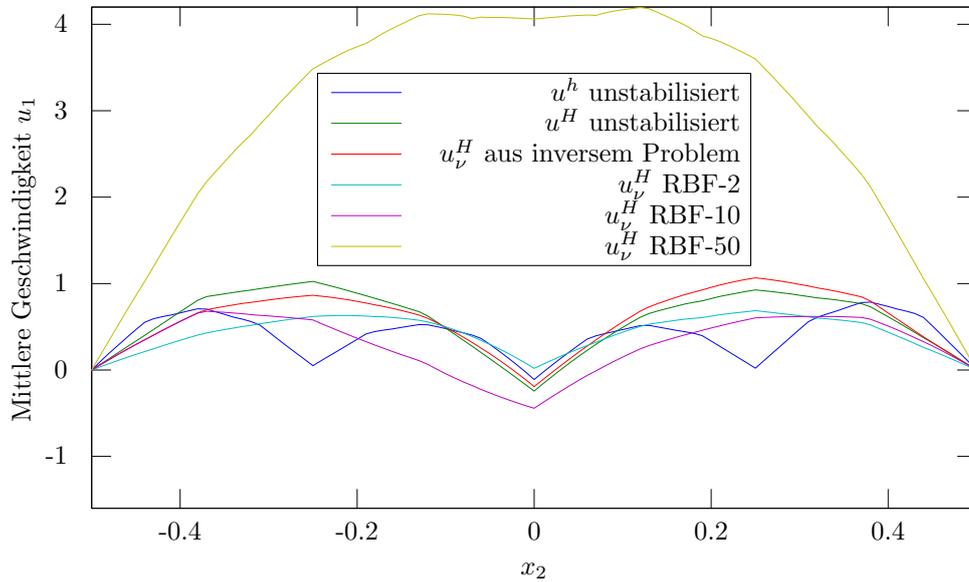


(b)

Abbildung 7.37: Geschwindigkeitsquerschnitte bei Modellproblem B mit den lokalen RBF-Ansätzen über den Zeitraum (a) $t \in (2.5, 3)$ sowie (b) $t \in (3.25, 3.5)$



(a)



(b)

Abbildung 7.38: Geschwindigkeitsquerschnitte bei Modellproblem C mit den lokalen RBF-Ansätzen über den Zeitraum (a) $t \in (3, 6)$ sowie (b) $t \in (8, 10)$

7.7 Zusammenfassung

In den Abbildungen 7.39 bis 7.41 sind verschiedene statistische Größen der Funktionale der verschiedenen Rechnungen im Vergleich dargestellt. Dazu wurden der jeweilige Mittelwert

$$\bar{j}_i = \frac{1}{T} \int_0^T \hat{j}(\hat{u}_{\nu,i}^H),$$

die Standardabweichung

$$\sigma_j = \left(\frac{1}{T} \int_0^T \left(\hat{j}(\hat{u}_{\nu,i}^H) - \bar{j}_i \right)^2 \right)^{\frac{1}{2}}$$

sowie der Fehler des Funktionals in der L_2 -Norm gegenüber der Referenzlösung

$$\varrho_j = \left(\frac{1}{T} \int_0^T \left(\hat{j}(\hat{u}_{\nu,i}^H) - \hat{j}(\hat{u}^h) \right)^2 \right)^{\frac{1}{2}}$$

aufgetragen. Um die instationäre Entwicklung zu berücksichtigen, wurde zusätzlich zu den Abbildungen (a), in denen die obigen Größen jeweils als Mittel über das gesamte zeitliche Intervall $(0, T)$ bestimmt wurden, auch die Abbildungen (b) zum Vergleich gestellt. Die dort betrachteten zeitlichen Intervalle sollen den jeweils quasistationären Fall beschreiben – die meisten der Rechnungen waren in dem entsprechenden Zeitraum eingeschwungen bzw. der Übergang in eine turbulente Strömung war abgeschlossen.

In Abbildung 7.39 ist deutlich zu erkennen, dass die grobskalige Rechnung ohne Modell deutlich von der Feinskaligen abweicht: Alle drei Größen sind hier deutlich zu groß. Sowohl das inverse Problem als auch die Modellierung mithilfe von Smagorinsky- sowie RBF-2-Ansatz liefern für die grobskalige Rechnung deutlich bessere Ergebnisse. Die Modellierungen RBF-10, RBF-50, Approximation 2 sowie Approximation 3 liefern allesamt zwar einen ähnlich schlechten Mittelwert wie die Rechnung ohne Modellierung, jedoch ist die Standardabweichung wesentlich näher an der feinskaligen Lösung und insbesondere auch der L_2 -Fehler wesentlich kleiner. Während diese Modellierungen nicht so zufriedenstellend sind wie RBF-2 und Smagorinsky, sind sie einer grobskaligen Rechnung ohne Modell immer noch vorzuziehen.

