

**Fehlerkontrollierte adaptive h - und
 hp -Finite-Elemente-Methoden
für Kontaktprobleme
mit Anwendungen in der Fertigungstechnik**

Andreas Schröder

ADRESSE DES AUTORS:

Andreas Schröder

Im Löttenkamp 34

59510 Lippetal, Deutschland

e-mail: andreas.schroeder@mathematik.uni-dortmund.de

Diese Arbeit wurde vom Fachbereich Mathematik der Universität Dortmund als Dissertation zur Erlangung des Grades eines Doktors der Naturwissenschaften genehmigt.

1. Gutachter:	Prof. Dr. Heribert Blum
2. Gutachter:	Prof. Dr. Franz-Theo Suttmeier
Tag der Einreichung:	31. 10. 2005
Tag des Kolloquiums:	15. 12. 2005

Inhaltsverzeichnis

Einleitung	v
I Elliptische Minimierungsprobleme	1
I.1 Notwendige und hinreichende Bedingungen, Existenz	2
I.1.1 Notwendige und hinreichende Bedingungen	2
I.1.2 Existenz von Lösungen	5
I.2 Sattelpunktformulierungen	6
I.2.1 Variationelle Formulierung, Existenz von Sattelpunkten	7
I.2.2 Sattelpunktformulierungen für Minimierungsprobleme	9
I.3 Elliptische Minimierungsprobleme	12
I.3.1 Elliptische Minimierungsprobleme erster Art	13
I.3.2 Elliptische Minimierungsprobleme zweiter Art	16
I.3.3 Existenz und Eindeutigkeit von Sattelpunkten	18
I.4 Variationsungleichungen auf Hilberträumen	21
I.4.1 Variationsungleichungen erster Art	22
I.4.2 Variationsungleichungen zweiter Art	24
II Modellierungen von Kontaktproblemen	31
II.1 Sobolevräume	32
II.2 Modellprobleme	35
II.2.1 Vereinfachte Signorini-Probleme	36
II.2.2 Idealisierte Reibungsprobleme	39
II.3 Linear-elastische Kontaktprobleme	41
II.3.1 Reibungsfreie Kontaktprobleme	45
II.3.2 Reibungsprobleme mit vorgegebener Normalspannung	47
II.3.3 Kontaktprobleme mit Reibung	50

II.3.4	Das Coulombsche Reibungsgesetz	51
II.4	Eine Anwendung in der Fertigungstechnik	53
II.4.1	Reibungsfreier Kontakt	57
II.4.2	Reibungsbehafteter Kontakt mit vorgegebenen Normalspannungen	60
II.4.3	Reibungsbehafteter Kontakt	62
II.5	Weitere Modellprobleme	63
II.5.1	Hindernisprobleme	64
II.5.2	Elastoplastische Torsionsprobleme	65
II.5.3	Bingham-Fluid-Probleme	66
III	Diskretisierungen	67
III.1	Finite-Elemente-Methoden	73
III.1.1	Modellprobleme	74
III.1.2	Linear-elastische Probleme	89
III.1.3	Weitere Modellprobleme	93
III.2	h - und hp -Finite-Elemente-Methoden	95
III.2.1	Hierarchische Basen in 2D und 3D	98
III.2.2	Irreguläre Knoten, Kanten, Facetten	105
III.3	Lösungsverfahren	110
III.3.1	Das SOR-Verfahren mit Projektion	111
III.3.2	Lösungsalgorithmen für Sattelpunktprobleme	117
IV	Fehlerkontrolle und Adaptivität	123
IV.1	Ein allgemeiner Zugang	124
IV.1.1	Elliptische Minimierungsprobleme erster Art	127
IV.1.2	Unrestringierte elliptische Minimierungsprobleme zweiter Art	129
IV.1.3	Restringierte elliptische Minimierungsprobleme zweiter Art	130
IV.2	Residuale Fehlerschätzer für Variationsgleichungen	132
IV.2.1	Modellprobleme	133
IV.2.2	Linear-elastische Probleme	138
IV.3	Fehlerschätzer für vereinfachte Kontaktprobleme	140
IV.3.1	Vereinfachte Signorini-Probleme	140
IV.3.2	Idealisierte Reibungsprobleme	144
IV.4	Fehlerschätzer für Kontaktprobleme	146
IV.4.1	Reibungsfreie Kontaktprobleme	146
IV.4.2	Reibungsprobleme mit vorgegebener Normalspannung	148

IV.4.3	Kontaktprobleme mit Reibung	148
IV.5	Fehlerschätzer für weitere Modellprobleme	150
IV.5.1	Hindernisprobleme	150
IV.5.2	Bingham-Fluid-Probleme	153
IV.6	Fehlerschätzer vom Goal-Oriented-Typ	154
IV.6.1	Variationsgleichungen	154
IV.6.2	Modellprobleme	156
IV.6.3	Ein allgemeiner Ansatz für elliptische Minimierungsprobleme erster und zweiter Art	158
IV.7	h - und hp -adaptive Verfeinerungen	159
IV.7.1	h -adaptive Verfeinerungsstrategien	162
IV.7.2	hp -adaptive Verfeinerungsstrategien	163
IV.7.3	Variationsgleichungen	166
IV.7.4	Fehlerschätzer vom Goal-Oriented-Typ	171
IV.7.5	Modellprobleme	179
IV.7.6	Linear-elastische Kontaktprobleme	184
IV.7.7	Hindernisprobleme	189
V	Ausblick	193
A	Beispiele, Kopplungsgewichte, Grundlagen	197
A.1	h - und hp -adaptive Gitter für Variationsgleichungen	197
A.2	Kopplungsgewichte für hierarchische Basen	201
A.3	Funktionalanalytische Grundlagen	203
	Literaturverzeichnis	207

Einleitung

Die Modellierung und die Simulation von Kontaktproblemen sind von zentraler Bedeutung für das Verständnis vieler Ingenieur-technischer Prozesse und für die Entwicklung neuer Produktionsverfahren.

Im Allgemeinen versteht man unter Kontakt die Berührung zweier (oder mehrerer) elastischer bzw. inelastischer Körper in einem Kontaktbereich. Aus strukturmechanischer Sicht verursacht der Kontakt in der Regel Deformationen der beteiligten Körper, die wiederum Reaktionskräfte im Kontaktbereich hervorrufen. Ein Kontaktproblem besteht in diesem Fall in der Erfassung der Kontaktzone, der Deformationen oder der auftretenden Reaktionskräfte. Darüber hinaus können weitere kontaktspezifische Aspekte innerhalb eines Kontaktproblems von Interesse sein.

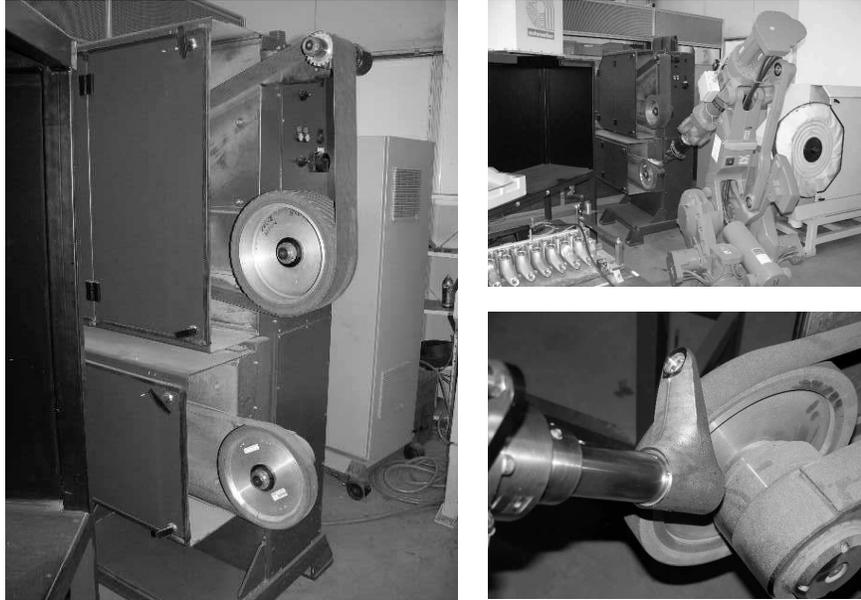


Abb. 1: Bandschleifmaschine mit zwei Kontaktscheiben, Industrieroboter mit Sanitärarmatur (oben), vergrößerter Ausschnitt mit Sanitärarmatur und Kontaktscheibe (unten).

Ein Beispiel aus der Fertigungstechnik

Ingenieur-technische sowie für die Industrie relevante Kontaktprobleme sind insbesondere in der Fertigungstechnik zu finden. Hier ist die Betrachtung von Kontaktproblemen entscheidend für die Gestaltung von Fertigungsverfahren. Dies gilt speziell für Fertigungsverfahren aus dem Bereich Umformen (z.B. Schmieden, Walzen oder Tiefziehen), Trennen (z.B. Schleifen, Fräsen oder Hobeln) oder Beschichten, da hier vor allem der Kontakt des Werkzeugs bzw. des Wirkmediums mit dem Werkstück die gewünschte Geometrieänderung des Werkstücks bewirkt.¹

Als ein repräsentatives Beispiel sei an dieser Stelle das Schleifen und speziell das Industrieroboter-gestützte Bandschleifen genannt. Schleifen ist ein abtragendes Fertigungsverfahren zur Bearbeitung von Oberflächen. Wichtige Einsatzbereiche sind das Glätten und die Feinbearbeitung (z.B. Profilschleifen) von gehärteten, insbesondere freigeformten Flächen.

¹vgl. DIN 8580, [43]

In vielen Industriebranchen, wie z.B. in der Sanitärarmaturindustrie oder im Dampf- und Gasturbinenbau, werden vor allem Bandschleifmaschinen zur Oberflächenbearbeitung eingesetzt.

Bandschleifen ist ein Fertigungsverfahren mit einem vielschneidigen Werkzeug aus Schleifkörnern, die auf einem Schleifband als Unterlage aufgebracht sind, das mindestens zwei rotierende Rollen umläuft und durch eine dieser Rollen (Kontaktscheibe), ein anderes Stützelement oder auch ohne Stützelement an das zu schleifende Werkstück angepresst wird. Die geometrisch unbestimmten Schneiden trennen mit hoher Geschwindigkeit den Werkstoff vom Werkstück ab.² Durch die Verwendung von Kontaktscheibenbelägen aus einem Gummi-elastischen Material können auch komplizierte freigeformte Oberflächengeometrien bearbeitet werden.

In der industriellen Fertigung werden Schleifvorgänge an Bandschleifmaschinen entweder manuell von einem Anlagebediener, der vor allem durch seine praktischen Erfahrungen geleitet wird, oder, dies betrifft insbesondere Branchen mit großen Losgrößen, wie etwa die Sanitärarmaturindustrie, durch Industrieroboter vollzogen.

Beim Industrieroboter-gestützten Bandschleifen führt ein Roboterarm ein Werkstück gegen das Schleifband. Für diese Form der Schleifbearbeitung ist im Vorhinein eine genaue Planung der vom Industrieroboter abzufahrenden Bahn bzw. Positionen zu erstellen, so dass ein möglichst optimales Bearbeitungsergebnis erzielt werden kann. Derartige Bahnplanungen erfordern im hohen Maße genaue Kenntnisse über die Kontaktverhältnisse, die aus dem Kontakt von Werkstück und Schleifband bzw. Kontaktscheibe resultieren. Maßgeblich sind hier vor allem die Kontaktkräfte und die Kontaktzone, mit deren Kenntnis der Abtrag am Werkstück bestimmt werden kann, um den Vorschub sowie die Andruckkraft des Roboterarms entlang der Schleifbahn zu regulieren. Die hinreichend genaue Bestimmung der Kontaktkräfte und der Kontaktzone erfordert im Allgemeinen die Anwendung simulativer, rechnergestützter Verfahren und kann erst mit Hilfe effizienter numerischer Methoden befriedigend durchgeführt werden.

Modellierung

Die Entwicklung von effizienten numerischen Methoden zur Simulation von Kontaktproblemen, wie dem oben beschriebenen Problem aus der Fertigungstechnik, ist ein wichtiger Forschungsbereich der angewandten Mathe-

²vgl. DIN 8589, [44]

matik.

Ausgangspunkt ist stets die Beschreibung des Problems in Form eines mathematischen Modells. Je nach Beschaffenheit des realen Problems, das z.B. durch Materialeigenschaften, Auftreten dynamischer Effekte usw. gekennzeichnet ist, stehen unterschiedliche vor allem mechanisch/physikalisch motivierte Zugänge zur Verfügung.

Im Fall des oben beschriebenen Kontaktproblems des Industrieroboter-gestützten Bandschleifens ist die Modellierung über einen Energieminimierungsansatz angemessen, dem ein linear-elastisches Materialgesetz zugrunde liegt. Hierdurch ist schließlich ein elliptisches Minimierungsproblem zu betrachten, dessen Lösung zum Beispiel die zu ermittelnde Deformation der elastischen Körper beschreibt.³ Die Einarbeitung von geometrischen Kontaktbedingungen, vor allem das Verhindern des gegenseitigen Eindringens der in Kontakt befindlichen Körper, wird durch die Angabe einer geeigneten Restriktionsmenge modelliert. Reibungsnebenbedingungen, wie sie bei vielen Kontaktproblemen und insbesondere beim Industrieroboter-gestützten Bandschleifen zu berücksichtigen sind, werden durch eine entsprechende Modifizierung des Zielfunktional formuliert. Die Betrachtung von sowohl reibungsfreien Kontaktproblemen als auch von Kontaktproblemen mit Reibung macht eine Unterscheidung der zugrunde liegenden elliptischen Minimierungsprobleme in elliptische Minimierungsprobleme erster und zweiter Art sinnvoll.

Ein zu Minimierungsproblemen äquivalenter Modellierungsansatz besteht in der Betrachtung von Kräftebilanzen, die in Form von partiellen Differentialgleichungssystemen dargestellt werden können. Während die Formulierung von Kräftebilanzen von Ingenieur-technischer Seite im Allgemeinen bevorzugt wird, da diese für die mechanische Interpretation der darin enthaltenen Größen zugänglicher sind, ist der Zugang über Minimierungsprobleme vor allem für mathematische Fragestellungen interessant. So können recht allgemein Fragen nach Existenz und Eindeutigkeit von Lösungen mit Hilfe von Aussagen der Konvexen Analysis geklärt werden. Zudem erlangt man über die Betrachtung von notwendigen und hinreichenden Bedingungen weitere Formulierungsansätze in Form von Variationsgleichungen bzw. Variationsungleichungen. Schließlich erhält man unmittelbar aufgrund des Minimierungszugangs eine vierte Darstellungsform des Kontaktproblems in Form von variationellen Sattelpunktformulierungen. Hier werden die Kontaktnebenbedingungen durch Hinzufügen von Lagrangeschen Multiplikatoren erfasst, die wiederum als Kontaktkräfte physikalisch interpretierbar sind.

³vgl. [45], [69]

Häufig können bestimmte Eigenschaften von komplizierten Problemen oder von deren Modellierungen bereits an wesentlich einfacheren Problemen studiert werden. Diese Probleme heißen Modellprobleme und sind ein wichtiger Bestandteil mathematischer Untersuchungen.

Fehlerkontrollierte adaptive Finite-Elemente-Methoden

Die variationelle Formulierung ermöglicht einen direkten Zugang zu effizienten numerischen Diskretisierungsverfahren zur approximativen Bestimmung der im Kontaktproblem geforderten Größen. Das in diesem Zusammenhang populärste Diskretisierungsverfahren ist die Finite-Elemente-Methode, die sich vor allem durch eine hohe Flexibilität im Umgang mit unterschiedlichen geometrischen Bedingungen und durch ein hohes Maß an problembezogener Variabilität gegenüber anderen Verfahren wie z.B. der Differenzenmethode auszeichnet.

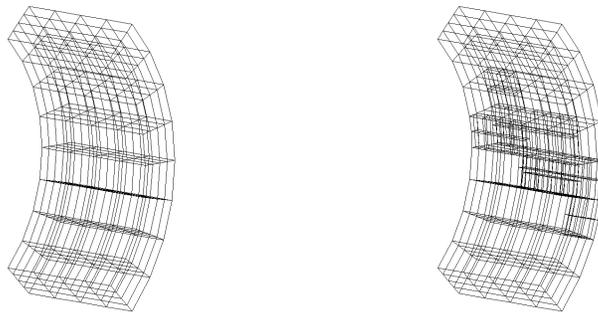


Abb. 2: Finite-Elemente-Gitter von einem Teil des Kontaktscheibenbelags, Gitter mit lokalen Verfeinerungen.

Die Finite-Elemente-Methode basiert auf einer Gitter-artigen Zerlegung des Rechengebiets, das im Wesentlichen durch die in Kontakt befindlichen Körper bestimmt wird, in einfache geometrische Objekte (z.B. Vierecke, Dreiecke, Hexaeder, ...) und auf der Festlegung eines darauf erklärten Diskretisierungsansatzes in Form stückweise definierter Funktionen, deren Einschränkung auf ein solches geometrisches Objekt in einem endlichdimensionalen (lokalen) Polynomraum liegt.

Im Fall elliptischer Minimierungsprobleme ohne Kontaktbedingungen führt

die Finite-Elemente-Methode stets auf lineare Gleichungssysteme, im Kontaktfall auf quadratische Minimierungsprobleme mit linearen oder nichtlinearen Nebenbedingungen bzw. auf Sattelpunktprobleme im \mathbb{R}^m . In jedem Fall sind effiziente Lösungsverfahren für lineare Gleichungssysteme bzw. Optimierungsverfahren anzuwenden.

Zur Steigerung der Approximationsgenauigkeit kann entweder die Maschenweite, definiert als der Durchmesser des größten Gitterelements, verkleinert oder der Polynomansatz erhöht werden. Im ersten Fall spricht man von h -Methoden, im zweiten von p -Methoden. Werden sowohl die Maschenweite verkleinert als auch der Polynomansatz erhöht, erhält man sogenannte hp -Methoden. Für alle drei Zugänge können je nach Problemtyp globale oder aber auch lokale Modifikationen des Diskretisierungsansatzes zu Approximationsverbesserungen führen. Insbesondere sind lokale Änderungen zweckmäßig, wenn hiermit etwa lokale Spitzen des Diskretisierungsfehlers beseitigt werden können.

Die Maschenweite kann zum Beispiel durch Zerlegen der Gitterelemente reduziert werden. Werden alle Gitterelemente zerlegt, spricht man von einer globalen Verfeinerung. Hingegen ist die Verfeinerung nur lokal, wenn nur Gitterelemente bestimmter Bereiche zerlegt werden (vgl. Abbildung 2, rechts). In analoger Weise können die Begriffe lokale und globale Polynomgraderhöhungen definiert werden.