Die Tendenz, dass die Modellierungen RBF-10, RBF-50, Approximation 2 sowie Approximation 3 nicht so gute Ergebnisse erzielen, lässt sich auch für Modellproblem B festhalten (Abbildung 7.40). Selbst die zuvor noch brauchbare Approximation RBF-2 schneidet hier eher schlecht ab. Lediglich das inverse Problem sowie die Smagorinsky-Modellierung bilden das Verhalten des Funktionals der feinskaligen Rechnung besser ab als die grobskalige Rechnung ohne Modellierung.

In Abbildung 7.41 fällt die Rechnung RBF-50 deutlich auf, die den Übergang in die turbulente Strömung nicht abgebildet hat. Ansonsten ist insbesondere beim Mittelwert von u^h und u^H für das Intervall $[8, 10]$ eine deutliche Abweichung zu erkennen, die beim inversen Problem sowie den übrigen Viskositätsmodellen geringer ausfällt.

Insbesondere das Smagorinsky-Modell erweist sich insgesamt als sehr stabil. Die RBF-2 bzw. RBF-10-Modellierungen lieferten zumindest in bestimmten Fällen noch brauchbare Resultate. Die Approximationen 2 und 3 sowie die RBF-50-Modellierung zeigten hingegen kein zufriedenstellendes Ergebnis, hier wurden zum Teil deutlich größere Fehler verursacht als

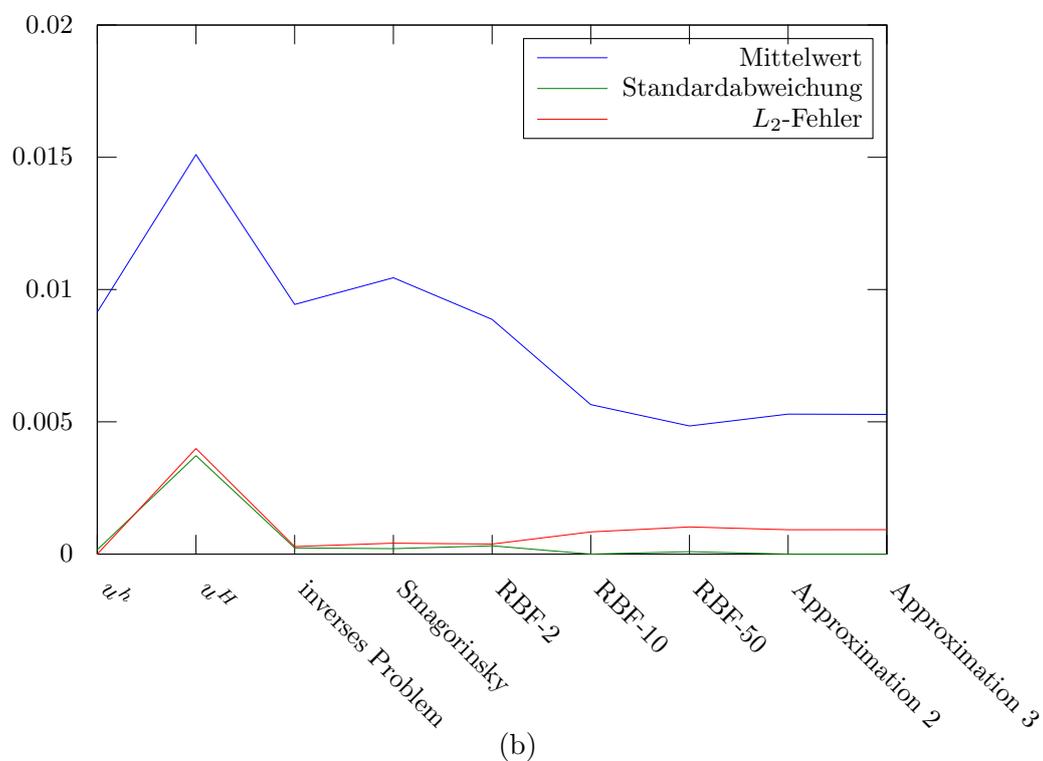
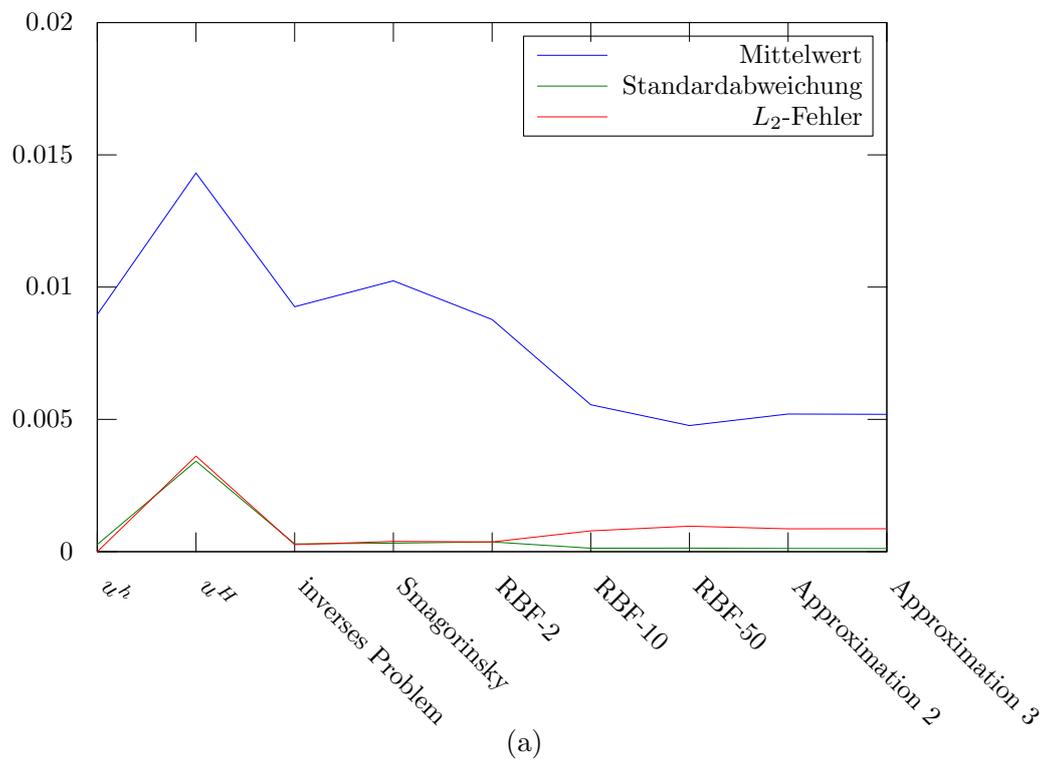


Abbildung 7.39: Zeitliche Mittelung des Funktionals aus den verschiedenen Approximationen bei Modellproblem A (a) über $I = (0, 20)$ und (b) über $I = (2, 20)$

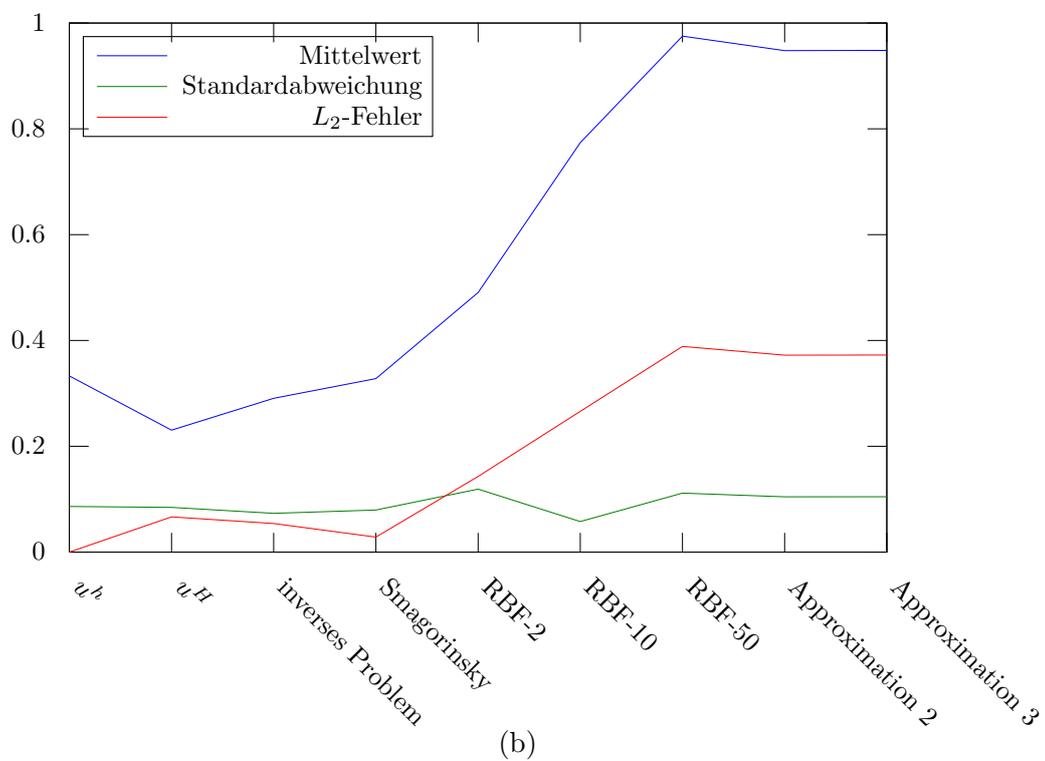
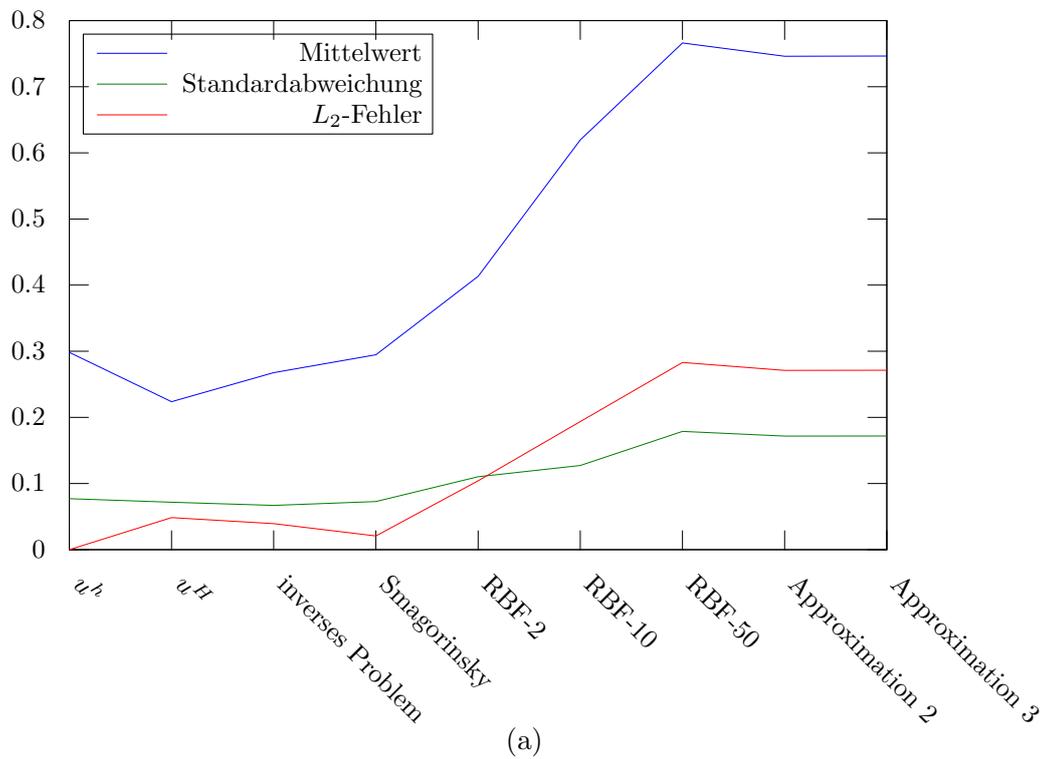