Zentral für die Beurteilung und somit für Einsatzmöglichkeiten von Näherungsverfahren wie die Finite-Elemente-Methode sind Untersuchungen bezüglich der Approximationsgüte, die unter dem Begriff Fehlerkontrolle zusammengefasst werden können.

Die Approximationsgüte der durch die Finite-Elemente-Methode gewonnenen Näherungslösungen ist im besonderen Maße abhängig von der Regularität der Lösung des kontinuierlichen Ausgangsproblems. Die Regularität wird wiederum wesentlich durch die Beschaffenheit des kontinuierlichen Problems, wie etwa durch Geometrieverhältnisse (z.B. einspringende Ecken im Rechengebiet) oder durch Kontaktbedingungen, beeinflusst.

Sind Informationen über die Regularität der kontinuierlichen Lösung bzw. über weitere Problemeigenschaften bekannt, kann für viele Problemklassen die Approximationsgüte bzw. das Konvergenzverhalten in Form von Konvergenzordnungen qualitativ beschrieben werden. Diese werden in der Regel abhängig von der Maschenweite und des verwendeten Polynomgrads angegeben. Zudem können häufig Maßnahmen in Form lokaler Verfeinerungen oder lokaler Polynomgraderhöhungen aufgezeigt werden, die eine Konvergenz verbessernde Wirkung haben.

Eine weitere Form der Fehlerkontrolle, die in der Regel ohne Regularitätsinformationen über die kontinuierliche Lösung auskommt, ist die sogenannte a-posteriori Fehlerkontrolle. Im Unterschied zur oben beschriebenen a-priori Fehlerkontrolle wird der Diskretisierungsfehler durch eine im Prinzip berechenbare obere Schranke, durch einen sogenannten Fehlerschätzer, abgeschätzt. Insbesondere geht in die Berechnung dieser Schranke die Näherungslösung ein. Ein wesentlicher Vorteil dieses Ansatzes ist, dass der Diskretisierungsfehler tatsächlich quantitativ bestimmt werden kann. Sind zudem noch Rückschlüsse auf lokale Fehlerverteilungen möglich, kann die a-posteriori Fehlerkontrolle für Konvergenz-verbessernde Maßnahmen genutzt werden. Die hierbei angewandte Strategie besteht darin, in Bereichen hoher lokaler Fehleranteile entweder die Maschenweite durch lokale Verfeinerungen zu verkleinern oder den lokalen Polynomgrad zu erhöhen. Auf dem so gewonnenen neuen Diskretisierungsansatz kann anschließend wiederum mittels Fehlerkontrolle die lokale Fehlerverteilung geschätzt werden, und erneut lokale Verfeinerungen oder Polynomgraderhöhungen vorgenommen werden. Verfahren, die Adaptierungsstrategien diesen Typs beinhalten, werden üblicherweise h - oder hp -adaptive Verfahren genannt. Diese sind wichtige Bestandteile moderner Finite-Elemente-Verfahren.

Problemstellung und Ziele

Das Ziel der vorliegenden Arbeit besteht darin, Modellierungen von Kontaktproblemen anzugeben, diese in den Kontext der Minimierungs- und Variationsprobleme einzuordnen, einen Überblick bezüglich Finite-Elemente-Diskretisierungen insbesondere vom Typ der h -, p - und hp -Methode inklusive der zugehörigen a-priori Fehleranalyse zu geben und schließlich Möglichkeiten zur a-posteriori Fehlerkontrolle und ihre Anwendbarkeit im Bereich der h - und hp -Adaptivität zu beschreiben. Die Untersuchungen konzentrieren sich hierbei im Wesentlichen auf Probleme für reibungsfreien sowie reibungsbehafteten einseitigen Kontakt, deren Modellierungen in Form sogenannter Signorini-Probleme mit und ohne Reibungsnebenbedingungen darstellbar sind und auf der linearen Elastizitätstheorie beruhen. Dieser Problemtyp entspricht den Anforderungen des oben beschriebenen Kontaktproblems bei der Bahnplanung des Industrieroboter-gestützten Bandschleifens. Dieses spezielle fertigungstechnische Kontaktproblem dient deshalb exemplarisch als Anwendungsbereich für die gewonnenen Resultate. Viele Phänomene können jedoch schon an einfachen Modellproblemen studiert werden.

Deshalb werden den Betrachtungen bezüglich des linear-elastischen Kontaktproblems Untersuchungen an vereinfachten Signorini-Problemen bzw. idealisierten Reibungsproblemen vorangestellt. Darüber hinaus wird die Übertragung der Resultate auch auf andere Modellprobleme wie dem Hindernisproblem, dem Torsionsproblem oder dem Bingham-Fluid-Problem diskutiert. Ein wichtiger Untersuchungsschwerpunkt ist stets die Kontrolle der theoretischen Resultate mit Hilfe numerischer Experimente.

Der zentrale Ansatzpunkt ist die Modellierung der hier untersuchten Kontaktprobleme als konvexe Minimierungsprobleme und insbesondere als elliptische Minimierungsprobleme erster und zweiter Art. Untersuchungen bezüglich der Existenz und Eindeutigkeit von Lösungen von Minimierungsproblemen dieses Typs werden typischerweise der Konvexen Analysis zugeordnet. In Kapitel I werden aus diesem Grund die wesentlichen Verallgemeinerungen von Kontaktproblemen aus Sicht der Konvexen Analysis formuliert und ein kurzer Überblick über die im Zusammenhang von Kontaktproblemen entscheidenden existenztheoretischen Aussagen gegeben. Hieraus werden insbesondere die variationellen Formulierungen der Kontaktprobleme als Variationsgleichungen bzw. -ungleichungen und als variationelle Sattelpunktprobleme gewonnen. Abschließend wird der funktionalanalytische Zusammenhang dieser Formulierungen zu Modellierungen über Kräftebilanzgleichungen mit Hilfe einer verallgemeinerten Greenschen Formel hergestellt. Als Ergebnis gewinnt man Systeme von Operatorgleichungen, die im Wesentlichen Systemen von partiellen Differential(un-)gleichungen entsprechen.

Kapitel II enthält die in dieser Arbeit zu untersuchenden linear-elastischen Modellierungen von reibungsfreien sowie reibungsbehafteten Kontaktproblemen, wobei der reibungsbehaftete Fall auf dem Coulombsche Reibungsgesetz basiert. Insbesondere wird zwischen reibungsfreien Kontaktproblemen, Kontaktproblemen mit vorgegebener Normalspannung und reibungsbehafteten Kontaktproblemen unterschieden. Im ersten Fall werden nur die geometrisch bedingten Nebenbedingungen berücksichtigt, das Kontaktproblem kann als elliptisches Minimierungsproblem erster Art modelliert werden. Im zweiten Fall werden Reibungsbedingungen, im dritten Fall sowohl geometrische Kontaktbedingungen als auch Reibungsbedingungen eingearbeitet. In diesen Fällen wird die Modellierung über ein elliptisches Minimierungsproblem zweiter Art angegeben.

Zudem werden in diesem Kapitel die dazugehörigen Modellprobleme vom Signorini-Typ und des idealisierten Reibungsproblems sowie weitere ver-

wandte Modellprobleme aufgeführt.

Das Ziel dieses Kapitels besteht vornehmlich in der Anwendung der in Kapitel I erzielten Resultate in Bezug auf Existenz und Eindeutigkeit von Lösungen und auf die Darstellbarkeit in Form variationeller Formulierungen. Insbesondere können die in Kapitel I formulierten Operatorgleichungen zu den in der ingenieurwissenschaftlichen Literatur bekannten Systemen von partiellen Differentialgleichungen konkretisiert werden. Sowohl für die linear-elastischen Modellierungen als auch für Modellprobleme werden die so gewonnenen Differentialgleichungssysteme für physikalisch/mechanische Interpretationen genutzt, um die Angemessenheit der verwendeten Modelle zu prüfen.

Ferner wird das einleitend dargestellte Problem zur Erfassung des Kontaktbereichs und der Kontaktkräfte beim Industrieroboter-gestützten Bandschleifen als fertigungstechnische Anwendung von Kontaktproblemen beschrieben. Darüber hinaus wird eine dreidimensionale Beispielkonfiguration hiervon vorgestellt, die in den weiteren Teilen dieser Arbeit exemplarisch verwendet wird. Ebenso werden auch für die angegebenen Modellprobleme zweidimensionale Beispielkonfigurationen formuliert.

Diskretisierungsverfahren und speziell Finite-Elemente-Methoden vom h -, p - oder hp -Typ werden in Kapitel III vorgestellt. Ferner werden die wesentlichen a-priori Resultate aus der Literatur zusammengestellt. Für Kontaktprobleme sind a-priori Konvergenzaussagen im Allgemeinen nur für h -Methoden mit Polynomgrad $p = 1$ oder $p = 2$ bekannt. Für Variationsgleichungen, die kontaktfreien Problemen zugeordnet sind, bestehen hingegen auch Aussagen für höhere Polynomgrade, die an dieser Stelle ebenfalls vorgestellt werden. Zudem existieren a-priori Adaptierungsmaßnahmen für spezielle Problemklassen, die hauptsächlich durch das Vorhandensein von Eckensingularitäten charakterisiert sind. Während die Konvergenzordnung bei Problemen mit Eckensingularitäten im Wesentlichen durch r^α -Singularitäten bestimmt wird, erhält man bei Anwendung von a-priori Adaptierungstechniken wieder die volle Konvergenzordnung $\mathcal{O}(h^p)$ (h -Methode) bzw. exponentielle Konvergenz (hp -Methode). Inwieweit derartige a-priori Aussagen auch auf Kontaktprobleme übertragbar sind, ist bislang ungeklärt. Die experimentell gewonnenen Resultate auf der Grundlage von a-posteriori Adaptierungsstrategien des nachfolgenden Kapitels lassen jedoch auf eine gewisse Übertragbarkeit schließen.⁴

⁴In diesem Fall sind allerdings a-priori Informationen über den Kontaktbereich notwendig.

Ein Problem bei Diskretisierungen von Sattelpunktproblemen ist die Gewährleistung der Stabilität des Verfahrens, womit im Wesentlichen die eindeutige Existenz von diskreten Lagrangeschen Multiplikatoren, in einem gewissen Sinne unabhängig von der Maschenweite bzw. dem gewählten Polynomgrad gemeint ist. Möglichkeiten zur Gewinnung stabiler Diskretisierungsverfahren werden in diesem Kapitel ebenfalls diskutiert und anhand von numerischen Experimenten überprüft.

Für die Implementierung von Finite-Elemente-Methoden, die Polynomansätze höherer Ordnung unterstützen, ist die Angabe von Basen auf Referenzelementen, sogenannte lokale Basen, entscheidend. Mit Hilfe von geeigneten Transformationen auf die Elemente des Finite-Elemente-Gitters werden hiermit die Basisfunktionen des gewählten Ansatzraumes konstruiert. Üblicherweise werden zwei Typen von lokalen Basen eingesetzt: Lagrangeartige Basen und hierarchische Basen, die sich aus aufintegrierten Legendre-Polynomen zusammensetzen. Beide Basistypen weisen spezifische Vor- und Nachteile auf, die bei Finite-Elemente-Implementierungen gegeneinander abzuwägen sind. Hierarchische Basen zeichnen sich vor allem durch eine stabile Gewinnung von Funktions- und Ableitungswerten aus. Außerdem können diese auf einfache Art für Finite-Elemente-Diskretisierungen mit unterschiedlichen Polynomgraden eingesetzt werden. In der Literatur existieren zum Teil sehr unterschiedliche Formulierungen von hierarchischen Basen. In dieser Arbeit wird eine konkrete und für den Einsatz in Finite-Elemente-Programmen verwertbare Darstellung von hierarchischen Basen beliebiger Ordnung für den zwei- und dreidimensionalen Fall mit Hilfe von Gegenbauerpolynomen beschrieben. Insbesondere werden programmiertechnische Probleme in Bezug auf die Basisnummerierung und die Einarbeitung von Stetigkeitsanforderungen auch über sogenannte irreguläre Kanten und Facetten hinweg mit Hilfe von Kopplungsgewichten und speziellen, rekursiv definierten Datenstrukturen gelöst.

Zum Abschluss dieses Kapitels werden Lösungsalgorithmen für quadratische Minimierungsprobleme diskutiert. Hierzu wird insbesondere das in der Finite-Elemente-Literatur weit verbreitete SOR-Verfahren mit Projektion vorgestellt und auf quadratische Minimierungsprobleme, die von linear-elastischen Signorini-Problemen herrühren, mit Hilfe einer Variablentransformation erweitert. Zudem werden auch Lösungsverfahren für Sattelpunktprobleme dargestellt. Es handelt sich hierbei um die standardmäßig eingesetzten Verfahren vom Uzawa- oder Arrow-Hurwicz-Typ. Darüber hinaus wird ein alternativer Ansatz vorgestellt, der die Einbindung von Optimierungsverfahren aus dem Bereich des Quadratic-Programmings bzw. der

nichtlinearen Optimierung ermöglicht. Hiermit können auch kommerzielle Optimierungsprogramme vorteilhaft genutzt werden. Auch in diesem Zusammenhang werden numerische Vergleichsstudien durchgeführt.

Kapitel IV beinhaltet einen neuen Ansatz zur a-posteriori Fehlerkontrolle von Kontaktproblemen. Dieser verallgemeinert eine erst vor kurzem formulierte Idee von Braess [26] zur a-posteriori Fehlerkontrolle von Hindernisproblemen auf allgemeine elliptische Minimierungsprobleme erster und zweiter Art. Wesentlicher Ausgangspunkt ist die Formulierung des Modells in Form einer variationellen Sattelpunktformulierung. Der Ansatz beruht auf der Erkenntnis, dass der diskreten Verschiebungslösung des Sattelpunktproblems die kontinuierliche Lösung eines Hilfsproblems zugeordnet werden kann. Das Hilfsproblem ist eine Variationsgleichung, deren Diskretisierung die gleiche Lösung wie das diskretisierte Ausgangsproblem hat, wodurch der Fehler zwischen diskreter Verschiebungslösung und kontinuierlicher Lösung des Hilfsproblems mit aus der Literatur bekannten Fehlerschätzern für Variationsgleichungen abgeschätzt werden kann. Dieser Zugang bietet den Vorteil, dass Fehlerabschätzungen in einem relativ allgemeinen, funktionalanalytischen Zusammenhang durchgeführt werden können. Die Fehlerkontrolle wird erst durch die Verwendung eines bestimmten Fehlerschätzers, z.B. eines residualen Fehlerschätzers oder eines ZZ-Fehlerschätzers, konkretisiert. Hiermit können Fehlerschätzer für reibungsfreie Kontaktprobleme, Kontaktprobleme mit vorgegebener Normalspannung und reibungsbehaftete Kontaktprobleme sowie für die zugehörigen Modellprobleme angegeben werden. Außerdem können Fehlerschätzer für Hindernisprobleme und Bingham-Fluid-Probleme formuliert werden. Numerische Untersuchungen bestätigen die richtige Asymptotik der Fehlerschätzungen bei globalen Verfeinerungen. Im Hinblick auf die Anwendung höherer Polynomgrade eignen sich vor allem residuale Fehlerschätzer für die Konkretisierung der für Kontaktprobleme formulierten Methoden zur Fehlerkontrolle. Hier kann insbesondere auf neuere Ergebnisse von Melenk [80] zurückgegriffen werden, in denen residuale Fehlerschätzer für hp -Methoden zumindest für den zweidimensionalen Fall hergeleitet werden. In den residualen Fehlerschätzern, die ursprünglich auf Babuška und Rheinboldt [14] zurückgehen, werden p -Abhängigkeiten noch nicht berücksichtigt.

Die Qualität des verwendeten residualen Fehlerschätzers wird mit Hilfe von numerischen Experimenten hinsichtlich des sogenannten Effektivitätsindex geprüft und die Verwendbarkeit bestätigt.

Die so gewonnenen Fehlerschätzer messen den Diskretisierungsfehler in der H^1 -Norm bzw. in der dazu äquivalenten Energie-Norm. Es handelt sich hier-

bei um gemittelte Fehlermaße. Für viele Ingenieur-technische Anwendungen sind jedoch auch lokale Fehlermaße wie der Fehler in einem bestimmten Punkt oder der gemittelte Fehler entlang einer Linie von Interesse. Bei dem oben dargestellten fertigungstechnischen Kontaktproblem kann zum Beispiel die Fehlerkontrolle von Integralmittelwerten von Normalspannungen im Kontaktbereich nutzbringend sein.

Neuere Techniken erlauben, Fehlerschätzer für die H^1 -Norm auch zur Kontrolle des Diskretisierungsfehlers, der bezüglich eines beliebigen Fehlermaßes gemessen werden soll, einzusetzen. Fehlerschätzer dieser Form sind vom sogenannten Goal-Oriented-Typ. Der wesentliche Ansatz besteht in der Darstellung des gewählten Fehlermaßes durch die Lösung eines dualen Problems im Sinne des Rieszschen Darstellungssatzes. In diesem Kapitel werden die wesentlichen Aspekte von Fehlerkontrolltechniken dieses Typs sowohl für Variationsgleichungen als auch für Variationsungleichungen skizziert. Hierbei werden insbesondere Vorschläge von Suttmeier [96] aufgegriffen, den Diskretisierungsfehler bezüglich eines a-posteriori korrigierten Fehlermaßes zu messen.

Ein wichtiger Anwendungsbereich von a-posteriori Fehlerschätzern ist die Erfassung von Fehlerverteilungen im Rechengebiet, mit denen insbesondere Strategien zur Gitteradaptierung entwickelt werden können. Die Formulierung von h - und hp -Adaptierungsstrategien sowie umfangreiche numerische Tests bilden den Untersuchungsschwerpunkt des abschließenden Teils dieses Kapitels.

Adaptierungsstrategien sind in zahlreichen Varianten in der Literatur zu finden. Alle Strategien beruhen auf dem mehr oder weniger heuristischen Ansatz, dass lokale Verfeinerungen, ob vom h - oder p -Typ, in Bereichen mit großen Fehleranteilen sinnvoll sind. Während h -Adaptierungsstrategien relativ leicht zu formulieren sind und in vielen Fällen fast optimale Gitter liefern, sind hp -Adaptierungsstrategien wesentlich komplizierter, da hier zusätzlich über die Art der lokalen Verfeinerung (h - oder p -Verfeinerung) entschieden werden muss. Zentral sind in diesem Zusammenhang vor allem Schätzungen über die lokale Regularität der Lösung.

Die hier vorgestellte hp -Adaptierungsstrategie beinhaltet eine im Vergleich zu bisherigen Strategien relativ vorsichtigen Mechanismus zur Schätzung der lokalen Regularität. Es handelt sich hierbei um eine Verfeinerung der von Süli, u.a. [93] vorgeschlagenen Strategie. Numerische Tests anhand des für Variationsgleichungen wohlbekanntes L-Gebiet-Problems belegen, dass die durch h - und hp -Adaptierungsmaßnahmen gewonnenen adaptiven Gitter stets zu optimalen, in Kapitel II bereits beschriebenen Konvergenzresulta-

ten führen.