Abbildung 7.40: Zeitliche Mittelung des Funktionals aus den verschiedenen Approximationen bei Modellproblem B (a) über $I = (0, 5)$ und (b) über $I = (1.5, 5)$

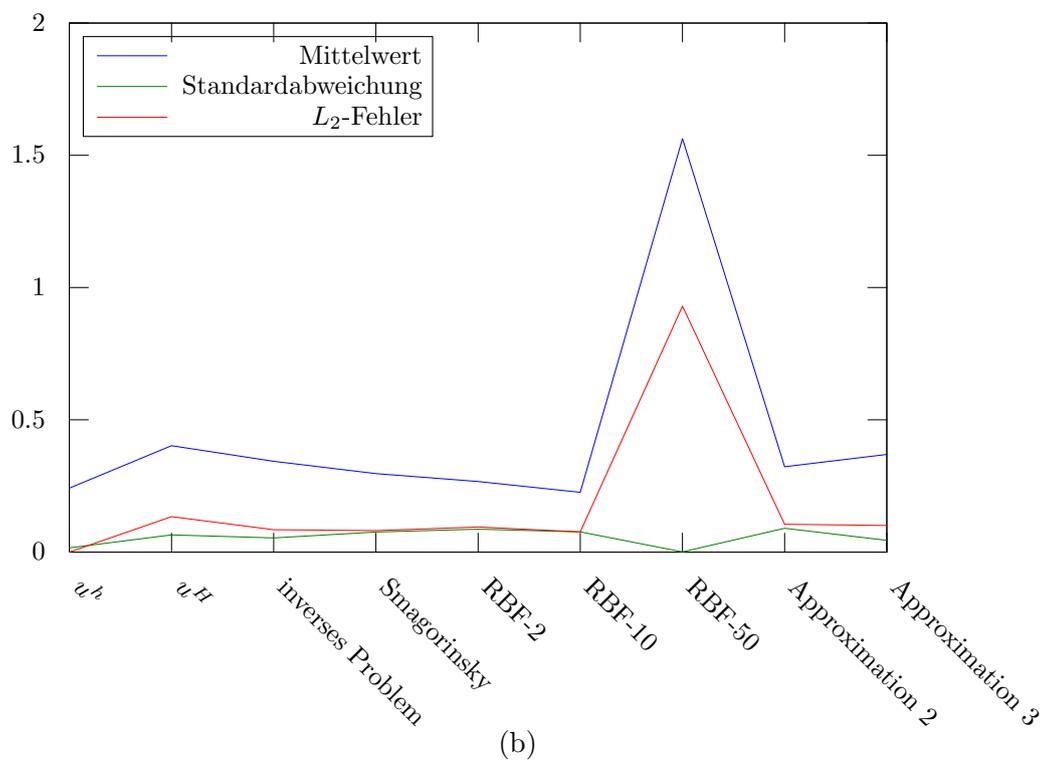
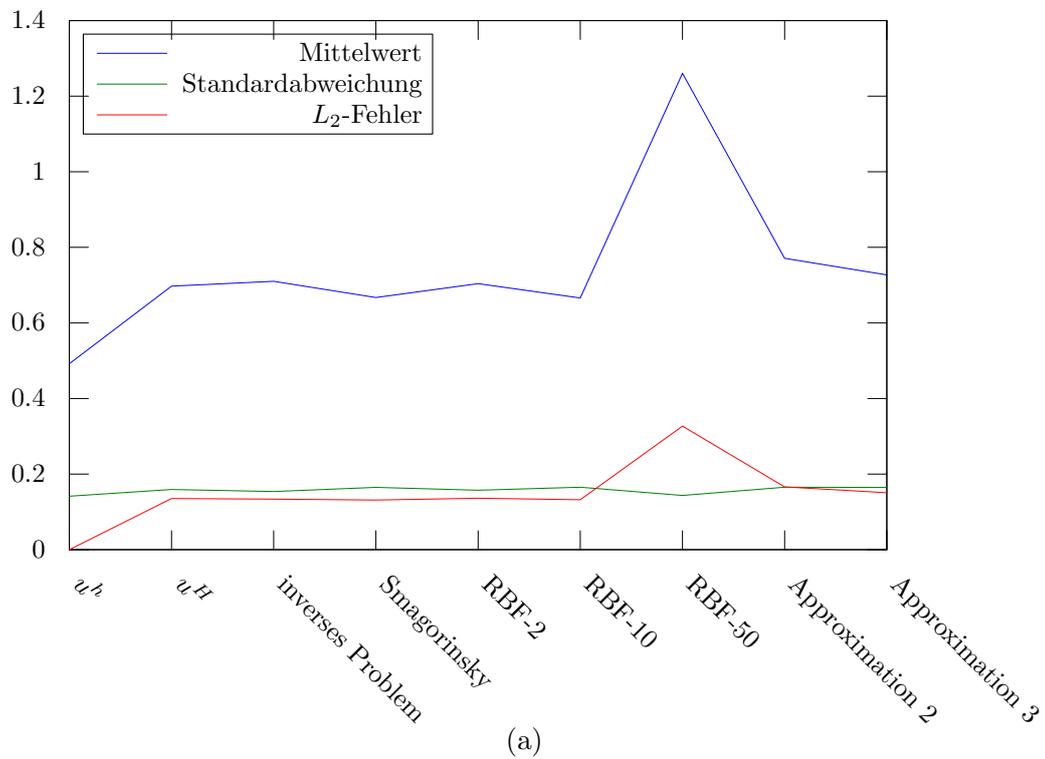


Abbildung 7.41: Zeitliche Mittelung des Funktionals aus den verschiedenen Approximationen bei Modellproblem C (a) über $I = (0, 10)$ und (b) über $I = (8, 10)$

beim Verzicht auf eine Modellierung.

Dies ist bemerkenswert, denn was Approximationsgüte an die Modelltrainingsdaten angeht, sind die Approximationen 2 und 3 sowie RBF-50 deutlich besser als die „einfacheren“ Modelle wie RBF-2, RBF-10 oder Smagorinsky. Wir hätte erwartet, dass etwa die Modellierung RBF-50 irgendwie „zwischen“ RBF-10 und dem inversen Problem liegt, oder entsprechend Approximation 3 zwischen Smagorinsky und dem inversen Problem. Dies ist offensichtlich nicht der Fall, eine Begründung hierfür ist jedoch nicht ersichtlich.

8 Zusammenfassung und Ausblick

Die vorliegende Arbeit stellt einen Beitrag zum optimierten Einsatz von Simulationsressourcen dar. Statt möglichst viel Kapazität darauf zu verwenden, „blind“ durch Erhöhung der Anzahl von Variablen und Gleichungen ein Simulationsproblem möglichst gut aufzulösen (*Direct Numerical Simulation*), wurde der Versuch unternommen, eine Modellierung zu entwickeln, die mit weniger Rechnung auskommt, aber – zumindest in Bezug auf gewisse *relevante* Daten – eine ähnliche Qualität liefert.

Nachdem in den ersten Kapiteln in Notation, betrachtete Probleme und gebräuchliche Techniken eingeführt wurde, erläuterte Kapitel 5 die zentrale Idee der vorliegenden Arbeit: Der dort vorgestellte Ansatz berechnet – ausgehend von einem vorgegebenen Funktional der Simulationsergebnisse – eine Modellierungsgröße. Es handelt sich gewissermaßen um ein inverses Problem, dessen Ziel es ist, den Effekt der feinen Skalen auf das vorgegebene Funktional zu beschreiben. Dies wird durch mehrskalige Rechnung sowie Betrachtung entsprechender dualer Probleme bewerkstelligt. Dieses Verfahren wurde zunächst theoretisch sowie anhand des Beispiels einer gewöhnlichen Differentialgleichung erläutert.