Im Hinblick auf hp -Adaptierungen weiß man, dass symmetrische Teilungen der Gitterelemente nicht optimal in Bezug auf das Konvergenzverhalten sind. Stattdessen liefern erst Teilungen mit einem Grading-Factor von etwa 0.15 nahezu optimale Gitter. Die hier vorgestellte hp -Adaptierungsstrategie erreicht jedoch durch eine spezielle Polynomgradverteilung ebenfalls fast optimale Gitter, trotz symmetrischer Teilungen.

Die entwickelten h - und hp -Adaptierungsstrategien werden auch im Zusammenhang mit Fehlerschätzern vom Goal-Oriented-Typ getestet. Hierbei zeigt sich, dass sowohl für getrennt durchgeführte Diskretisierungen für das primale und duale Problem als auch für gleiche Diskretisierungen nachvollziehbare Gitter und Polynomgradverteilungen entstehen.

Hauptaugenmerk dieses Kapitelabschnitts gilt jedoch dem Einsatz der vorgestellten Fehlerkontrolltechniken und der hiermit verbundenen h - und hp -Adaptierungsstrategien im Zusammenhang mit Kontaktproblemen. Es zeigt sich, dass lokale Verfeinerungen vor allem im Bereich der Kontaktzone und insbesondere im Übergangsbereich von Kontaktzone und kontaktfreien Gebieten auftreten. Die Polynomgrade sind in der Regel in der Nähe des Kontaktbereichs klein und steigen außerhalb des Kontaktbereichs an. Damit weisen sowohl h - als auch hp -adaptive Gitterstrukturen, die bei Kontaktproblemen entstehen, große Ähnlichkeiten zu den a-priori adaptierten Gittern für Variationsgleichungen auf.

Numerische Konvergenzstudien belegen, dass durch a-posteriori Adaptierungsmaßnahmen wesentliche Konvergenzverbesserungen bewirkt werden können. Bei der Anwendung von h -adaptiven Strategien wird in den meisten Fällen eine Konvergenzordnung von $\mathcal{O}(h^p)$ erreicht. Bei Anwendung hp -adaptiver Strategien erhält man exponentielle Konvergenz.

Durch die Tendenz der h - und hp -Adaptierungsstrategien, vor allem den Kontaktbereich durch lokale Verfeinerungen zu erfassen, erweist sich der Gebrauch von adaptiven Finite-Elemente-Methoden für Anwendungen aus dem Bereich der Fertigungstechnik und speziell beim Industrieroboter-gestützten Bandschleifen als gewinnbringend. Sowohl Kontaktzone als auch Kontaktkräfte können hiermit wesentlich genauer bestimmt werden als mit herkömmlichen nicht-adaptiven Verfahren.

Anmerkung

Die vorliegende Arbeit ist im Rahmen der Forschergruppe 366 „Simulationsgestützte Offline-Prozessplanung und -optimierung bei der Fertigung von Freiformflächen“ der Deutschen Forschungsgemeinschaft (DFG) entstanden. Besonderer Dank gilt Herrn Prof. Dr. Blum für die wissenschaftliche Unterstützung und den Mitarbeitern des Lehrstuhls „Wissenschaftliches Rechnen“ des Fachbereichs Mathematik der Universität Dortmund sowie Frau Ursula Meyring für das Korrekturlesen der Arbeit.

Für Tuğba und Mia Tuana.

Kapitel I

Elliptische Minimierungsprobleme

Physikalisch- bzw. ingenieurtechnische Modellierungen nutzen in der Regel das sogenannte Minimalprinzip, das voraussetzt, dass physikalische Systeme oder Vorgänge stets den energetisch günstigsten Zustand anstreben, und führen demnach auf die Bestimmung des Minimums

$$\min_{v \in K} E(v)$$

einer definierten Energiegröße, die durch ein (eigentliches¹) Zielfunktional $E : V \rightarrow \overline{\mathbb{R}}$ dargestellt wird. Die Eingangsgröße v ist im Allgemeinen ein Element eines Funktionenraums V und beschreibt Verschiebungs-, Geschwindigkeitsfelder oder Ähnliches. Der energetisch günstigste Zustand wird dann durch eine Lösung $u \in K$ mit

$$E(u) = \min_{v \in K} E(v) \tag{I.1}$$

dargestellt. Die Menge $K \subset V$ entspricht entweder V selbst, also $K := V$, oder K ist eine nichtleere Teilmenge von V , die durch die Angabe von Restriktionen definiert ist, die die modellbedingten Begrenzungen der Eingangsgröße beschreiben. Derartige Begrenzungen sind zum Beispiel maximale Verschiebungen, die durch das Vorhandensein eines Hindernisses bedingt sind. Je nachdem, ob $K := V$ oder $K \subset V$ vorausgesetzt wird, spricht man

¹ $E : V \rightarrow \overline{\mathbb{R}}$ heißt eigentlich, falls $E \not\equiv \infty$ und $E(v) > -\infty$ für alle $v \in V$ ist.

entweder von einem unrestringierten oder restringierten Minimierungsproblem. Die Menge K heißt Restriktionsmenge.

Ziel dieses Kapitels ist, die wesentlichen Eigenschaften von Lösungen von Minimierungsproblemen und insbesondere von elliptischen Minimierungsproblemen erster und zweiter Art zusammenzufassen. Hierzu gehören notwendige wie hinreichende Bedingungen zur Existenz von Lösungen sowie äquivalente Formulierungen in Form von Variationsgleichungen bzw. -ungleichungen und Sattelpunktformulierungen. Hierbei werden im Wesentlichen bekannte Resultate der Konvexen Analysis u.a. aus Cea [34] sowie Ekeland und Temam [49] zusammengetragen.

I.1 Notwendige und hinreichende Bedingungen, Existenz von Lösungen

Notwendige und hinreichende Bedingungen für eine Lösung u des Minimierungsproblems (I.1) erhält man üblicherweise über den Begriff der Fréchet-Ableitung. Ist V ein normierter Raum mit der Norm $\|\cdot\|$, dann heißt E in $w \in V$ Fréchet-differenzierbar, falls ein $E'(w) \in V'$ existiert, so dass

$$\forall v \in V : E(w + v) - E(w) = \langle E'(w), v \rangle + \varrho(v),$$

mit einer Abbildung $\varrho : V \rightarrow \mathbb{R}$ und $|\varrho(v)| = o(\|v\|)$ für $\|v\| \rightarrow 0$ gilt. Die Abbildung $E'(w)$ heißt Fréchet-Ableitung in w . Dabei bezeichnet $\langle \cdot, \cdot \rangle : V' \times V \rightarrow \mathbb{R}$ das übliche Inzidenzprodukt $\langle \lambda, v \rangle := \lambda(v)$ für $(\lambda, v) \in V' \times V$.

Für die Herleitung von notwendigen und hinreichenden Bedingungen an eine Lösung u werden im Folgenden zwei Fälle betrachtet: Minimierungsprobleme mit Fréchet-differenzierbarem Zielfunktional und Minimierungsprobleme, deren Zielfunktional sich aus einem Fréchet-differenzierbaren Anteil und einem im Allgemeinen nicht Fréchet-differenzierbaren, aber konvexen Anteil zusammensetzt.

Die Frage nach der Existenz von Lösungen wird anschließend geklärt.

I.1.1 Notwendige und hinreichende Bedingungen

Ist K konvex und E in u Fréchet-differenzierbar, dann folgt aus

$$0 \leq E(u + \epsilon(v - u)) - E(u) = \epsilon \langle E'(u), v - u \rangle + \varrho(\epsilon(v - u))$$

für jedes $\epsilon \in (0, 1]$ und jedes $v \in K$, dass $0 \leq \langle E'(u), v - u \rangle + \varrho(\epsilon(v - u))/\epsilon$ ist. Für $\epsilon \rightarrow 0$ erhält man daraus die notwendige Bedingung

$$\forall v \in K : \langle E'(u), v - u \rangle \geq 0 \quad (\text{I.2})$$

für die Minimalität von $E(u)$. Die Bedingung (I.2) ist im Allgemeinen nicht hinreichend, es sei denn, man setzt für das Zielfunktional E eine zusätzliche Eigenschaft voraus: Ist E konvex und in $w \in V$ Fréchet-differenzierbar, dann folgt für ein beliebiges $v \in V$ und ein $\epsilon \in (0, 1]$ aus

$$E(w + \epsilon(v - w)) = E(\epsilon v + (1 - \epsilon)w) \leq \epsilon E(v) + (1 - \epsilon)E(w)$$

die Ungleichung

$$E(v) - E(w) \geq \frac{E(w + \epsilon(v - w)) - E(w)}{\epsilon} = \langle E'(w), v - w \rangle + \frac{\varrho(\epsilon(v - w))}{\epsilon},$$

und mit $\epsilon \rightarrow 0$ erhält man schließlich

$$\forall v \in K : E(v) - E(w) \geq \langle E'(w), v - w \rangle. \quad (\text{I.3})$$

Demnach ist die Bedingung (I.2) auch hinreichend dafür, dass u eine Lösung des Minimierungsproblems ist.

Ist E streng konvex, so lässt sich die Bedingung (I.3) für $v \neq w$ zu $E(v) - E(w) > \langle E'(w), v - w \rangle$ für alle $v \in K$ verschärfen, woraus unmittelbar folgt, dass E in diesem Fall nur höchstens eine Lösung besitzen kann.

Wenn K ein affiner Unterraum von V ist, also $K = v^* + U$ mit einem linearen Unterraum U und $v^* \in V$, sind mit der Lösung $u \in K$ des Minimierungsproblems auch $u + v$ und $u - v$ für alle $v \in U$ in K , und es folgt $\langle E'(u), v \rangle \geq 0$ und $\langle E'(u), v \rangle \leq 0$. In diesem Fall ist

$$\forall v \in U : \langle E'(u), v \rangle = 0$$

notwendig, und wenn E konvex ist, auch hinreichend.

Die hergeleiteten notwendigen und hinreichenden Bedingungen führen zu äquivalenten Variationsformulierungen des Minimierungsproblems (I.1), die in dem folgenden Satz zusammengefasst werden:

Satz I.1 *Es sei V ein normierter Raum und $K \subset V$ konvex sowie $u \in K$ und $E : V \rightarrow \overline{\mathbb{R}}$ ein konvexes Funktional, das in u Fréchet-differenzierbar ist. Genau dann ist u Lösung des Minimierungsproblems (I.1), wenn gilt:*

$$\forall v \in K : \langle E'(u), v - u \rangle \geq 0. \quad (\text{I.4})$$

Im Fall, dass $K = v^* + U$ mit einem Unterraum U und $v^* \in V$ ist, ist u genau dann Lösung von (I.1), wenn gilt:

$$\forall v \in U : \langle E'(u), v \rangle = 0. \quad (\text{I.5})$$

Ist E strikt konvex, dann existiert höchstens eine Lösung u .

Bemerkung I.1 Variationsformulierungen der Form (I.4) werden Variationsungleichungen genannt, Variationsformulierungen der Form (I.5) heißen Variationsgleichungen.

Einen Spezialfall bilden Minimierungsprobleme, deren Zielfunktionale die Form

$$E(v) := H(v) + j(v) \quad (\text{I.6})$$

besitzen. Hierbei sind $H : V \rightarrow \overline{\mathbb{R}}$ und $j : V \rightarrow \overline{\mathbb{R}}$ konvexe Funktionale, wobei j im Allgemeinen nicht Fréchet-differenzierbar ist. Zum Beispiel führt die Modellierung von Reibung auf ein Zielfunktional wie in (I.6) mit einem konvexen, aber nicht Fréchet-differenzierbaren Funktional j .

Ist H in der Lösung u des Minimierungsproblems (I.1) mit einem Zielfunktional wie in (I.6) Fréchet-differenzierbar, und ist K konvex, dann ist für jedes $\epsilon \in (0, 1]$ und jedes $v \in K$

$$\begin{aligned} E(u + \epsilon(v - u)) - E(u) &= H(u + \epsilon(v - u)) - H(u) + j(u + \epsilon(v - u)) - j(u) \\ &\leq \epsilon \langle H'(u), v - u \rangle + \varrho(\epsilon(v - u)) + \epsilon(j(v) - j(u)). \end{aligned}$$

Also ist $0 \leq \langle H'(u), v - u \rangle + j(v) - j(u) + \varrho(\epsilon(v - u))/\epsilon$. Mit $\epsilon \rightarrow 0$ folgt daraus die für eine Lösung u notwendige Bedingung

$$\forall v \in K : \langle H'(u), v - u \rangle + j(v) - j(u) \geq 0. \quad (\text{I.7})$$

Für ein konvexes H ist (I.7) auch hinreichend. Denn dann ist E ebenfalls konvex. Analog zu (I.3) gilt für jedes $w \in V$, in dem H Fréchet-differenzierbar ist, und für jedes $v \in V$

$$\begin{aligned} E(v) - E(w) &\geq \frac{E(w + \epsilon(v - w)) - E(w)}{\epsilon} \\ &\geq \langle H'(w), v - w \rangle + j(v) - j(w) + \frac{\varrho(\epsilon(v - w))}{\epsilon}. \end{aligned}$$

Hieraus folgt schließlich für $\epsilon \rightarrow 0$

$$E(v) - E(w) \geq \langle H'(w), v - w \rangle + j(v) - j(w). \quad (\text{I.8})$$

Die notwendigen und hinreichenden Bedingungen zusammen führen damit auf eine Variationsungleichung der folgenden Form:

Satz I.2 *Es sei V ein normierter Raum und $K \subset V$ konvex, sowie $E : V \rightarrow \overline{\mathbb{R}}$ definiert wie in (I.6) mit einem in $u \in K$ Fréchet-differenzierbaren, konvexen Funktional $H : V \rightarrow \overline{\mathbb{R}}$ und mit einem konvexen Funktional $j : V \rightarrow \overline{\mathbb{R}}$. Genau dann ist u Lösung des Minimierungsproblems (I.1), wenn gilt:*

$$\forall v \in K : \langle H'(u), v - u \rangle + j(v) - j(u) \geq 0.$$

Ist H strikt konvex, dann existiert höchstens eine Lösung u .

I.1.2 Existenz von Lösungen

Für den Nachweis der Existenz einer Lösung $u \in K$ des Minimierungsproblems (I.1) sind neben der Konvexität weitere Eigenschaften von E bzw. K erforderlich.

Eine wichtige Eigenschaft von Funktionalen ist in diesem Zusammenhang die Unterhalb-Stetigkeit. Ein Funktional $E : V \rightarrow \overline{\mathbb{R}}$ heißt unterhalb stetig, falls

$$\forall v \in V : E(v) \leq \liminf_{\bar{v} \rightarrow v} E(\bar{v}).$$

Ist ein Funktional E konvex und unterhalb stetig, dann ist der Epigraph $\text{epi}(E)$ konvex und abgeschlossen und damit nach Satz A.5 des Anhangs auch schwach abgeschlossen. Also ist E auch schwach unterhalb stetig². Ist V reflexiv und K abgeschlossen, folgt aus Satz A.3, dass eine beschränkte Folge $\{u_k\} \subset K$ eine schwach konvergente Teilfolge $\{u_{k_i}\} \subset K$ mit $u_{k_i} \rightharpoonup u \in K$ besitzt. Ist $\{u_k\}$ eine Minimalfolge³ bzgl. E , folgt schließlich $E(u) \leq \lim_{i \rightarrow \infty} E(u_{k_i}) = \inf_{v \in K} E(v)$, und man erhält:

Lemma I.1 *Es sei V ein reflexiver Raum und $K \subset V$ abgeschlossen und konvex. Das Funktional E sei konvex und unterhalb stetig. Sind die Minimalfolgen des Minimierungsproblems (I.1) beschränkt, dann existiert mindestens eine Lösung $u \in K$.*

Ist $K \subset V$ konvex, abgeschlossen und beschränkt, dann folgt sofort die Existenz einer Lösung $u \in K$ aus Lemma I.1. Im Fall, dass K unbeschränkt ist,

² $E : V \rightarrow \overline{\mathbb{R}}$ heißt schwach unterhalb stetig, falls $\forall v \in V : E(v) \leq \liminf_{\bar{v} \rightarrow v} E(\bar{v})$ gilt.

³ $\{u_k\} \subset K \subset V$ heißt Minimalfolge (bzw. Maximalfolge) bzgl. $E : V \rightarrow \overline{\mathbb{R}}$, falls $E(u_k) \rightarrow \inf_{v \in K} E(v)$ (bzw. $E(u_k) \rightarrow \sup_{v \in K} E(v)$).

sichert die Koerzivität eines Funktionals die Beschränktheit einer Minimalfolge und damit gemäß Lemma I.1 die Existenz einer Lösung. Ein Funktional $E : V \rightarrow \overline{\mathbb{R}}$ heißt koerzitiv, falls

$$\lim_{\|v\| \rightarrow \infty} E(v) = \infty$$

ist. Es sei $\{v_k\}$ eine Folge mit $\lim_{k \rightarrow \infty} E(v_k) < \infty$. Angenommen $\{v_k\}$ wäre unbeschränkt, dann folgte direkt aus der Koerzivität $\lim_{k \rightarrow \infty} E(v_k) = \infty$. Also muss $\{v_k\}$ beschränkt sein. Insbesondere sind damit Minimalfolgen beschränkt. Zusammenfassend erhält man damit:

Satz I.3 *Es sei V ein reflexiver Raum und $K \subset V$ abgeschlossen und konvex. Das Funktional E sei konvex und unterhalb stetig. Ist zudem K beschränkt oder E koerzitiv, dann existiert mindestens eine Lösung $u \in K$ des Minimierungsproblems (I.1).*

Die in Satz I.3 genannten Voraussetzungen können in der folgenden Form kombiniert werden:

Satz I.4 *Für $i = 0, 1$ seien V_i reflexive Räume sowie $K_i \subset V_i$ abgeschlossen und konvex, wobei K_1 beschränkt ist. Die Funktionale $E_i : V_i \rightarrow \overline{\mathbb{R}}$ seien konvex und unterhalb stetig, wobei E_0 koerzitiv ist. Das Funktional $E : V \rightarrow \overline{\mathbb{R}}$ mit $V := V_0 \times V_1$ sei konvex und unterhalb stetig, und es gelte:*

$$\forall (v_0, v_1) \in V_0 \times V_1 : E(v_0, v_1) \geq E_0(v_0) + E_1(v_1). \quad (\text{I.9})$$

Dann existiert mindestens eine Lösung $u \in K := K_0 \times K_1$ des Minimierungsproblems (I.1).