In Kapitel 6 wurde dieses inverse Problem konkret für instationäre inkompressible Strömungen formuliert. Die zu bestimmende Modellierung übernimmt hierbei die Rolle eines Turbulenzmodells. Der beschriebene Algorithmus wurde auf drei verschiedene Modellprobleme angewendet und die Ergebnisse veranschaulicht. Dabei wurde deutlich, dass die Werte des Funktionals bei grobskaliger Rechnung mit Modellierung aus dem inversen Problem sehr viel besser zu den Funktionalwerten der feinskaligen Rechnung passen, als die der grobskaligen Rechnung ohne jegliche Modellierung. Der gesuchte Einfluss der Skalen auf das Funktional konnte in den untersuchten Fällen also größtenteils durch die Modellierung nachgebildet werden.

Schließlich wurde in Kapitel 7 untersucht, inwiefern die Ergebnisse des inversen Problems verwendet werden können, um unmittelbar ein Turbulenzmodell zu formulieren. Die Auswertung eines derartigen Turbulenzmodells wäre wesentlich einfacher als die Berechnung des inversen Problems, auf das dann verzichtet werden könnte. Um derartige Turbulenzmodelle zu gewinnen, wurden die aus Kapitel 6 gewonnenen Daten aus dem inversen Problem der drei Modellprobleme verwendet. Dazu wurden Approximationen mit reduzierter Basis, vorgegebenen Ansatzfunktionen sowie radialen Basisfunktionen untersucht und anschließend auch als Turbulenzmodell in der grobskaligen Rechnung verwendet. Interessanterweise haben hier die einfacheren Turbulenzmodelle wie Smagorinsky oder RBF-2 deutlich bessere Ergebnisse erzielt als die mit zum Teil deutlich besserer Approximationsgüte. Folglich ist aus den Ergebnissen hier kein neues Turbulenzmodell hervorgegangen, das für den Einsatz bei einer Simulation mit vorgegebenem Funktional empfehlenswerter wäre. Allerdings können wir festhalten, dass die vorgestellte Vorgehensweise eine Möglichkeit darstellt, konkrete Parameter eines herkömmlichen Turbulenzmodells zu bestimmen. Immerhin ist die Wahl der konkreten Zahlen an dieser Stelle ohne weitere Referenzen nicht trivial.

8.1 Konkrete in der Arbeit aufgetretene Beschränkungen

In Abschnitt 6.1 wurde bereits diskutiert, dass die Vorgehensweise der zeitlichen Diskretisierung und Untersuchung des stationären dualen Problems eine Vereinfachung ist, die vor allem aus Gründen der Praktikabilität gemacht wurde. Es ist nicht unwahrscheinlich, dass insbesondere die instationären Diskrepanzen zwischen feinskaliger und grobskaliger Lösung des inversen Problems hierdurch verstärkt wurden. Mit steigender Rechenleistung ist zu erwarten, dass die Behandlung des kompletten instationären Problems (sowie dessen dualen Problems) in Zukunft handhabbarer wird und der vorgestellte Ansatz darauf angewendet werden kann.

In Abschnitt 7.4.3 wurde darauf verwiesen, dass die verwendeten Modelltrainingsdaten für einige bestimmte Verfahren der Approximation in gebräuchlicher Form zu groß sind. Bei den Modelltrainingsdaten B war die darzustellende Matrix zu groß, um sie mithilfe einer gewöhnlichen Singulärwertzerlegung analysieren zu können. Dies ist aber sicherlich mit spezialisierter Technik auch heute möglich: Denkbar ist etwa ein Front-Lösungs-Verfahren (vgl. [Iro70]), bei dem die Matrix jeweils nur stückweise im Speicher gehalten werden muss, und die Lösung iterativ korrigiert wird. Auch eine Darstellung und Berechnung mithilfe sogenannter hierarchischer Matrizen (vgl. [Hac09]) ist hier denkbar. Die erwähnten Verfahren wurden in dieser Arbeit jedoch nicht angewendet, da dies erheblichen Mehraufwand in der Implementierung bedeutete und der Fokus der Untersuchung auf anderen Gebieten lag.

8.2 Mögliche Weiterführungen der grundlegenden Idee

Als mögliche Weiterentwicklung sei an dieser Stelle noch einmal darauf hingewiesen, dass der in Kapitel 5 vorgestellte Ansatz der Modellierungsgewinnung aus einem vorgegebenen Funktional grundsätzlich auch auf ganz andere Probleme angewendet werden kann. Neben anderen Differentialgleichungen und anderen Modellierungsansätzen sind auch andere Probleme denkbar, die sich variationell formulieren lassen. Möglicherweise muss dabei besondere Aufmerksamkeit auf die konkrete Lokalisierung der variationellen Formulierung gewidmet werden, dies sollte jedoch kein grundlegendes Hindernis darstellen.