Beweis: Da K_1 beschränkt ist, existiert nach Satz I.3 ein $u_1 \in K_1$ mit $E_1(u_1) = \inf_{v_1 \in V_1} E_1(v_1)$. Es sei $\{(v_{0,k}, v_{1,k})\} \subset K_0 \times K_1$ eine Minimalfolge. Angenommen $\{(v_{0,k}, v_{1,k})\}$ ist unbeschränkt, dann muss, da K_1 beschränkt ist, $\{v_{0,k}\}$ unbeschränkt sein. Damit folgt aber aus der Koerzivität von E_0

$$\lim_{k \rightarrow \infty} E(v_{0,k}, v_{1,k}) \geq \lim_{k \rightarrow \infty} E_0(v_{0,k}) + E_1(v_{1,k}) \geq \lim_{k \rightarrow \infty} E_0(v_{0,k}) + E(u_1) = \infty.$$

Also ist $\{(v_{0,k}, v_{1,k})\}$ beschränkt, womit aus Lemma I.1 die Existenz einer Lösung $u \in K$ des Minimierungsproblems folgt. \square

I.2 Sattelpunktformulierungen

Für theoretische Betrachtungen und für die Herleitung numerischer Algorithmen zur Lösung des Minimierungsproblems (I.1) ist es häufig vorteilhaft,

statt des Minimierungsproblems ein entsprechendes Sattelpunktproblem zu untersuchen.

Das Paar $(u, \lambda) \in V \times \Lambda$ heißt Sattelpunkt des Funktionals $\mathcal{L} : V \times \Lambda \rightarrow \overline{\mathbb{R}}$, falls

$$\forall v \in V, \forall \mu \in \Lambda : \mathcal{L}(u, \mu) \leq \mathcal{L}(u, \lambda) \leq \mathcal{L}(v, \lambda) \quad (\text{I.10})$$

ist. Für λ ist im Zusammenhang von Minimierungsproblemen die Bezeichnung Lagrangescher Multiplikator gebräuchlich, \mathcal{L} wird Lagrange-Funktional genannt.

Ziel ist die Herleitung von Sattelpunktformulierungen, die hinreichend für die Lösung des Minimierungsproblems (I.1) sind. Diskutiert wird insbesondere, in welcher Weise die Restriktionsmenge K und der in Zielfunktionalen der Form (I.6) enthaltene Anteil j mit Hilfe Lagrangescher Multiplikatoren erfasst werden können.

Zuvor wird eine zum Sattelpunktproblem (I.10) äquivalente variationelle Sattelpunktformulierung angegeben und die Existenz von Sattelpunkten geklärt.

I.2.1 Variationelle Formulierung, Existenz von Sattelpunkten

Zur Herleitung einer zu (I.10) äquivalenten variationellen Sattelpunktformulierung lassen sich die in Abschnitt I.1 aufgeführten Aussagen direkt anwenden. Durch Umformulieren von (I.10) in Minimierungsprobleme folgt aus Satz I.1:

Satz I.5 *Es seien V und U' normierte Räume und $\emptyset \neq \Lambda \subset U'$ konvex. Für $(u, \lambda) \in V \times \Lambda$ und das Funktional $\mathcal{L} : V \times \Lambda \rightarrow \overline{\mathbb{R}}$ gelte:*

- (i) $\mathcal{L}_u := \mathcal{L}(u, \cdot)$ sei konkav und in λ Fréchet-differenzierbar.
- (ii) $\mathcal{L}_\lambda := \mathcal{L}(\cdot, \lambda)$ sei konvex und in u Fréchet-differenzierbar.

Genau dann ist (u, λ) ein Sattelpunkt von \mathcal{L} , wenn gilt:

$$\begin{aligned} \forall v \in V : \langle \mathcal{L}'_\lambda(u), v \rangle &= 0 \\ \forall \mu \in \Lambda : \langle \mathcal{L}'_u(\lambda), \mu - \lambda \rangle &\leq 0. \end{aligned}$$

Für die Existenz von Sattelpunkten werden im Weiteren Voraussetzungen untersucht, die insbesondere für die in Abschnitt I.3 diskutierten elliptischen

Minimierungsprobleme geeignet sind.
Allgemein gilt, dass aus

$$\max_{\mu \in \Lambda} \inf_{v \in V} \mathcal{L}(v, \mu) = \min_{v \in V} \sup_{\mu \in \Lambda} \mathcal{L}(v, \mu) \quad (\text{I.11})$$

bereits die Existenz eines Sattelpunkts folgt.⁴ Denn für $(u, \lambda) \in V \times \Lambda$ mit

$$\begin{aligned} \inf_{v \in V} \mathcal{L}(v, \lambda) &= \max_{\mu \in \Lambda} \inf_{v \in V} \mathcal{L}(v, \mu) \\ \sup_{\mu \in \Lambda} \mathcal{L}(u, \mu) &= \min_{v \in V} \sup_{\mu \in \Lambda} \mathcal{L}(v, \mu) \end{aligned}$$

erhält man unmittelbar

$$\max_{\mu \in \Lambda} \inf_{v \in V} \mathcal{L}(v, \mu) \leq \mathcal{L}(u, \lambda) \leq \min_{v \in V} \sup_{\mu \in \Lambda} \mathcal{L}(v, \mu). \quad (\text{I.12})$$

Wegen (I.11) besteht Gleichheit in (I.12), also ist (u, λ) ein Sattelpunkt.

Zentral für die folgenden Überlegungen ist die Aussage:

Lemma I.2 *Es seien V und U' reflexive Räume und $\Lambda \subset U'$ abgeschlossen und konvex. Für das Funktional $\mathcal{L} : V \times \Lambda \rightarrow \overline{\mathbb{R}}$ gelte:*

- (i) $\mathcal{L}_v := \mathcal{L}(v, \cdot)$ sei konkav und oberhalb stetig⁵ für alle $v \in V$.
- (ii) $\mathcal{L}_\mu := \mathcal{L}(\cdot, \mu)$ sei konvex und unterhalb stetig für alle $\mu \in \Lambda$.
- (iii) \mathcal{L}_{μ_0} sei koerziv für ein $\mu_0 \in \Lambda$.

Dann gilt

$$\min_{v \in V} \sup_{\mu \in \Lambda} \mathcal{L}(v, \mu) = \sup_{\mu \in \Lambda} \inf_{v \in V} \mathcal{L}(v, \mu).$$

Beweis: Prop.VI.2.3 in [49] □

Hieraus erhält man nun unmittelbar eine Existenzaussage: Wenn Λ als beschränkt oder die Abbildung $\mu \mapsto \sup_{v \in V} (-\mathcal{L}(v, \mu))$ als koerziv vorausgesetzt wird, so folgt aus Satz I.3, dass ein $\lambda \in \Lambda$ mit der Eigenschaft

$$\sup_{v \in V} (-\mathcal{L}(v, \lambda)) = \inf_{\mu \in \Lambda} \sup_{v \in V} (-\mathcal{L}(v, \mu))$$

⁴Andererseits ist leicht einzusehen, dass (I.11) auch notwendig für die Existenz eines Sattelpunkts ist.

⁵d.h., $-\mathcal{L}_v$ ist unterhalb stetig.

existiert. Unter den Voraussetzungen von Lemma I.2 folgt daraus

$$\max_{\mu \in \Lambda} \inf_{v \in V} \mathcal{L}(v, \mu) = \sup_{\mu \in \Lambda} \inf_{v \in V} \mathcal{L}(v, \mu) = \min_{v \in V} \sup_{\mu \in \Lambda} \mathcal{L}(v, \mu).$$

Zusammen mit den Vorüberlegungen ist damit die Existenz eines Sattelpunkts gezeigt. Insgesamt gilt demnach der Satz:

Satz I.6 *Es gelten die Voraussetzungen von Lemma I.2. Zudem sei Λ beschränkt oder $\mu \mapsto \sup_{v \in V} (-\mathcal{L}(v, \mu))$ koerziv. Dann besitzt \mathcal{L} mindestens einen Sattelpunkt.*

Analog erhält man im Spezialfall $\Lambda := \Lambda_0 \times \Lambda_1$ aus Satz I.4:

Satz I.7 *Es seien V sowie U'_0 und U'_1 reflexive Räume und $\Lambda_0 \subset U'_0$ und $\Lambda_1 \subset U'_1$ abgeschlossen und konvex. Mit $\Lambda := \Lambda_0 \times \Lambda_1$ gelten die Voraussetzungen von Lemma I.2. Zudem habe $(\mu_0, \mu_1) \mapsto \sup_{v \in V} (-\mathcal{L}(v, \mu_0, \mu_1))$ die Eigenschaft (I.9). Dann besitzt \mathcal{L} mindestens einen Sattelpunkt.*

I.2.2 Sattelpunktformulierungen für Minimierungsprobleme

Die Umformulierung des Minimierungsproblems (I.1) in ein Sattelpunktproblem wird durch die Angabe eines Multiplikatorfunktionals vollzogen.

Definition I.1 *Es seien V und U'_0 lineare Räume, $\Lambda_0 \subset U'_0$ und $K \subset V$. Das Funktional $\Phi_0 : V \times \Lambda_0 \rightarrow \mathbb{R}$ heißt Multiplikatorfunktional bzgl. K , wenn gilt:*

$$\sup_{\mu_0 \in \Lambda_0} \Phi_0(v, \mu_0) = \begin{cases} 0, & v \in K \\ \infty, & v \notin K. \end{cases}$$

Die Motivation zur Einführung des Multiplikatorfunktionals ergibt sich aus der folgenden Beobachtung: Für ein Multiplikatorfunktional Φ_0 bzgl. K wie in Definition I.1 gilt

$$\inf_{v \in K} E(v) = \inf_{v \in V} \left(E(v) + \sup_{\mu_0 \in \Lambda_0} \Phi_0(v, \mu_0) \right) = \inf_{v \in V} \sup_{\mu_0 \in \Lambda_0} (E(v) + \Phi_0(v, \mu_0)),$$

woraus unmittelbar die Formulierung des Minimierungsproblems (I.1) als Sattelpunktproblem herzuleiten ist:

Satz I.8 *Es seien V und U'_0 lineare Räume, $\Lambda_0 \subset U'_0$ und $K \subset V$. Das Funktional $\Phi_0 : V \times \Lambda_0 \rightarrow \mathbb{R}$ sei ein Multiplikatorfunktional bzgl. K . Dann ist $u \in K$ Lösung des Minimierungsproblems (I.1), wenn ein $\lambda_0 \in \Lambda_0$ existiert, so dass (u, λ_0) Sattelpunkt von $\mathcal{L} : V \times \Lambda_0 \rightarrow \overline{\mathbb{R}}$ mit $\mathcal{L}(v, \mu_0) := E(v) + \Phi_0(v, \mu_0)$ ist.*

Ein einfaches Kriterium für Multiplikatorfunktionale liefert der folgende Satz:

Satz I.9 *Es seien V und U'_0 lineare Räume, $\Lambda_0 \subset U'_0$ ein Kegel⁶ und $K \subset V$. Für das Funktional $\Phi_0 : V \times \Lambda_0 \rightarrow \mathbb{R}$ gelte:*

$$\forall \alpha \geq 0, \forall (v, \mu_0) \in V \times \Lambda_0 : \Phi_0(v, \alpha \mu_0) = \alpha \Phi_0(v, \mu_0) \quad (\text{I.13})$$

$$v \in K \Leftrightarrow \forall \mu_0 \in \Lambda_0 : \Phi_0(v, \mu_0) \leq 0. \quad (\text{I.14})$$

Dann ist Φ_0 ein Multiplikatorfunktional bzgl. K .

Beweis: Da 0 in Λ_0 enthalten ist, ist wegen (I.13) $\Phi_0(v, 0) = 0$ für alle $v \in V$. Aus (I.14) folgt dann $\sup_{\mu_0 \in \Lambda_0} \Phi_0(v, \mu_0) = 0$. Ist $v \notin K$, dann muss es ein $\tilde{\mu}_0 \in \Lambda_0$ geben, so dass $\Phi_0(v, \tilde{\mu}_0) > 0$ ist. Folglich ist $\sup_{\mu_0 \in \Lambda_0} \Phi_0(v, \mu_0) \geq \sup_{\alpha > 0} \Phi_0(v, \alpha \tilde{\mu}_0) = \sup_{\alpha > 0} \alpha \Phi_0(v, \tilde{\mu}_0) = \infty$. \square

Durch Verwenden eines Multiplikatorfunktionals erhält man nun mit Hilfe von Satz I.5 eine variationelle Sattelpunktformulierung, die hinreichend für die Lösung des Minimierungsproblems (I.1) ist.

Korollar I.1 *Es seien V und U'_0 normierte Räume sowie $\Lambda_0 \subset U'_0$ und $K \subset V$ konvex. Die Abbildung $\Phi_0 : V \times \Lambda_0 \rightarrow \mathbb{R}$ sei ein Multiplikatorfunktional bzgl. K . Für $(u, \lambda_0) \in V \times \Lambda_0$ und das konvexe Funktional $E : V \rightarrow \overline{\mathbb{R}}$ gelte:*

- (i) E sei in u Fréchet-differenzierbar.
- (ii) $\Phi_{0,u} := \Phi_0(u, \cdot)$ sei konkav und in λ_0 Fréchet-differenzierbar.
- (iii) $\Phi_{0,\lambda_0} := \Phi_0(\cdot, \lambda_0)$ sei konvex und in u Fréchet-differenzierbar.

Dann ist u Lösung des Minimierungsproblems (I.1), wenn gilt:

$$\forall v \in V : \langle E'(u) + \Phi'_{0,\lambda_0}(u), v \rangle = 0 \quad (\text{I.15})$$

$$\forall \mu_0 \in \Lambda_0 : \langle \Phi'_{0,u}(\lambda_0), \mu_0 - \lambda_0 \rangle \leq 0. \quad (\text{I.16})$$

⁶Die Menge Λ heißt Kegel (mit Spitze in 0), falls $0 \in \Lambda$ und $\forall \mu \in \Lambda, \forall \alpha \geq 0 : \alpha \mu \in \Lambda$ ist.

Der wesentliche Vorteil der variationellen Sattelpunktformulierung (I.15)-(I.16) gegenüber der Variationsformulierung (I.4) besteht darin, dass der lineare Raum V , und nicht die konvexe Menge K , sowohl die Menge der zulässigen Funktionen also auch die Menge der Testfunktionen bildet. Erkauft wird dies durch Einführen Lagrangescher Multiplikatoren aus dem konvexen Kegel Λ_0 , der aber im Allgemeinen wesentlich einfachere Restriktionsbedingungen aufweist.

In der Situation von (I.6) mit einem in u Fréchet-differenzierbaren Funktional H und einem konvexen Funktional j erhält man unter den Voraussetzungen von Satz I.1 an das Multiplikatorfunktional Φ_0 gemäß Satz I.5 statt einer Variationsgleichung wie in (I.15) nur die Variationsungleichung

$$\forall v \in V : \langle H'(u) + \Phi'_{0,\lambda_0}(u), v \rangle + j(v) - j(u) \geq 0.$$

Abhilfe verschafft hier die Verwendung eines Begriffs ähnlich wie in Definition I.1.

Definition I.2 *Es seien V und U'_1 lineare Räume, $\Lambda_1 \subset U'_1$ und $j : \Lambda_1 \rightarrow \overline{\mathbb{R}}$. Das Funktional $\Phi_1 : V \times \Lambda_1 \rightarrow \mathbb{R}$ heißt Multiplikatorfunktional bzgl. j , wenn gilt:*

$$\forall v \in V : j(v) = \sup_{\mu_1 \in \Lambda_1} \Phi_1(v, \mu_1). \quad (\text{I.17})$$

Auch hier ist die Motivation offensichtlich. Wenn Φ_0 ein Multiplikatorfunktional bzgl. K ist, und Φ_1 ein Multiplikatorfunktional bzgl. j , dann ist

$$\begin{aligned} \inf_{v \in K} E(v) &= \inf_{v \in V} \sup_{\mu_0 \in \Lambda_0} (E(v) + \Phi_0(v, \mu_0)) \\ &= \inf_{v \in V} \left(\sup_{\mu_0 \in \Lambda_0} (H(v) + \Phi_0(v, \mu_0)) + \sup_{\mu_1 \in \Lambda_1} \Phi_1(v, \mu_1) \right) \\ &= \inf_{v \in V} \sup_{\mu_0 \in \Lambda_0, \mu_1 \in \Lambda_1} (H(v) + \Phi_0(v, \mu_0) + \Phi_1(v, \mu_1)). \end{aligned}$$

Demnach erhält man die folgende Formulierung des Minimierungsproblems (I.1) als Sattelpunktproblem:

Satz I.10 *Es seien V , U'_0 und U'_1 lineare Räume sowie $\Lambda_0 \subset U'_0$ und $\Lambda_1 \subset U'_1$. Zudem seien $K \subset V$ und $j : V \rightarrow \overline{\mathbb{R}}$. Die Funktionale $\Phi_0 : V \times \Lambda_0 \rightarrow \mathbb{R}$ und $\Phi_1 : V \times \Lambda_1 \rightarrow \mathbb{R}$ seien Multiplikatorfunktionale bzgl. K bzw. j . Ferner sei $E := H + j$ mit einem Funktional $H : V \rightarrow \overline{\mathbb{R}}$. Dann ist $u \in K$ Lösung des Minimierungsproblems (I.1), wenn ein Paar $(\lambda_0, \lambda_1) \in \Lambda_0 \times \Lambda_1$ existiert, so dass $(u, \lambda_0, \lambda_1)$ Sattelpunkt von $\mathcal{L} : V \times \Lambda_0 \times \Lambda_1 \rightarrow \overline{\mathbb{R}}$ mit $\mathcal{L}(v, \mu_0, \mu_1) := H(v) + \Phi_0(v, \mu_0) + \Phi_1(v, \mu_1)$ ist.*

Aus Satz I.10 folgt schließlich eine variationelle Sattelpunktformulierung, die dem System (I.15)-(I.16) entspricht, also insbesondere aus einer Variationsgleichung und einer Variationsungleichung besteht:

Korollar I.2 *Es seien V , U'_0 und U'_1 normierte Räume sowie $\Lambda_0 \subset U'_0$, $\Lambda_1 \subset U'_1$ und $K \subset V$ konvex. Zudem sei $j : V \rightarrow \overline{\mathbb{R}}$ konvex. Die Funktionale $\Phi_0 : V \times \Lambda_0 \rightarrow \mathbb{R}$ und $\Phi_1 : V \times \Lambda_1 \rightarrow \mathbb{R}$ seien Multiplikatorfunktionale bzgl. K bzw. j . Ferner sei $E := H + j$ mit einem konvexen Funktional $H : V \rightarrow \overline{\mathbb{R}}$. Für $(u, \lambda_0, \lambda_1) \in V \times \Lambda_0 \times \Lambda_1$ gelte:*

(i) H sei in u Fréchet-differenzierbar.

(ii) $\Phi_{0,u} := \Phi(u, \cdot)$ und $\Phi_{1,u} := \Phi_1(u, \cdot)$ seien konkav und in λ_0 bzw. λ_1 Fréchet-differenzierbar.