Bei der Diskretisierung des Verfahrens aus Kapitel 5 tritt die Freiheit auf, die Funktionsräume der feinen und groben Diskretisierung geeignet zu wählen. Um die Anzahl der möglichen Variationen gering zu halten, wurden in dieser Arbeit lediglich derartige Diskretisierungen untersucht, bei denen die feine Diskretisierung durch einen Schritt der Gitterverfeinerung gegenüber der groben Diskretisierung entstehen. Hier sind aber auch mehrere Schritte, eine Verfeinerung des lokalen Polynomgrades, sowie eine Kombination dieser Verfeinerungstechniken denkbar. Eine umfassende Untersuchung dieser Möglichkeiten bleibt an dieser Stelle zukünftigen Arbeiten überlassen.

Auch die in Kapitel 7 vorgestellten Ansätze zur Modellierungsgewinnung erheben keinen Anspruch auf Vollständigkeit. Es gibt dabei sicherlich auch noch weitere Möglichkeiten, die vorgestellten Ideen zu vertiefen. Insbesondere wäre zu untersuchen, woher die beobachtete Diskrepanz zwischen Approximations- und Modellierungsgüte, die in Abschnitt 7.7 aufgezeigt wurde, stammt und ob hier eine bessere Lösung zu finden ist.

Literaturverzeichnis

- [AF03] ADAMS, Robert A. ; FOURNIER, John J.: *Sobolev spaces*. Bd. 140. Academic press, 2003
- [Ale77] ALEXANDER, Roger: Diagonally implicit Runge-Kutta methods for stiff O. D. E.'s. In: *SIAM Journal on Numerical Analysis* 14 (1977), Nr. 6, S. 1006–1021
- [Ang13] ANGERMANN, Lutz: A posteriori estimates for errors of functionals on finite volume approximations to solutions of elliptic boundary value problems. In: BEILINA, Larisa (Hrsg.) ; SHESTOPALOV, Yury V. (Hrsg.): *Inverse Problems and Large-Scale Computation* Bd. 53. New York - Heidelberg - Dordrecht - London : Springer-Verlag, 2013, S. 41–56
- [BBD⁺13a] BASTIAN, Peter ; BLATT, Markus ; DEDNER, Andreas ; ENGWER, Christian ; FAHLKE, Jorrit ; GRÄSER, Carsten ; KLÖFKORN, Robert ; NOLTE, Martin ; OHLBERGER, Mario ; SANDER, Oliver: *DUNE PDELab Web page*. 2013. – <http://www.dune-project.org/pdelab>
- [BBD⁺13b] BASTIAN, Peter ; BLATT, Markus ; DEDNER, Andreas ; ENGWER, Christian ; FAHLKE, Jorrit ; GRÄSER, Carsten ; KLÖFKORN, Robert ; NOLTE, Martin ; OHLBERGER, Mario ; SANDER, Oliver: *DUNE Web page*. 2013. – <http://www.dune-project.org>
- [BBT11] BRAACK, Malte ; BURMAN, Erik ; TASCHEBERGER, Nico: Duality based a posteriori error estimation for quasi-periodic solutions using time averages. In: *SIAM Journal on Scientific Computing* 33 (2011), Nr. 5, S. 2199–2216
- [BE03] BRAACK, Malte ; ERN, Alexandre: A posteriori control of modeling errors and discretization errors. In: *Multiscale Model. Simul.* 1 (2003), Nr. 1, S. 221–238
- [BG06] BOCHEV, Pavel ; GUNZBURGER, Max: Least-squares finite element methods. In: *Proceedings of the International Congress of Mathematicians (ICM)* Springer, 2006, S. 1137–1162
- [Bra07] BRAESS, Dietrich: *Finite elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer, 2007
- [Bru00] BRUNEAU, Charles-Henri: Boundary conditions on artificial frontiers for incompressible and compressible Navier-Stokes equations. In: *Mathematical Modelling and Numerical Analysis* 34 (2000), Nr. 2, S. 303–314
- [Buh03] BUHMANN, Martin D.: *Radial basis functions: theory and implementations*. Bd. 5. Cambridge university press Cambridge, 2003
- [CR73] CROUZEIX, Michel ; RAVIART, Pierre-Arnaud: Conforming and nonconforming finite element methods for solving the stationary Stokes equations, I. In: *RAIRO Analyse Numérique* 7 (1973), S. R-3:33–76