(iii) $\Phi_{0,\lambda_0} := \Phi_0(\cdot, \lambda_0)$ und $\Phi_{1,\lambda_1} := \Phi_1(\lambda_1, \cdot)$ seien konvex und in u Fréchet-differenzierbar.

Dann ist u Lösung des Minimierungsproblems (I.1), wenn gilt:

$$\forall v \in V : \langle H'(u) + \Phi'_{0,\lambda_0}(u) + \Phi'_{1,\lambda_1}(u), v \rangle = 0 \quad (\text{I.18})$$

$$\forall (\mu_0, \mu_1) \in \Lambda_0 \times \Lambda_1 : \langle \Phi'_{0,u}(\lambda_0), \mu_0 - \lambda_0 \rangle + \langle \Phi'_{1,u}(\lambda_1), \mu_1 - \lambda_1 \rangle \leq 0. \quad (\text{I.19})$$

Bemerkung I.2 Nach Satz I.8 und Satz I.10 kann die Frage nach der Existenz einer Lösung des Minimierungsproblems (I.1) auch über die Frage nach der Existenz von Sattelpunkten beantwortet werden. Erfüllen die in den Sätzen (I.8) bzw. (I.10) definierten Lagrange-Funktionale die in Satz I.6 vorgegebenen Voraussetzungen, dann existiert ein Sattelpunkt (u, λ_0) bzw. $(u, \lambda_0, \lambda_1)$, wobei u dann gleichzeitig Lösung des Minimierungsproblems (I.1) ist.

I.3 Elliptische Minimierungsprobleme

Bei der Modellierung Ingenieur-technischer Anwendungen spielen elliptische Minimierungsprobleme eine zentrale Rolle. So führt die Betrachtung des Deformationsverhaltens von Körpern auf der Grundlage der linearen Elastizitätstheorie auf eine Modellierung in Form eines elliptischen Minimierungsproblems.

Elliptische Minimierungsprobleme sind durch die Angabe einer symmetrischen, stetigen und elliptischen Bilinearform $a : V \times V \rightarrow \mathbb{R}$ charakterisiert. Dabei heißt die Bilinearform a auf einem normierten Raum stetig, wenn gilt:

$$\exists \nu_0 \in \mathbb{R}_{\geq 0} : \forall v, w \in V : a(v, w) \leq \nu_0 \|v\| \|w\|,$$

und elliptisch, wenn gilt:

$$\exists \nu_1 \in \mathbb{R}_{> 0} : \forall v \in V : a(v, v) \geq \nu_1 \|v\|^2.$$

In diesem Abschnitt werden zwei Typen von elliptischen Minimierungsproblemen vorgestellt: Elliptische Minimierungsprobleme erster und zweiter Art. Diese werden bezüglich ihrer Zugehörigkeit zu den bisher genannten Problemklassen und bezüglich ihrer Eigenschaften - notwendige und hinreichende Bedingungen sowie Existenz von Lösungen, variationelle Sattelpunktformulierungen - untersucht.

Die Existenz von Sattelpunkten bezüglich elliptischer Minimierungsprobleme wird abschließend in einem allgemeinen Zusammenhang geklärt.

I.3.1 Elliptische Minimierungsprobleme erster Art

Definition I.3 *Es sei V ein reflexiver Raum und $K \subset V$ konvex und abgeschlossen. Es sei $a : V \times V \rightarrow \mathbb{R}$ eine symmetrische, stetige und elliptische Bilinearform und $\ell \in V'$. Das Minimierungsproblem (I.1) heißt elliptisches Minimierungsproblem erster Art, falls*

$$E(v) := \frac{1}{2}a(v, v) - \langle \ell, v \rangle. \quad (\text{I.20})$$

Das Zielfunktional E eines elliptischen Minimierungsproblems erster Art ist für alle $w \in V$ Fréchet-differenzierbar, denn für alle $v \in V$ ist

$$\begin{aligned} E(w+v) - E(w) &= \frac{1}{2}a(w+v, w+v) - \ell(w+v) - \frac{1}{2}a(w, w) - \langle \ell, w \rangle \\ &= a(w, v) - \langle \ell, v \rangle + \frac{1}{2}a(v, v). \end{aligned}$$

Da die Bilinearform a als stetig vorausgesetzt wird, erhält man $a(v, v)/\|v\| \leq \nu_0 \|v\|$, also ist $|a(v, v)| = o(\|v\|)$ für $\|v\| \rightarrow 0$. Für die Fréchet-Ableitung von E in w gilt demnach

$$\langle E'(w), v \rangle = a(w, v) - \langle \ell, v \rangle. \quad (\text{I.21})$$

Nach Satz I.3 folgt die Existenz einer Lösung u , wenn das Zielfunktional E konvex, koerziv und unterhalb stetig ist und die Restriktionsmenge K als abgeschlossen und konvex vorausgesetzt wird. Hinreichend für die Eindeutigkeit der Lösung ist nach Satz I.1 die strikte Konvexität von E . Da sowohl die Bilinearform a als auch die Linearform ℓ als stetig und damit als unterhalb stetig vorausgesetzt werden, ist E unterhalb stetig. Auch die Koerzivität von E ist leicht einzusehen, denn es ist aufgrund der Elliptizität von a

$$E(v) \geq \frac{1}{2}\nu_1\|v\|^2 - \|\ell\|\|v\| = \frac{1}{2}\nu_1 \left(\|v\| - \frac{\|\ell\|}{\nu_1} \right)^2 - \frac{\|\ell\|^2}{2\nu_1}.$$

Die strikte Konvexität von E folgt direkt aus der Tatsache, dass die Bilinearform a elliptisch ist. Denn für $v, w \in V$ mit $v \neq w$ ist $0 < a(v-w, v-w) = a(v, v) - 2a(v, w) + a(w, w)$, und damit ist $2a(v, w) < a(v, v) + a(w, w)$. Demnach ist für $\epsilon \in (0, 1)$

$$\begin{aligned} & a(\epsilon v + (1-\epsilon)w, \epsilon v + (1-\epsilon)w) \\ &= \epsilon^2 a(v, v) + 2\epsilon(1-\epsilon)a(v, w) + (1-\epsilon)^2 a(w, w) \\ &< \epsilon^2 a(v, v) + \epsilon(1-\epsilon)(a(v, v) + a(w, w)) + (1-\epsilon)^2 a(w, w) \\ &= \epsilon a(v, v) + (1-\epsilon)a(w, w). \end{aligned}$$

Zusammen mit der Linearität von ℓ folgt daraus die strikte Konvexität von E . Damit ist insgesamt die eindeutige Existenz einer Lösung gezeigt. Aus Satz I.1 und aus (I.21) erhält man zudem eine äquivalente Variationsformulierung.

Satz I.11 *Das elliptische Minimierungsproblem erster Art besitzt genau eine Lösung $u \in K$. Genau dann ist u eine Lösung, wenn gilt:*

$$\forall v \in K : a(u, v - u) \geq \langle \ell, v - u \rangle. \quad (\text{I.22})$$

Ist $K = V$, dann ist u genau dann eine Lösung, wenn gilt:

$$\forall v \in V : a(u, v) = \langle \ell, v \rangle. \quad (\text{I.23})$$

Zur Formulierung von Sattelpunktproblemen für elliptische Minimierungsprobleme ist es sinnvoll, die Restriktionsmenge K in der Form

$$K := \{v \in V \mid g - \beta_0(v) \in G\} \quad (\text{I.24})$$

vorauszusetzen. Dabei sei $G \subset U_0$ ein abgeschlossener und konvexer Kegel, U_0 ein normierter Raum und $g \in U_0$. Außerdem sei $\beta_0 \in L(V, U_0)$.

Die Menge K ist konvex, denn für $v, w \in K$ und $\epsilon \in [0, 1]$ ist

$$\begin{aligned} g - \beta_0(\epsilon v + (1 - \epsilon)w) &= g - \epsilon\beta_0(v) + (1 - \epsilon)\beta_0(w) \\ &= \epsilon(g - \beta_0(v)) + (1 - \epsilon)(g - \beta_0(w)) \in G. \end{aligned}$$

Darüber hinaus ist K abgeschlossen: Für eine Folge $\{v_n\} \subset K$ mit $v_n \rightarrow v$ konvergiert $\{g - \beta_0(v_n)\} \subset G$ gegen $g - \beta_0(v)$, weil β_0 als stetig vorausgesetzt wird. Da G abgeschlossen ist, folgt $g - \beta_0(v) \in G$.

Für die Restriktionsmenge K wie in (I.24) lässt sich ein Multiplikatorfunktional angeben. Dazu sei im Folgenden $\Lambda_0 := G' := \{\mu \in U_0' \mid \forall w \in G : \langle \mu, w \rangle \geq 0\}$. Die Menge G' ist wegen der Linearität und Stetigkeit von $\langle \cdot, w \rangle$ für ein $w \in V$ ein abgeschlossener Kegel und heißt dualer Kegel von G .

Lemma I.3 *Es sei V ein normierter Raum und $K \subset V$ wie in (I.24). Dann ist das Funktional $\Phi_0 : V \times G' \rightarrow \mathbb{R}$ mit*

$$\Phi_0(v, \mu_0) := \langle \mu_0, \beta_0(v) - g \rangle \quad (\text{I.25})$$

ein Multiplikatorfunktional bzgl. K .

Beweis: Es sei $\alpha \geq 0$, dann ist für ein $v \in V$ und ein $\mu_0 \in \Lambda_0$

$$\Phi_0(v, \alpha\mu_0) = \langle \alpha\mu_0, \beta_0(v) - g \rangle = \alpha\langle \mu_0, \beta_0(v) - g \rangle = \alpha\Phi_0(v, \mu_0),$$

woraus die Eigenschaft (I.13) folgt. Nach Satz A.6 ist $g - \beta_0(v) \in G$ und damit $v \in K$, genau dann, wenn $0 \leq \langle \mu, g - \beta_0(v) \rangle = -\Phi_0(v, \mu)$ für alle $\mu \in G'$ ist. Wegen $G' = \Lambda_0$ gilt damit auch die Eigenschaft (I.14). \square

Bemerkung I.3 Für die Aussage von Lemma I.3 ist nicht entscheidend, dass $\beta_0 \in L(V, U_0)$ ist. Lemma I.3 bleibt gültig, wenn β_0 so gewählt wird, dass K abgeschlossen und konvex ist.

Für das Multiplikatorfunktional Φ_0 wie in (I.25) sind $\Phi_{0,u} := \Phi_0(u, \cdot) = \langle \cdot, \beta_0(u) - g \rangle$ und $\Phi_{0,\lambda_0} := \Phi_0(\cdot, \lambda_0) = \langle \lambda_0, \beta_0(\cdot) - g \rangle$ linear bzw. affin linear und damit konvex und konkav zugleich. Wegen $\Phi_{0,u}(\lambda_0 + \mu_0) - \Phi_{0,u}(\lambda_0) = \langle \mu_0, \beta_0(u) - g \rangle$ und $\Phi_{\lambda_0}(u+v) - \Phi_{\lambda_0}(u) = \langle \lambda_0, \beta_0(v) \rangle$ folgt $\langle \Phi'_{0,u}(\lambda_0), \mu_0 \rangle = \langle \mu_0, \beta_0(u) - g \rangle$ und $\langle \Phi'_{0,\lambda_0}(u), v \rangle = \langle \lambda_0, \beta_0(v) \rangle$. Nach Korollar I.1 gilt demnach:

Satz I.12 *Es sei K wie in (I.24). Dann ist $u \in V$ Lösung des elliptischen Minimierungsproblems erster Art, wenn ein $\lambda_0 \in \Lambda_0$ existiert, so dass:*

$$\forall v \in V : a(u, v) = \langle \ell, v \rangle - \langle \lambda_0, \beta_0(v) \rangle \quad (\text{I.26})$$

$$\forall \mu_0 \in \Lambda_0 : \langle \mu_0 - \lambda_0, \beta_0(u) - g \rangle \leq 0. \quad (\text{I.27})$$

I.3.2 Elliptische Minimierungsprobleme zweiter Art

Definition I.4 *Es seien V ein reflexiver Raum und $K \subset V$ konvex und abgeschlossen. Weiter seien $a : V \times V \rightarrow \mathbb{R}$ eine symmetrische, stetige und elliptische Bilinearform, $\ell \in V'$ und $j : V \rightarrow \mathbb{R}$ ein konvexes, stetiges Funktional. Das Minimierungsproblem (I.1) heißt elliptisches Minimierungsproblem zweiter Art, falls*

$$E(v) := \frac{1}{2}a(v, v) - \ell(v) + j(v). \quad (\text{I.28})$$

Da das Funktional $H : V \rightarrow \mathbb{R}$ mit $H(v) := \frac{1}{2}a(v, v) - \ell(v)$ strikt konvex und unterhalb stetig ist, ist ebenfalls E strikt konvex und unterhalb stetig. Für den Nachweis einer eindeutigen Lösung des elliptischen Minimierungsproblems zweiter Art bleibt demnach nur noch die Koerzivität von E nachzuweisen. Wesentlich hierbei ist, dass ein (unterhalb) stetiges Funktional durch eine stetige, affine Abbildung beschränkt werden kann.

Lemma I.4 *Es seien V ein normierter Raum und $j : V \rightarrow \mathbb{R}$ (unterhalb) stetig und konvex. Dann existieren ein $\mu \in V'$ und ein $\alpha \in \mathbb{R}$, so dass:*

$$\forall v \in V : j(v) > \langle \mu, v \rangle + \alpha.$$

Beweis: Der Epigraph $\text{epi}(j)$ ist konvex und nach Satz A.2 abgeschlossen. Es sei $(w, c) \notin \text{epi}(j)$, dann existiert nach dem Trennungssatz von Hahn-Banach A.4 ein $\zeta \in (V \times \mathbb{R})'$, so dass

$$\langle \zeta, (w, c) \rangle < \inf_{(v, d) \in \text{epi}(j)} \langle \zeta, (v, d) \rangle \quad (\text{I.29})$$

ist. Die Menge $U := \{(v, d) \in V \times \mathbb{R} \mid \langle \zeta, (v, d) \rangle = \langle \zeta, (w, c) \rangle\} = (w, c) + \ker \zeta$ ist ein abgeschlossener affiner Unterraum. Es existieren demnach ein $\mu \in V'$ und ein $\alpha \in \mathbb{R}$ mit

$$\{(v, d) \in V \times \mathbb{R} \mid d = \langle \mu, v \rangle + \alpha\} = U. \quad (\text{I.30})$$

Angenommen, es existierte ein $\tilde{v} \in V$ mit $j(\tilde{v}) \leq \langle \mu, \tilde{v} \rangle + \alpha$, dann wäre $(\tilde{v}, \langle \mu, \tilde{v} \rangle + \alpha) \in \text{epi}(j)$. Nach (I.29) wäre dann $\langle \zeta, (w, c) \rangle < \langle \zeta, (\tilde{v}, \langle \mu, \tilde{v} \rangle + \alpha) \rangle$. Wegen (I.30) ist aber $(\tilde{v}, \langle \mu, \tilde{v} \rangle + \alpha) \in U$. \square

Aus der Elliptizität von a und aus Lemma I.4 folgt

$$\begin{aligned} E(v) &\geq \frac{1}{2}\nu_1\|v\|^2 - \|\ell\|\|v\| + \langle\mu, v\rangle + \alpha \\ &\geq \frac{1}{2}\nu_1\|v\|^2 - (\|\ell\| + \|\mu\|)\|v\| + \alpha \\ &= \frac{1}{2}\nu_1\left(\|v\| - \frac{\|\ell\| + \|\mu\|}{\nu_1}\right)^2 - \frac{(\|\ell\| + \|\mu\|)^2}{2\nu_1} + \alpha \end{aligned}$$

und damit die Koerzivität von E . Zusammenfassend erhält man demnach:

Satz I.13 *Das elliptische Minimierungsproblem zweiter Art besitzt genau eine Lösung $u \in K$. Genau dann ist u eine Lösung, wenn gilt:*

$$\forall v \in K : a(u, v - u) + j(v) - j(u) \geq \langle\ell, v - u\rangle. \quad (\text{I.31})$$

Für die Formulierung von Sattelpunktproblemen elliptischer Minimierungsprobleme zweiter Art ist es zweckmäßig, nur solche Multiplikatorfunktionale $\Phi_1 : V \times \Lambda_1 \rightarrow \mathbb{R}$ bzgl. j zu betrachten, die einer Darstellung der Form

$$\Phi_1(v, \mu_1) = \langle\mu_1, \beta_1(v)\rangle \quad (\text{I.32})$$

mit einer Abbildung $\beta_1 \in L(V, U_1)$ genügen. Hierbei ist U_1 ein normierter Raum und $\Lambda_1 \subset U_1'$. Für den unrestringierten Fall erhält man dann nach Korollar I.2:

Satz I.14 *Es existiere ein Multiplikatorfunktional Φ_1 bzgl. $j : V \rightarrow \mathbb{R}$ mit der Darstellung wie in (I.32). Dann ist u Lösung des elliptischen Minimierungsproblems zweiter Art mit $K = V$, wenn ein $\lambda_1 \in \Lambda_1$ existiert, so dass:*

$$\forall v \in V : a(u, v) = \langle\ell, v\rangle - \langle\lambda_1, \beta_1(v)\rangle \quad (\text{I.33})$$

$$\forall \mu_1 \in \Lambda_1 : \langle\mu_1 - \lambda_1, \beta_1(u)\rangle \leq 0. \quad (\text{I.34})$$

Für den restringierten Fall erhält man:

Satz I.15 *Das Multiplikatorfunktional Φ_0 bzgl. K sei wie in (I.25) gewählt, und es existiere ein Multiplikatorfunktional Φ_1 bzgl. j wie in (I.32). Dann ist u Lösung des elliptischen Minimierungsproblems zweiter Art, wenn ein $(\lambda_0, \lambda_1) \in \Lambda_0 \times \Lambda_1$ existiert, so dass:*

$$\forall v \in V : a(u, v) = \langle\ell, v\rangle - \langle\lambda_0, \beta_0(v)\rangle - \langle\lambda_1, \beta_1(v)\rangle \quad (\text{I.35})$$

$$\forall(\mu_0, \mu_1) \in \Lambda_0 \times \Lambda_1 : \langle\mu_0 - \lambda_0, \beta_0(u) - g\rangle + \langle\mu_1 - \lambda_1, \beta_1(u)\rangle \leq 0. \quad (\text{I.36})$$

I.3.3 Existenz und Eindeutigkeit von Sattelpunkten

Zur Herleitung von Aussagen bezüglich Existenz und Eindeutigkeit von Lösungen der Systeme (I.26)-(I.27), (I.33)-(I.34) und (I.35)-(I.36) wird im Folgenden das allgemeine Lagrange-Funktional $\mathcal{L} : V \times \Lambda \rightarrow \mathbb{R}$ mit

$$\mathcal{L}(v, \mu) := \frac{1}{2}a(v, v) - \langle \ell, v \rangle + b(v, \mu) - \langle \mu, \tilde{g} \rangle$$

betrachtet, wobei V und U' reflexive Räume sind, $\Lambda \subset U'$ abgeschlossen und konvex sowie $b : V \times U' \rightarrow \mathbb{R}$ eine stetige Bilinearform mit $b(v, \mu) \leq c\|v\|\|\mu\|$ ($c > 0$) und $\tilde{g} \in U$ ist.

Offenbar ist \mathcal{L}_μ strikt konvex, (unterhalb) stetig und wegen

$$\begin{aligned} \mathcal{L}_\mu &\geq \frac{1}{2}\nu_1\|v\|^2 - \|\ell\|\|v\| - c\|v\|\|\mu\| - \langle \mu, \tilde{g} \rangle \\ &= \frac{1}{2}\nu_1 \left(\|v\| - \frac{\|\ell\| + c\|\mu\|}{\nu_1} \right)^2 - \frac{(\|\ell\| + c\|\mu\|)^2}{2\nu_1} - \langle \mu, \tilde{g} \rangle \end{aligned}$$

für alle $\mu \in \Lambda$ koerziv. Die Abbildung \mathcal{L}_v ist konkav und (oberhalb) stetig. Ist Λ beschränkt, folgt direkt aus Satz I.6, dass \mathcal{L} einen Sattelpunkt hat. Ist Λ unbeschränkt, dann ist für den Nachweis der Existenz eines Sattelpunkts nach Satz I.6 die Koerzivität von $\mu \mapsto \sup_{v \in V} (-\mathcal{L}(v, \mu))$ erforderlich. Eine zentrale Rolle spielt hierbei die Bedingung

$$\exists \alpha \in \mathbb{R}_{>0} : \forall \mu \in U' : \alpha\|\mu\| \leq \sup_{v \in V \setminus \{0\}} \frac{b(v, \mu)}{\|v\|}, \quad (\text{I.37})$$

die auch als Babuška-Brezzi Bedingung bekannt ist.⁷

Lemma I.5 *Ist (I.37) erfüllt, dann ist $\mu \mapsto \sup_{v \in V} (-\mathcal{L}(v, \mu))$ koerziv.*

Beweis: Da \mathcal{L}_μ (unterhalb) stetig und koerziv für $\mu \in U'$ ist, existiert nach Satz I.3 ein v_μ , so dass $\mathcal{L}_\mu(v_\mu) = \inf_{v \in V} \mathcal{L}_\mu(v)$ ist. Aus Satz I.1 folgt, dass $a(v_\mu, v) = \langle \ell, v \rangle - b(v, \mu)$ für alle $v \in V$ ist. Demnach ist

$$-\mathcal{L}_{v_\mu}(\mu) = \frac{1}{2}a(v_\mu, v_\mu) + \langle \mu, \tilde{g} \rangle \geq \frac{1}{2}\nu_1\|v_\mu\|^2 - \|\mu\|\|\tilde{g}\|. \quad (\text{I.38})$$

Andererseits ist

$$\alpha\|\mu\| \leq \sup_{v \in V \setminus \{0\}} \frac{b(v, \mu)}{\|v\|} = \sup_{v \in V} \frac{\langle \ell, v \rangle - a(v_\mu, v)}{\|v\|} \leq \|\ell\| + \nu_0\|v_\mu\|. \quad (\text{I.39})$$

⁷vgl. Lemma 3.2. in [72]

Wenn $\|\mu\| \rightarrow \infty$, dann folgt aus (I.39), dass auch $\|v_\mu\| \rightarrow \infty$ gilt. Aus (I.38) und (I.39) folgt dann $\sup_{v \in V} (-\mathcal{L}(v, \mu)) \rightarrow \infty$. \square

Bedingungen für die Eindeutigkeit einer Lösung ergeben sich aus den folgenden Aussagen. Aus der Koerzivität von a und aus Satz I.1 folgt:

Lemma I.6 *Es sei $(u, \lambda) \in V \times U'$ ein Sattelpunkt von \mathcal{L} . Dann ist u eindeutig. Gilt für $\mu \in U'$*

$$\forall v \in V : b(\mu, v) = 0 \quad \Rightarrow \quad \mu = 0, \quad (\text{I.40})$$

dann ist auch λ eindeutig. Ist $(u, \lambda^) \in V \times U'$ ein weiterer Sattelpunkt, dann gilt:*

$$\forall v \in V : b(\lambda - \lambda^*, v) = 0.$$

Lemma I.7 *Ist (I.37) erfüllt, dann gilt (I.40).*

Beweis: Es sei $\mu \in U'$ mit $b(\mu, v) = 0$ für alle $v \in V$, dann ist $\|\mu\| \leq \alpha^{-1} \sup_{v \in V} b(v, \mu) / \|v\| = 0$. Also ist $\mu = 0$. \square

Insgesamt gilt damit:

Satz I.16 *Ist Λ beschränkt oder gilt (I.37), dann besitzt \mathcal{L} einen Sattelpunkt (u, λ) , wobei u eindeutig ist. Wenn (I.40) erfüllt ist, ist auch λ eindeutig.*

Aus Satz I.7 folgt außerdem:

Satz I.17 *Es seien U'_0 und U'_1 reflexive Räume und $\Lambda_0 \subset U'_0$ und $\Lambda_1 \subset U'_1$ abgeschlossen und konvex. Für $\Lambda := \Lambda_0 \times \Lambda_1$ sei zudem $b(v, \mu_0, \mu_1) := b_0(\mu_0, v) + b_1(\mu_1, v)$ mit den stetigen Bilinearformen b_0 und b_1 . Ist (I.37) bzgl. b_0 erfüllt, und ist Λ_1 beschränkt, dann besitzt \mathcal{L} einen Sattelpunkt $(u, \lambda_0, \lambda_1)$, wobei u eindeutig ist. Ist (I.40) bzgl. b erfüllt, dann sind auch λ_0 und λ_1 eindeutig.*

Im Fall elliptischer Minimierungsprobleme erster Art ist $U' := U'_0 / \ker \beta'_0$, $\Lambda := \{\mu_0 + \ker \beta'_0 \mid \mu_0 \in \Lambda_0\}$ und $b(v, [\mu_0]) := \langle \mu_0, \beta_0(v) \rangle$ zu setzen.⁸ Ist $g \in \text{rg}(\beta_0)$, dann ist $\tilde{g} := g \in U$.⁹

Da Λ_0 unbeschränkt ist, ist die Bedingung (I.37) nachzuweisen. Ein wichtiges Kriterium liefert hierfür der Satz A.11 vom abgeschlossenen Bild:

⁸Der Operator $\beta'_0 \in L(U'_0, V')$ ist der zu β_0 adjungierte Operator. Mit $[\cdot]$ werden hier die Elemente des Quotientenraums bezeichnet.

⁹Hierbei ist $\text{rg}(\beta_0) := \beta_0(V)$.

Lemma I.8 *Ist $\text{rg}(\beta_0)$ abgeschlossen in U_0 , dann ist die Bedingung (I.37) erfüllt.*

Beweis: Da $\text{rg}(\beta_0)$ abgeschlossen ist, existiert nach Satz A.10 ein $c > 0$, so dass

$$\forall w \in \text{rg}(\beta_0) : \exists v \in V : \beta_0(v) = w, \|v\| \leq c\|w\|$$

ist. Für $\tilde{V} := \{v \in V \mid \|v\| \leq c\|\beta_0(v)\|\}$ gilt demnach $\beta_0(\tilde{V}) = \text{rg}(\beta_0)$, womit für $\mu_0 \in U'_0$ gilt:

$$\begin{aligned} \sup_{v \in V \setminus \{0\}} \frac{\langle \mu_0, \beta_0(v) \rangle}{\|v\|} &\geq \sup_{v \in \tilde{V} \setminus \{0\}} \frac{\langle \mu_0, \beta_0(v) \rangle}{\|v\|} \geq c^{-1} \sup_{v \in \tilde{V} \setminus \{0\}} \frac{\langle \mu_0, \beta_0(v) \rangle}{\|\beta_0(v)\|} \\ &= c^{-1} \sup_{w \in \text{rg}(\beta_0) \setminus \{0\}} \frac{\langle \mu_0, w \rangle}{\|w\|} = c^{-1} \|\mu_0|_{\text{rg}(\beta_0)}\|. \end{aligned}$$

Nach dem Satz A.11 vom abgeschlossenen Bild ist $\text{rg}(\beta_0) = (\ker \beta'_0)_\perp$.¹⁰ Demnach ist nach den Sätzen A.7 und A.8 $(\text{rg}(\beta_0))' \simeq U'_0 / \text{rg}(\beta_0)^\perp = U'_0 / ((\ker \beta'_0)_\perp)^\perp = U'_0 / \ker \beta'_0$ und damit $\|\mu_0|_{\text{rg}(\beta_0)}\| = \|\mu_0\|$. \square

Aus Satz I.16 erhält man damit unmittelbar:

Satz I.18 *Ist $g \in \text{rg}(\beta_0)$ und $\text{rg}(\beta_0)$ abgeschlossen in U_0 , dann besitzt das System (I.26)-(I.27) eine Lösung (u, λ_0) , wobei u eindeutig ist. Ist β_0 surjektiv, dann ist auch λ_0 eindeutig.*

Im Fall elliptischer Minimierungsprobleme zweiter Art mit $K = V$ setzt man $U' := U'_1$, $\Lambda := \Lambda_1$ und $b(v, \mu_1) := \langle \mu_1, \beta_1(v) \rangle$ und $\tilde{g} = 0$. Aus Satz I.16 folgt:

Satz I.19 *Ist Λ_1 beschränkt, dann besitzt das System (I.33)-(I.34) eine Lösung (u, λ_1) , wobei u eindeutig ist. Ist β_1 surjektiv, dann ist auch λ_1 eindeutig.*

Im Fall elliptischer Minimierungsprobleme zweiter Art mit $K \subset V$ setzt man: $U' := U'_0 / \ker \beta'_0 \times U'_1$, $\Lambda := \{\mu_0 + \ker \beta'_0 \mid \mu_0 \in \Lambda_0\} \times \Lambda_1$ und $b(v, [\mu_0], \mu_1) := \langle \mu_0, \beta_0(v) \rangle + \langle \mu_1, \beta_1(v) \rangle$ und $\tilde{g} = (g, 0)$. Aus Satz I.17 erhält man:

Satz I.20 *Ist $g \in \text{rg}(\beta_0)$ und $\text{rg}(\beta_0)$ in U_0 abgeschlossen sowie Λ_1 beschränkt, dann besitzt das System (I.35)-(I.36) einen Sattelpunkt $(u, \lambda_0, \lambda_1)$, wobei u eindeutig ist. Gilt zudem $\beta_0(\ker \beta_1) = U_0$ und $\beta_1(\ker \beta_0) = U_1$, dann sind auch λ_0 und λ_1 eindeutig.*

¹⁰Für einen normierten Raum V , Unterräume $M \subset V$ und $N \subset V'$ heißen die Mengen $M^\perp := \{\phi \in V' \mid \forall v \in M : \langle \phi, v \rangle = 0\}$ und $N_\perp := \{v \in V \mid \forall \phi \in N : \langle \phi, v \rangle = 0\}$ die Annihilatoren von M bzw. N .

I.4 Variationsungleichungen auf Hilberträumen

Im Folgenden seien U , V und W Hilberträume und $\beta \in L(V, U)$ sowie $a : V \times V \rightarrow \mathbb{R}$ eine stetige Bilinearform. Zudem sei W mit seinem Dualraum identifiziert, β surjektiv, V mit einer stärkeren Topologie in W enthalten und $\ker \beta$ dicht in W .

Für ein $u \in V$ sei $A(u) \in (\ker \beta)'$ definiert als $(A(u), \cdot) = \langle A(u), \cdot \rangle := a(u, \cdot)$. Offenbar ist $A \in L(V, (\ker \beta)')$. Es sei $\text{dom}(A) := \{v \in V \mid A(v) \in W\}$, dann ist $A : V \supset \text{dom}(A) \rightarrow W$ abgeschlossen. Nach Satz A.12 ist $\text{dom}(A)$, versehen mit der Norm $\|u\|_{\text{dom}(A)}^2 := \|u\|_V^2 + \|A(u)\|_W^2$, ein Hilbertraum sowie $A \in L(\text{dom}(A), W)$.

Satz I.21 *Es existiert ein eindeutiger Operator $\alpha \in L(\text{dom}(A), U')$ mit*

$$\forall w \in \text{dom}(A), \forall v \in V : a(w, v) = (A(w), v) + \langle \alpha(w), \beta(v) \rangle.$$

Beweis: Ch.6, Th.2.1 in [5]. □

Der Operator A heißt der zu a assoziierte Operator. Mit seiner Hilfe können elliptische Minimierungsprobleme auch über Kräftebilanzen gedeutet werden, die häufig aus Ingenieur-technischer Sicht vorteilhafter erscheinen, da die darin beschriebenen Größen einer physikalischen Interpretation leichter zugänglich sind¹¹. Dabei repräsentiert $A(u)$ für eine Lösung u im Allgemeinen die in einem System auftretenden inneren Kräfte, die mit auf das System wirkenden äußeren Kräften $f \in W$ und $q \in U'$ im Gleichgewicht stehen müssen. Das aus physikalischer Sicht zu erwartende Gleichgewicht zwischen inneren und äußeren Kräften wird bei elliptischen Minimierungsproblemen ohne Restriktionen durch die folgenden Aussagen wiedergegeben:

Lemma I.9 *Gilt für $u \in \text{dom}(A)$*

$$\forall v \in \ker \beta : a(u, v) = (f, v),$$

dann ist $A(u) = f$.

Beweis: Für alle $v \in \ker \beta$ ist $(A(u) - f, v) = a(u, v) - (f, v) = 0$, also ist $A(u) - f \perp \ker \beta$. Da $\ker \beta \subset W$ dicht ist, folgt $A(u) - f \perp W$. □

Mit $\langle \ell, v \rangle := (f, v) + \langle q, \beta(v) \rangle$ gilt dann:

¹¹vgl. Abschnitt II.3

Satz I.22 *Es sei $u \in \text{dom}(A)$. Genau dann ist u Lösung von (I.23) mit $K = V$, wenn gilt:*

$$A(u) = f \tag{I.41}$$

$$\alpha(u) = q. \tag{I.42}$$

Beweis: Es gelte (I.23), dann folgt aus Lemma I.9 die Gleichung (I.41). Außerdem ist für $v \in V$

$$\langle q, \beta(v) \rangle = a(u, v) - \langle f, v \rangle = \langle \alpha(u), \beta(v) \rangle + \langle A(u), v \rangle - \langle f, v \rangle = \langle \alpha(u), \beta(v) \rangle.$$

Da β surjektiv ist, folgt (I.42). Andererseits, wenn u (I.41)-(I.42) erfüllt, dann ist

$$\langle f, v \rangle + \langle q, \beta(v) \rangle = \langle A(u), v \rangle + \langle \alpha(u), \beta(v) \rangle = a(u, v)$$

für alle $v \in V$, also erfüllt u auch (I.23). \square

Ziel dieses Abschnitts ist, Kräftebilanz-Systeme herzuleiten, die allgemeinen elliptischen Minimierungsproblemen erster und zweiter Art bzw. den dazu äquivalenten Variationsungleichungen (I.22) und (I.31) zugeordnet sind. Im Folgenden wird (I.22) als (elliptische) Variationsungleichung erster Art und (I.31) als (elliptische) Variationsungleichung zweiter Art bezeichnet. Die Tatsache, dass elliptische Minimierungsprobleme elliptische und symmetrische Bilinearformen implizieren, wird bei den folgenden Betrachtungen nicht weiter berücksichtigt. Dementsprechend bleiben die Aussagen auch für Variationsungleichungen mit nicht symmetrischen und nicht elliptischen Bilinearformen gültig.

I.4.1 Variationsungleichungen erster Art

Im Folgenden sei $U = U_0 \times U_1$ mit den Hilberträumen U_0 und U_1 und $\beta = (\beta_0, \beta_1)$ mit $\beta_0 \in L(V, U_0)$ und $\beta_1 \in L(V, U_1)$. Für $\mu \in U'$ seien $\mu_0 \in U_0'$ und $\mu_1 \in U_1'$ definiert für $v_0 \in U_0$ und $v_1 \in U_1$ als $\langle \mu_0, v_0 \rangle := \langle \mu, (v_0, 0) \rangle$ und $\langle \mu_1, v_1 \rangle := \langle \mu, (0, v_1) \rangle$. Für $v := (v_0, v_1)$ ist dann offenbar $\langle \mu, v \rangle = \langle \mu_0, v_0 \rangle + \langle \mu_1, v_1 \rangle$. Für $i = 0, 1$ sei $\alpha_i(u) := (\alpha(u))_i$. Wenn nicht anders angegeben, seien außerdem K wie (I.24) und $\Lambda_0 := G'$.

Satz I.23 *Es sei $u \in K \cap \text{dom}(A)$ und $\beta_1(\ker \beta_0) = U_1$. Genau dann ist u Lösung der Variationsungleichung erster Art (I.22), wenn gilt:*

$$A(u) = f \quad (\text{I.43})$$

$$q_0 - \alpha_0(u) \in \Lambda_0 \quad (\text{I.44})$$

$$\langle q_0 - \alpha_0(u), \beta_0(u) - g \rangle = 0 \quad (\text{I.45})$$

$$\alpha_1(u) = q_1 \quad (\text{I.46})$$

Beweis: Es sei $w \in \ker \beta$, dann ist $v := u \pm w \in K$. Also ist $a(u, \pm w) \geq (f, \pm w)$ und damit ist $a(u, w) = (f, w)$. Aus Lemma I.9 folgt dann (I.43). Aus (I.22) erhält man für $v \in K$

$$\begin{aligned} & \langle q_0 - \alpha_0(u), \beta_0(v - u) \rangle + \langle q_1 - \alpha_1(u), \beta_1(v - u) \rangle \\ &= \langle q - \alpha(u), \beta(v - u) \rangle \\ &= \langle q, \beta(v - u) \rangle - a(u, v - u) + (A(u), v - u) \\ &= \langle q, \beta(v - u) \rangle - a(u, v - u) + (f, v - u) \\ &\leq 0. \end{aligned} \quad (\text{I.47})$$

Es sei $v \in \ker \beta_0$, dann ist $u \pm v \in K$. Einsetzen in (I.47) liefert (I.46), da $\beta_1(\ker \beta_0) = U_1$ ist. Es sei $w \in V$ mit $\beta_0(w) = g$, dann sind w und $2u - w$ in K , da $0 \in G$ und $g - \beta_0(2u - w) = 2(g - \beta_0(u)) \in G$. Einsetzen für v in (I.47) liefert (I.45). Ist $\beta_0(v) \in G$, so ist $w - v \in K$. Aus (I.47) und (I.45) folgt

$$\langle q_0 - \alpha_0(u), \beta_0(v) \rangle = -(\langle q_0 - \alpha_0(u), \beta_0(w - v - u) \rangle + \langle q_0 - \alpha_0(u), \beta_0(u) - g \rangle) \geq 0$$

Da β_0 surjektiv ist, folgt schließlich (I.44).