- [DAG⁺11] DENARO, Filippo M. ; ABBÀ, Antonella ; GERMANO, Massimo ; ICARDI, Matteo ; MARCHISIO, Daniele ; ROLFO, Stefano ; LAMPITELLA, Paolo ; COLOMBO, Emanuela ; INZOLI, Fabio ; APROVITOLA, Andrea ; MARRA, Francesco S. ; IOVIENO, Michele ; TORDELLA, Daniela: A comparative test for assessing the performances of large-eddy simulation codes. In: *Proc. XX AIMETA Conference*, 2011
- [DHO12] DROHMANN, Martin ; HAASDONK, Bernard ; OHLBERGER, Mario: Reduced Basis Approximation for Nonlinear Parametrized Evolution Equations based on Empirical Operator Interpolation. In: *SIAM Journal on Scientific Computing* 34 (2012), Nr. 2, S. A937–A969
- [Fef00] FEFFERMAN, Charles L.: Existence and smoothness of the Navier-Stokes equation. In: *The millennium prize problems* (2000), S. 57–67
- [Gal94] GALDI, Giovanni P.: *An Introduction to the Mathematical Theory of the Navier-Stokes Equations: Linearized steady problems*. Springer-Verlag, 1994 (Springer tracts in natural philosophy)
- [Hac09] HACKBUSCH, Wolfgang: *Hierarchische Matrizen: Algorithmen und Analysis*. Springer-Verlag, 2009
- [HJ10] HOFFMAN, Johan ; JOHNSON, Claes: *Computational Turbulent Incompressible Flow: Applied Mathematics: Body and Soul 4*. Springer, 2010 (Applied Mathematics: Body and Soul). – ISBN 9783642079863
- [HRT96] HEYWOOD, John G. ; RANNACHER, Rolf ; TUREK, Stefan: Artificial boundaries and flux and pressure conditions for the incompressible Navier–Stokes equations. In: *International Journal for Numerical Methods in Fluids* 22 (1996), Nr. 5, S. 325–352
- [HW90] HONERKAMP, Josef ; WEESE, Jürgen: Tikhonovs regularization method for ill-posed problems. In: *Continuum Mechanics and Thermodynamics* 2 (1990), Nr. 1, S. 17–30
- [Iro70] IRONS, Bruce M.: A frontal solution program for finite element analysis. In: *International Journal for Numerical Methods in Engineering* 2 (1970), Nr. 1, S. 5–32
- [Joh06] JOHN, Volker: On Large Eddy Simulation and Variational Multiscale Methods in the Numerical Simulation of Turbulent Incompressible Flows. In: *Applications of Mathematics* 50 (2006), Nr. 4, S. 321–353
- [KMM87] KIM, John ; MOIN, Parviz ; MOSER, Robert: Turbulence statistics in fully developed channel flow at low Reynolds number. In: *Journal of fluid mechanics* 177 (1987), S. 133–166
- [Lay08] LAYTON, William J.: *Computational science & engineering*. Bd. 6: *Introduction to the numerical analysis of incompressible viscous flows*. Philadelphia, Pa. : SIAM, 2008
- [MKM99] MOSER, Robert D. ; KIM, John ; MANSOUR, Nagi N.: Direct numerical simulation of turbulent channel flow up to $Re_\tau = 590$. In: *Phys. Fluids* 11 (1999), Nr. 4, S. 943–945

- [MP94] MOHAMMADI, Bijan ; PIRONNEAU, Olivier: *Analysis of the K-Epsilon Turbulence Model*. New York : John Wiley & Sons, 1994
- [MT98] MARION, Martine ; TEMAM, Roger M. ; CIARLET, Philippe G. (Hrsg.) ; LIONS, Jacques-Louis (Hrsg.): *Handbook of Numerical Analysis*. Bd. 6: *Navier-Stokes Equations: Theory and Approximation*. Elsevier Science, 1998
- [OD96] ODEN, J. T. ; DEMKOWICZ, Leszek F.: *Applied Functional Analysis*. 1996
- [OD06] OHLBERGER, Mario ; DEDNER, Andreas: *Wissenschaftliches Rechnen und Anwendungen in der Strömungsmechanik*. 2006. – http://aam.mathematik.uni-freiburg.de/IAM/Teaching/ubungen/sci.com_SS06/skriptum.pdf
- [RA08] RANG, Joachim ; ANGERMANN, Lutz: New Rosenbrock Methods of order 3 for PDAEs of Index 2. In: *Advances in Differential Equations and Control Processes* 1 (2008), Nr. 2, S. 193–217
- [Rem10] REMPKKE, Arne: *Turbulenz-Simulation unter Verwendung variationeller Multiskalenmethoden*. Diplomarbeit, TU Clausthal, 2010
- [Sag02] SAGOUT, Pierre: *Large Eddy Simulation for Incompressible Flows*. Springer, 2002
- [SDT06] SAGOUT, Pierre ; DECK, Sebastian ; TERRACOL, Marc: *Multiscale und Multiresolution Approaches in Turbulence*. Imperial College Press, London, 2006
- [STD⁺96] SCHÄFER, Michael ; TUREK, Stefan ; DURST, Franz ; KRAUSE, Egon ; RANNACHER, Rolf: *Benchmark computations of laminar flow around a cylinder*. Springer, 1996
- [TH73] TAYLOR, Cedric ; HOOD, P.: A numerical solution of the Navier-Stokes equations using the finite element technique. In: *Computers & Fluids* 1 (1973), Nr. 1, S. 73–100