Es erfülle u die Bedingungen (I.43)-(I.46). Ist $v \in K$, dann ist $w - v \in G$, und damit gilt

$$\begin{aligned} & (f, v - u) + \langle q, \beta(v - u) \rangle \\ &= (A(u), v - u) + \langle q - \alpha(u), \beta(v - u) \rangle + \langle \alpha(u), \beta(v - u) \rangle \\ &= a(u, v - u) - \langle q_0 - \alpha_0(u), \beta_0(w - v) \rangle + \langle q_0 - \alpha_0(u), g - \beta_0(u) \rangle \\ &\leq a(u, v - u) \end{aligned}$$

Also erfüllt u die Bedingung (I.22). \square

Satz I.24 *Es sei $u \in \text{dom}(A)$ und $\beta_1(\ker \beta_0) = U_1$. Genau dann ist (u, λ_0) Lösung von (I.26)-(I.27), wenn $u \in K$ ist, u (I.22) oder (I.43)-(I.46) erfüllt, und wenn gilt:*

$$\lambda_0 = q_0 - \alpha_0(u). \quad (\text{I.48})$$

Beweis: Es sei (u, λ_0) Lösung von (I.26)-(I.27). Dann liefert Einsetzen von 0 und $2\lambda_0$

$$\langle \lambda_0, \beta_0(u) - g \rangle = 0. \quad (\text{I.49})$$

Daraus folgt, dass $\langle \mu_0, g - \beta_0(u) \rangle \geq 0$ für alle $\mu_0 \in \Lambda_0$ ist. Wegen $\Lambda_0 = G'$ folgt damit aus Satz A.6 $g - \beta_0(u) \in G$, also $u \in K$. Es sei $v \in K$, dann liefern Subtrahieren von (I.26) und Anwenden von (I.49)

$$\begin{aligned} a(u, v - u) &= (f, v - u) + \langle q, \beta(v - u) \rangle - \langle \lambda_0, \beta_0(v - u) \rangle \\ &= (f, v - u) + \langle q, \beta(v - u) \rangle + \langle \lambda_0, g - \beta_0(v) \rangle \\ &\geq (f, v - u) + \langle q, \beta(v - u) \rangle, \end{aligned}$$

da $\lambda_0 \in G'$ und $g - \alpha(v) \in G$ ist. Demnach gilt (I.22). Nach Satz I.23 ist dann auch das System (I.43)-(I.46) erfüllt. Für $v \in V$ folgt dann

$$\begin{aligned} \langle q_0 - \alpha_0(u), \beta_0(v) \rangle &= \langle q - \alpha(u), \beta(v) \rangle \\ &= \langle q, \beta(v) \rangle - a(u, v) + (A(u), v) \\ &= \langle q, \beta(v) \rangle - a(u, v) + (f, v) \\ &= \langle \lambda_0, \beta_0(v) \rangle. \end{aligned}$$

Da β_0 surjektiv ist, ist damit auch (I.48) erfüllt.

Wenn andererseits $u \in K$ ist, und (u, λ_0) das System (I.43)-(I.46) und (I.48) erfüllt, dann folgt für $v \in V$

$$\begin{aligned} (f, v) + \langle q, \beta(v) \rangle - \langle \lambda_0, \beta_0(v) \rangle &= (A(u), v) + \langle q_1, \beta_1(v) \rangle + \langle \alpha_0(u), \beta_0(v) \rangle \\ &= (A(u), v) + \langle \alpha(u), \beta(v) \rangle \\ &= a(u, v). \end{aligned}$$

Demnach gilt (I.26). Wegen (I.45) ist $\langle \lambda_0, \beta_0(u) - g \rangle = 0$. Damit ist für alle $\mu_0 \in \Lambda_0$

$$\langle \mu_0 - \lambda_0, \beta_0(u) - g \rangle = -\langle \mu_0, g - \beta_0(u) \rangle \leq 0,$$

da $g - \beta_0(u) \in G$ ist. Also gilt auch (I.27). \square

I.4.2 Variationsungleichungen zweiter Art

Im Folgenden habe das Funktional $j : V \rightarrow \mathbb{R}_{\geq 0}$ die Eigenschaften:

$$\forall v \in \ker \beta_1 : j(v) = 0, \quad (\text{I.50})$$

$$\forall v \in V : j(-v) = j(v), \quad (\text{I.51})$$

$$\forall v, w \in V : j(v + w) \leq j(v) + j(w), \quad (\text{I.52})$$

und es sei

$$\Lambda_1 := \{\mu_1 \in U_1' \mid \forall v \in V : |\langle \mu_1, \beta_1(v) \rangle| \leq j(v)\}. \quad (\text{I.53})$$

Satz I.25 *Es sei $u \in \text{dom}(A)$ und $\beta_0(\ker \beta_1) = U_0$. Genau dann ist u Lösung der Variationsungleichung zweiter Art (I.31) mit $K = V$, wenn gilt:*

$$A(u) = f \quad (\text{I.54})$$

$$q_1 - \alpha_1(u) \in \Lambda_1 \quad (\text{I.55})$$

$$\langle q_1 - \alpha_1(u), \beta_1(u) \rangle = j(u) \quad (\text{I.56})$$

$$\alpha_0(u) = q_0. \quad (\text{I.57})$$

Beweis: Einsetzen von $\pm w \in \ker \beta$ liefert $a(u, w) = (f, w)$. Aus Lemma I.9 folgt dann (I.54). Aus (I.31) folgt für $v \in V$

$$\begin{aligned} & \langle q_0 - \alpha_0(u), \beta_0(v - u) \rangle + \langle q_1 - \alpha_1(u), \beta_1(v - u) \rangle \\ &= \langle q - \alpha(u), \beta(v - u) \rangle \\ &= \langle q, \beta(v - u) \rangle - a(u, v - u) + (A(u), v - u) \\ &= \langle q, \beta(v - u) \rangle - a(u, v - u) + (f, v - u) \\ &\leq j(v) - j(u). \end{aligned} \quad (\text{I.58})$$

Es sei $v \in \ker \beta_1$, dann liefert Einsetzen von $u \pm v$ in (I.58) unter Ausnutzen von (I.51), (I.52) und $\beta_0(\ker \beta_1) = U_0$ die Bedingung (I.57). Wird 0 und $2u$ in (I.58) eingesetzt, erhält man wegen (I.51) und (I.52) die Bedingung (I.56). Für $\pm v \in V$ folgt aus (I.58)

$$\begin{aligned} \pm \langle q_1 - \alpha_1(u), \beta_1(v) \rangle &= \langle q_1 - \alpha_1(u), \beta(\pm v - u) \rangle + \langle q_1 - \alpha_1(u), \beta_1(u) \rangle \\ &\leq j(\pm v) - j(u) + j(u) \\ &= j(v), \end{aligned}$$

und damit gilt (I.55).

Wenn andererseits ein $u \in V$ das System (I.54)-(I.57) erfüllt, dann ist für $v \in V$

$$\begin{aligned} & (f, v - u) + \langle q, \beta(v - u) \rangle \\ &= (A(u), v - u) + \langle q, \beta(v - u) \rangle \\ &= a(u, v - u) - \langle \alpha(u), \beta(v - u) \rangle + \langle q, \beta(v - u) \rangle \\ &= a(u, v - u) + \langle q_1 - \alpha_1(u), \beta_1(v) \rangle - \langle q_1 - \alpha_1(u), \beta_1(u) \rangle \\ &\leq a(u, v - u) + j(v) - j(u). \end{aligned}$$

Also erfüllt u auch (I.31). □

Ein Kriterium für die Eigenschaften (I.50)-(I.52) erhält man aus der folgenden Aussage:¹²

Lemma I.10 *Gilt für $j : V \rightarrow \mathbb{R}_{\geq 0}$*

$$j(v) = \sup_{\mu_1 \in \Lambda_1} \langle \mu_1, \beta_1(v) \rangle, \quad (\text{I.59})$$

dann sind die Eigenschaften (I.50)-(I.52) erfüllt.

Beweis: Für $v \in \ker \beta_1$ ist offenbar $j(v) = 0$, und es gilt für $v, w \in V$

$$\begin{aligned} j(v+w) &= \sup_{\mu_1 \in \Lambda_1} \langle \mu_1, \beta_1(v+w) \rangle \\ &\leq \sup_{\mu_1 \in \Lambda_1} \langle \mu_1, \beta_1(v) \rangle + \sup_{\mu_1 \in \Lambda_1} \langle \mu_1, \beta_1(w) \rangle \\ &= j(v) + j(w). \end{aligned}$$

Mit $\mu_1 \in \Lambda_1$ ist auch $-\mu_1 \in \Lambda_1$, woraus $j(-v) = \sup_{\mu_1 \in \Lambda_1} \langle -\mu_1, \beta_1(v) \rangle = j(v)$ folgt. \square

Satz I.26 *Es seien $u \in \text{dom}(A)$ und $\beta_0(\ker \beta_1) = U_0$ sowie j wie in (I.59). Genau dann ist (u, λ_1) Lösung von (I.33)-(I.34), wenn u (I.31) oder (I.54)-(I.57) erfüllt und wenn gilt:*

$$\lambda_1 = q_1 - \alpha_1(u). \quad (\text{I.60})$$

Beweis: Es sei (u, λ_1) Lösung von (I.33)-(I.34) sowie $v \in V$, dann ist nach Subtrahieren von (I.33) und Anwenden von (I.34) für $\mu_1 \in \Lambda_1$

$$\begin{aligned} a(u, v-u) &= (f, v-u) + \langle q, \beta(v-u) \rangle - \langle \lambda_1, \beta_1(v) \rangle + \langle \lambda_1, \beta_1(u) \rangle \\ &\geq (f, v-u) + \langle q, \beta(v-u) \rangle - j(v) + \langle \mu_1, \beta_1(u) \rangle. \end{aligned}$$

Also gilt

$$\begin{aligned} a(u, v-u) &\geq (f, v-u) + \langle q, \beta(v-u) \rangle - j(v) + \sup_{\mu_1 \in \Lambda_1} \langle \mu_1, \beta_1(u) \rangle \\ &= (f, v-u) + \langle q, \beta(v-u) \rangle - j(v) + j(u). \end{aligned}$$

Demnach folgt (I.31). Nach Satz I.25 ist dann auch das System (I.54)-(I.57) erfüllt. Für alle $v \in V$ folgt dann

$$\begin{aligned} \langle q_1 - \alpha_1(u), \beta_1(v) \rangle &= \langle q - \alpha(u), \beta(v) \rangle \\ &= \langle q, \beta(v) \rangle - a(u, v) + (A(u), v) \\ &= \langle q, \beta(v) \rangle - a(u, v) + (f, v) \\ &= \langle \lambda_1, \beta_1(v) \rangle. \end{aligned}$$

¹²vgl. (I.32)

Da β_1 surjektiv ist, ist demnach auch (I.60) erfüllt.

Wenn andererseits (u, λ_1) (I.54)-(I.57) und (I.60) erfüllt, dann folgt für $v \in V$

$$\begin{aligned} \langle f, v \rangle + \langle q, \beta(v) \rangle - \langle \lambda_1, \beta_1(v) \rangle &= \langle A(u), v \rangle + \langle q_0, \beta_0(v) \rangle + \langle \alpha_1(u), \beta_1(v) \rangle \\ &= \langle A(u), v \rangle + \langle \alpha(u), \beta(v) \rangle \\ &= a(u, v). \end{aligned}$$

Demnach gilt (I.33). Wegen (I.56) ist $\langle \lambda_1, \beta_1(u) \rangle = j(u)$. Damit ist für alle $\mu_1 \in \Lambda_1$

$$\langle \mu_1 - \lambda_1, \beta_1(u) \rangle = \langle \mu_1, \beta_1(u) \rangle - j(u) \leq 0.$$

Also gilt auch (I.34). \square

Satz I.27 *Es seien $u \in K \cap \text{dom}(A)$, $\beta_1(\ker \beta_0) = U_1$ und $\beta_0(\ker \beta_1) = U_0$. Genau dann ist u Lösung der Variationsungleichung zweiter Art (I.31), wenn gilt:*

$$A(u) = f \tag{I.61}$$

$$(q_0 - \alpha_0(u), q_1 - \alpha_1(u)) \in \Lambda_0 \times \Lambda_1 \tag{I.62}$$

$$\langle q_0 - \alpha_0(u), \beta_0(u) - g \rangle = 0 \tag{I.63}$$

$$\langle q_1 - \alpha_1(u), \beta_1(u) \rangle = j(u). \tag{I.64}$$

Beweis: Einsetzen von $\pm w \in \ker \beta$ liefert $a(u, w) = \langle f, w \rangle$. Aus Lemma I.9 folgt dann (I.61). Analog zu (I.58) folgt aus (I.31) für $v \in K$

$$\langle q_0 - \alpha_0(u), \beta_0(v - u) \rangle + \langle q_1 - \alpha_1(u), \beta_1(v - u) \rangle \leq j(v) - j(u). \tag{I.65}$$

Wegen $\beta_0(\ker \beta_1) = U_0$ und $\beta_1(\ker \beta_0) = U_1$ existieren $u_0 \in \ker \beta_1$ und $u_1 \in \ker \beta_0$ mit $\beta_0(u_0) = \beta_0(u)$ und $\beta_1(u_1) = \beta_1(u)$. Für $v \in K \cap \ker \beta_1$ ist $v + u - u_0 \in K$, damit folgt aus (I.65)

$$\begin{aligned} \langle q_0 - \alpha_0(u), \beta_0(v - u_0) \rangle &\leq j(v + u - u_0) - j(u) - \langle q_1 - \alpha_1(u), \beta_1(v - u_0) \rangle \\ &\leq j(v - u_0) \\ &= 0. \end{aligned} \tag{I.66}$$

Es sei $w \in \ker \beta_1$ mit $\beta_0(w) = g$, dann sind w und $2u_0 - w$ in $K \cap \ker \beta_1$, da $0 \in G$ und $g - \beta_0(2u_0 - w) = 2(g - \beta_0(u)) \in G$. Einsetzen in (I.66) liefert

$$0 = \langle q_0 - \alpha_0(u), \beta_0(w - u_0) \rangle = \langle q_0 - \alpha_0(u), g - \beta_0(u) \rangle,$$

also (I.63). Für $v \in \ker \beta_0$ ist $v+u-u_1 \in K$, und man erhält durch Einsetzen in (I.65)

$$\begin{aligned} \langle q_1 - \alpha_1(u), \beta_1(v - u_1) \rangle &\leq j(v + u - u_1) - j(u) - \langle q_0 - \alpha_0(u), \beta_0(v - u_1) \rangle \\ &\leq j(v - u_1). \end{aligned} \quad (\text{I.67})$$

Einsetzen von 0 und $2u_1$ in (I.67) liefert zusammen mit (I.51)

$$j(u_1) = \langle q_1 - \alpha_1(u), \beta_1(v - u_1) \rangle = \langle q_1 - \alpha_1(u), \beta_1(v - u) \rangle.$$

Aus (I.52) folgt wegen $u - u_1 \in \ker \beta_1$

$$j(u_1) - j(u) = j(u_1 - u + u) - j(u) \leq j(u - u_1) = 0.$$

Also gilt $j(u_1) = j(u)$ und damit schließlich (I.64). Es sei $\beta_0(v) \in G$ mit $v \in \ker \beta_1$, dann ist $w - v \in K \cap \ker \beta_1$, und es folgt aus (I.66)

$$\begin{aligned} \langle q_0 - \alpha_0(u), \beta_0(v) \rangle \\ = -(\langle q_0 - \alpha_0(u), \beta_0(w - v - u_0) \rangle + \langle q_0 - \alpha_0(u), \beta_0(u) - g \rangle) \geq 0. \end{aligned}$$

Da $\beta_0(\ker \beta_1) = U_0$ ist, folgt daraus $q_0 - \alpha_0(u) \in \Lambda_0$. Es sei $v \in \ker \beta_0$, dann folgt aus (I.64) und (I.67)

$$\begin{aligned} \pm \langle q_1 - \alpha_1(u), \beta_1(v) \rangle &= \langle q_1 - \alpha_1(u), \beta_1(\pm v - u_1) \rangle + \langle q_1 - \alpha_1(u), \beta_1(u) \rangle \\ &\leq j(\pm v + u - u_1) \\ &\leq j(\pm v) + j(u - u_1) \\ &= j(v). \end{aligned}$$

Wegen $\beta_1(\ker \beta_0) = U_1$ gilt $q_1 - \alpha_1(u) \in \Lambda_1$. Demnach gilt insgesamt (I.62). Es erfülle u die Bedingungen (I.43)-(I.46). Ist $v \in K$, dann ist $w - v \in G$, und damit gilt

$$\begin{aligned} &(f, v - u) + \langle q, \beta(v - u) \rangle \\ &= (A(u), v - u) + \langle q - \alpha(u), \beta(v - u) \rangle + \langle \alpha(u), \beta(v - u) \rangle \\ &= a(u, v - u) - \langle q_0 - \alpha_0(u), \beta_0(w - v) \rangle + \langle q_0 - \alpha_0(u), g - \beta_0(u) \rangle \\ &\quad + \langle q_1 - \alpha_1(u), \beta_1(v) \rangle - \langle q_1 - \alpha_1(u), \beta_1(u) \rangle \\ &\leq a(u, v - u) + j(v) - j(u). \end{aligned}$$

Also erfüllt u die Bedingung (I.31). □

Satz I.28 *Es seien $u \in \text{dom}(A)$, $\beta_0(\ker \beta_1) = U_0$ und $\beta_1(\ker \beta_0) = U_1$ sowie j wie in (I.59). Genau dann ist $(u, \lambda_0, \lambda_1)$ Lösung von (I.35)-(I.36), wenn $u \in K$ ist, u (I.31) oder (I.61)-(I.64) erfüllt, und wenn gilt:*

$$\lambda_0 = q_0 - \alpha_0(u) \quad (\text{I.68})$$

$$\lambda_1 = q_1 - \alpha_1(u). \quad (\text{I.69})$$

Beweis: Es sei $(u, \lambda_0, \lambda_1)$ Lösung von (I.35)-(I.36). Dann liefert Einsetzen von $(0, \lambda_1)$ und $(2\lambda_0, \lambda_1)$ in (I.36)

$$\langle \lambda_0, \beta_0(u) - g \rangle = 0. \quad (\text{I.70})$$

Daraus folgt, dass $\langle \mu_0, g - \beta_0(u) \rangle \geq 0$ für alle $\mu_0 \in \Lambda_0$ ist. Wegen $\Lambda_0 = G'$ folgt damit aus Satz A.6 $g - \beta_0(u) \in G$, also $u \in K$. Es sei $v \in K$ und $\beta_0(w) = g$, dann ist nach Subtrahieren von (I.35) und Anwenden von (I.36) für $\mu_1 \in \Lambda_1$

$$\begin{aligned} a(u, v - u) &= (f, v - u) + \langle q, \beta(v - u) \rangle - \langle \lambda_0, \beta_0(v - u) \rangle - \langle \lambda_1, \beta_1(v - u) \rangle \\ &= (f, v - u) + \langle q, \beta(v - u) \rangle + \langle \lambda_0, \beta_0(w - v) \rangle + \langle \lambda_0, \beta_0(u) - g \rangle \\ &\quad - \langle \lambda_1, \beta_1(v) \rangle + \langle \lambda_1, \beta_1(u) \rangle \\ &\geq (f, v - u) + \langle q, \beta(v - u) \rangle - j(v) + \langle \mu_1, \beta_1(u) \rangle. \end{aligned}$$

Also gilt

$$\begin{aligned} a(u, v - u) &\geq (f, v - u) + \langle q, \beta(v - u) \rangle - j(v) + \sup_{\mu_1 \in \Lambda_1} \langle \mu_1, \beta_1(u) \rangle \\ &= (f, v - u) + \langle q, \beta(v - u) \rangle - j(v) + j(u). \end{aligned}$$

Demnach gilt (I.31). Nach Satz I.27 ist dann auch das System (I.61)-(I.64) erfüllt. Für $v \in \ker \beta_1$ folgt dann

$$\begin{aligned} \langle q_0 - \alpha_0(u), \beta_0(v) \rangle &= \langle q - \alpha(u), \beta(v) \rangle \\ &= \langle q, \beta(v) \rangle - a(u, v) + (A(u), v) \\ &= \langle q, \alpha(v) \rangle - a(u, v) + (f, v) \\ &= \langle \lambda_0, \beta_0(v) \rangle. \end{aligned}$$

Analog zeigt man, dass für alle $v \in \ker \beta_0$ gilt:

$$\langle q_1 - \alpha_1(u), \beta_1(v) \rangle = \langle \lambda_1, \beta_1(v) \rangle.$$

Wegen $\beta_0(\ker \beta_1) = U_0$ und $\beta_1(\ker \beta_0) = U_1$ sind damit (I.68) und (I.69) erfüllt.

Wenn andererseits $u \in K$ ist und $(u, \lambda_0, \lambda_1)$ (I.61)-(I.64) sowie (I.68) und (I.69) erfüllt, dann folgt für $v \in V$

$$\begin{aligned} \langle f, v \rangle + \langle q, \beta(v) \rangle - \langle \lambda_0, \beta_0(v) \rangle - \langle \lambda_1, \beta_1(v) \rangle &= \langle A(u), v \rangle + \langle \alpha(u), \beta(v) \rangle \\ &= a(u, v). \end{aligned}$$

Demnach gilt (I.26). Wegen (I.63) ist $\langle \lambda_0, \beta_0(u) - g \rangle = 0$, und wegen (I.64) ist $\langle \lambda_1, \beta_1(u) \rangle = j(u)$. Damit gilt für $(\mu_0, \mu_1) \in \Lambda_0 \times \Lambda_1$

$$\begin{aligned} \langle \mu_0 - \lambda_0, \beta_0(u) - g \rangle + \langle \mu_1 - \lambda_1, \beta_1(u) \rangle \\ = -\langle \mu_0, g - \beta_0(u) \rangle + \langle \mu_1, \beta_1(u) \rangle - j(u) \leq 0, \end{aligned}$$

da $g - \beta_0(u) \in G$ ist. Also gilt auch (I.36). \square

Eine Vereinfachung der Menge Λ_1 ist in der folgenden Weise möglich:

Satz I.29 *Es sei U_1 dicht in \tilde{U}_1 enthalten sowie $p : \tilde{U}_1 \rightarrow \mathbb{R}_{\geq 0}$ sublinear mit*

$$\forall w \in \tilde{U}_1 : p(w) = p(-w) \tag{I.71}$$

$$\exists c > 0 : \forall w \in \tilde{U}_1 : p(w) \leq c\|w\| \tag{I.72}$$

$$\forall v \in V : j(v) = p(\beta_1(v)) \tag{I.73}$$

Zudem sei $\tilde{\Lambda}_1 := \{\tilde{\mu}_1 \in \tilde{U}_1' \mid \forall w \in \tilde{U}_1 : |\langle \tilde{\mu}_1, w \rangle| \leq p(w)\}$. Dann ist $\Lambda_1 = \tilde{\Lambda}_1$.

Beweis: Die Relation $\tilde{\Lambda}_1 \subset \Lambda_1$ ist klar. Nach dem Fortsetzungssatz von Hahn-Banach A.13 und wegen der Surjektivität von β_1 sowie (I.73) existiert zu jedem $\mu_1 \in \Lambda_1$ eine lineare Fortsetzung $\tilde{\mu}_1 : \tilde{U}_1 \rightarrow \mathbb{R}$ mit $\tilde{\mu}_1(w) \leq p(w)$ für alle $w \in \tilde{U}_1$. Aus (I.71) folgt $-\tilde{\mu}_1(w) = \tilde{\mu}_1(-w) \leq p(-w) = p(w)$ und damit $|\tilde{\mu}_1(w)| \leq p(w)$. Wegen (I.72) ist $\tilde{\mu}_1 \in \tilde{U}_1'$. Da U_1 dicht in \tilde{U}_1 enthalten ist, ist nach Satz A.14 $\tilde{\mu}_1$ eindeutig, also gilt auch die Relation $\Lambda_1 \subset \tilde{\Lambda}_1$. \square

Kapitel II

Modellierungen von Kontaktproblemen

Ziel dieses Kapitels ist die Formulierung von Kontaktproblemen als elliptische Minimierungsprobleme erster und zweiter Art.

Im Mittelpunkt stehen statische Kontaktprobleme vom Signorini-Typ und modellhafte Vereinfachungen hiervon, die das Deformationsverhalten elastischer Körper bei Starrkörperkontakt beschreiben. Hierzu werden sowohl reibungsfreie als auch reibungsbehaftete Formulierungen betrachtet und mit Hilfe der in Kapitel I dargestellten Resultate hinsichtlich der Existenz und Eindeutigkeit von Lösungen und hinsichtlich variationeller Sattelpunktformulierungen untersucht. Darüber hinaus werden auf der Grundlage von Abschnitt I.4 die zugehörigen Kräftebilanzgleichungen und entsprechende physikalische Interpretationen angegeben.

Exemplarisch wird in Abschnitt II.4 der fertigungstechnische Prozess des Roboter-gestützten Bandschleifens als reibungsbehaftetes Kontaktproblem studiert.

Neben den elastischen Kontaktproblemen vom Signorini-Typ werden Modellprobleme weiterer Problemklassen untersucht, die sich ebenfalls als Minimierungsprobleme erster und zweiter Art formulieren lassen. Hierzu gehören Hindernis-Probleme, Torsionsprobleme und Bingham-Fluid-Probleme. Die Ergebnisse aus Kapitel I lassen sich allerdings nur eingeschränkt auf diese Probleme anwenden.

Für weitere Details und eine entsprechende Ausweitung auch auf dynami-

sche Kontaktprobleme sei zum Beispiel auf [47], [55], [56], [68] und [72] verwiesen.

Den folgenden Untersuchungen ist eine kurze Darstellung der für die Formulierungen erforderlichen Funktionenräume vorangestellt.

II.1 Sobolevräume

Die in Kapitel I beschriebenen Zielfunktionale, die zur Minimierung von Energiegrößen eingesetzt werden, sind im Allgemeinen Integralmittelwerte bestimmter Differentialoperatoren. Adäquate Funktionenräume, die in diesem Zusammenhang Verwendung finden und den in Kapitel I aufgeführten Voraussetzungen für Existenz und Eindeutigkeit von Lösungen genügen, sind sogenannte Sobolevräume.

Im Folgenden werden die wesentlichen Notationen und Eigenschaften dieser Funktionenräume angegeben. Für weitere Details in diesem Kontext wird auf [1] verwiesen.

Ausgangspunkt ist der Raum $L^2(\Omega) := \mathcal{L}^2(\Omega)/\mathcal{N}(\Omega)$, wobei $\Omega \subset \mathbb{R}^k$ beschränkt, offen und zusammenhängend und $\mathcal{N}(\Omega) := \{f \in \mathcal{L}^2(\Omega) \mid \int_{\Omega} f^2 dx = 0\}$ ist. Im Weiteren wird in der Notation für $[f] \in L^2(\Omega)$ nicht mehr zwischen $[f]$ und f unterschieden. Der Raum $L^2(\Omega)$ ist vermöge des Skalarprodukts $(u, v)_0 := (u, v)_{0, \Omega} := \int_{\Omega} uv dx$ mit der zugehörigen Norm $\|u\|_0^2 := \|u\|_{0, \Omega}^2 := (u, u)_{0, \Omega}$ ein Hilbertraum.

Für den Rand $\Gamma := \partial\Omega$ ist der Hilbertraum $L^2(\Gamma)$ entsprechend mit einem auf Γ erweiterten Integral-Begriff unter geeigneten Glätte-Voraussetzungen zu definieren.¹ Das zugehörige Skalarprodukt wird mit $(\cdot, \cdot)_{0, \Gamma}$ und die Norm mit $\|\cdot\|_{0, \Gamma}$ bezeichnet.

Der klassische Differenzierbarkeitsbegriff wird in der Regel punktweise definiert und ist deshalb für L^2 -Funktionen zu restriktiv. Der für L^2 -Funktionen angemessene Differenzierbarkeitsbegriff basiert auf einer distributiven Übertragung der klassischen Differenzierbarkeit und führt auf den Begriff der schwachen Ableitung. Für einen Multiindex $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ mit $|\alpha| := \sum_{i=1}^n \alpha_i$ und für eine Funktion $\phi \in C^{|\alpha|}(\Omega)$ bezeichnet $\partial^\alpha \phi := \partial^{|\alpha|} / \partial^{\alpha_1} x_1 \cdots \partial^{\alpha_n} x_n \phi$ die partielle Ableitung nach α im klassischen Sinn. Die Funktion $u \in L^2(\Omega)$ hat die schwache partielle Ableitung $\partial^\alpha u \in L^2(\Omega)$, falls

$$\forall \phi \in C_0^\infty(\Omega) : (\phi, \partial^\alpha u)_0 = (-1)^{|\alpha|} (\partial^\alpha \phi, u)_0$$

¹vgl. S.114 in [1]

gilt. Die schwache partielle Ableitung ist wohl definiert und stimmt mit der klassischen partiellen Ableitung überein, falls $u \in C^{|\alpha|}(\Omega)$ ist.

Die Menge aller Funktionen aus $L^2(\Omega)$, die alle Ableitungen bis zu einem bestimmten Grad $m \in \mathbb{N}_0$ besitzen, werden in dem Raum $H^m(\Omega) := \{u \in L^2(\Omega) \mid \forall \alpha \in \mathbb{N}_0^n, |\alpha| \leq m : \partial^\alpha u \in L^2(\Omega)\}$ zusammengefasst, für den der Begriff Sobolevraum gebräuchlich ist.² Vermöge des Skalarprodukts $(u, v)_m := (u, v)_{m, \Omega} := \sum_{|\alpha| \leq m} (\partial^\alpha u, \partial^\alpha v)_{0, \Omega}$ ist der Raum $H^m(\Omega)$ ein Hilbertraum. Die zugehörige Norm ist definiert durch $\|u\|_m^2 := \|u\|_{m, \Omega}^2 := (u, u)_{m, \Omega}$.

Unter Verwendung geeigneter Koordinatensysteme von Γ wird der Hilbertraum $H^m(\Gamma)$ über den Raum $H^m(\mathbb{R}^{k-1})$ definiert. Zur genauen Angabe der Definition sei auf [76] verwiesen. Das dazugehörige Skalarprodukt sei mit $(u, v)_{m, \Gamma}$ bezeichnet, die Norm mit $\|u\|_{m, \Gamma}$.

Bedingungen, die auf dem Rand von Ω vorgeschrieben werden, erfordern den Begriff der Einschränkung auf Teilmengen des Randes von Ω . Da die Gleichheit von L^2 -Funktionen punktweise nur fast überall erfüllt sein muss, ist die punktweise Einschränkung einer L^2 -Funktion auf den Rand $\Gamma := \partial\Omega$ nicht sinnvoll. Es existiert jedoch stets genau eine stetige, lineare Abbildung $\gamma : H^1(\Omega) \rightarrow L^2(\Gamma)$, die mit der punktweisen Einschränkung für Funktionen aus $C^0(\overline{\Omega}) \cap H^1(\Omega)$ übereinstimmt:³

$$\forall u \in C^0(\overline{\Omega}) \cap H^1(\Omega) : u|_\Gamma = \gamma(u).$$

Die Abbildung γ wird auch Spuoperator genannt. Dieser ist nicht surjektiv, jedoch ist der Hilbertraum $H^{1/2}(\Gamma) := \gamma(H^1(\Omega))$ versehen mit der Norm

$$\|w\|_{1/2} := \inf_{\substack{v \in H^1(\Omega) \\ \gamma(v) = w}} \|v\|_1,$$

dicht in $L^2(\Gamma)$. Der Dualraum von $H^{1/2}(\Gamma)$ wird mit $H^{-1/2}(\Gamma)$ bezeichnet, die dazugehörige Norm mit $\|\cdot\|_{-1/2}$.

Eine besondere Bedeutung haben Sobolevräume, die auf Teilmengen des Randes verschwinden. Man definiert mit Hilfe des Spuoperators γ für eine abgeschlossene Teilmenge $\Gamma_0 \subset \Gamma$ mit $\int_{\Gamma_0} ds > 0$ die Räume $H^1(\Omega, \Gamma_0) := \{v \in H^1(\Omega) \mid \gamma(v)|_{\Gamma_0} = 0\}$ und $H_0^1(\Omega) := H^1(\Omega, \Gamma)$. Offenbar ist $\ker \gamma =$

²Für diese Menge ist ebenfalls die Bezeichnung $W^m(\Omega)$ verbreitet, vgl. Bemerkungen auf S.162 in [106].

³vgl. Th.5.22 in [1]. Vorausgesetzt wird, dass Ω die sogenannte Kegelbedingung erfüllt.

$H_0^1(\Omega)$. Wenn der Rand Γ hinreichend glatt ist, ist $H_0^1(\Omega)$ die Vervollständigung von $C_0^\infty(\Omega)$ in $L^2(\Omega)$ bzgl. der $\|\cdot\|_1$ -Norm⁴, also ist insbesondere $H_0^1(\Omega)$ dicht in $L^2(\Omega)$ enthalten. Der zu $H_0^1(\Omega)$ gehörige Dualraum wird mit $H^{-1}(\Omega)$ bezeichnet.

Darüber hinaus definiert man für eine zu Γ_0 disjunkte, offene Teilmenge $\Gamma_1 \subset \Gamma$ mit $\Gamma_0 \cup \Gamma_1 = \Gamma$ den Raum $H^{1/2}(\Gamma_1) := \gamma(H^1(\Omega, \Gamma_0))$, der als abgeschlossener Unterraum von $H^{1/2}(\Gamma)$ mit der Norm $\|\cdot\|_{1/2, \Gamma_1}$ ein Hilbertraum ist. Der Dualraum wird mit $H^{-1/2}(\Gamma_1)$ bezeichnet, die zugehörige Norm ist $\|\cdot\|_{-1/2, \Gamma_1}$. Der Raum $L^2(\Gamma_1) := \{v \in L^2(\Gamma) \mid v|_{\Gamma_0} = 0\}$ ist als abgeschlossener Unterraum mit dem von $L^2(\Gamma)$ induzierten Skalarprodukt ebenfalls ein Hilbertraum. Das dazugehörige Skalarprodukt wird mit $(\cdot, \cdot)_{0, \Gamma_1}$ bezeichnet. Der Raum $H^{1/2}(\Gamma_1)$ ist eine dichte Teilmenge von $L^2(\Gamma_1)$.

Werden die Räume $L^2(\Omega)$ und $L^2(\Gamma_1)$ mit ihren Dualräumen identifiziert, erhält man die jeweils dichten Inklusionen

$$\begin{aligned} H_0^1(\Omega) &\subset L^2(\Omega) = L^2(\Omega)' \subset H^{-1}(\Omega), \\ H^{1/2}(\Gamma_1) &\subset L^2(\Gamma_1) = L^2(\Gamma_1)' \subset H^{-1/2}(\Gamma_1). \end{aligned}$$

Ein wesentlicher Zusammenhang zwischen Funktionen aus $H^1(\Omega)$ und $H^{1/2}(\Gamma)$ wird durch die Greensche Formel hergestellt. Hierzu ist der Gradient $\nabla : H^1(\Omega) \rightarrow L^2(\Omega)^k$ für ein $v \in H^1(\Omega)$ definiert als $(\nabla v)_i := \partial v / \partial x_i$. Ferner ist der Divergenz-Operator $\operatorname{div} : H(\operatorname{div}, \Omega) \rightarrow L^2(\Omega)$ definiert über

$$\forall \phi \in C_0^\infty(\Omega) : (\operatorname{div} w, \phi)_0 = -(w, \nabla \phi)_0 \quad (\text{II.1})$$

mit $H(\operatorname{div}, \Omega) := \{w \in L^2(\Omega)^k \mid \operatorname{div} w \in L^2(\Omega)\}$ sowie $(w, v)_0 := (w_i, v_i)_0$ für $w, v \in L^2(\Omega)^k$. Für $w \in H^1(\Omega)^k$ ist $\operatorname{div} w = \partial w_i / \partial x_i$.⁵ Es gilt für $w \in H(\Omega)^k$ und $v \in H^1(\Omega)$:⁶

$$(w, \nabla v)_0 = (-\operatorname{div} w, v)_0 + (\gamma(w)n, \gamma(v))_{0, \Gamma}. \quad (\text{II.2})$$

Hierbei bezeichnet $n \in L^2(\Gamma)^k$ den äußeren Normalenvektor mit Länge 1 bzgl. Γ .

Der Gradient für mehrkomponentige Funktionen $\nabla : H^1(\Omega)^k \rightarrow L^2(\Omega)^{k \times k}$ ist für ein $v \in H^1(\Omega)^k$ definiert als $(\nabla v)_{ij} := \partial v_i / \partial x_j$, sowie der entsprechende Divergenz-Operator $\operatorname{div} : H(\operatorname{div}, \Omega) \rightarrow L^2(\Omega)^k$ analog zu (II.1) mit

⁴vgl. Satz 6.2.42 in [61]

⁵im Sinne der Einsteinschen Summenkonvention

⁶vgl. z.B. Th.A.29 in [89]

$H(\operatorname{div}, \Omega) := \{\tau \in L^2(\Omega)^{k \times k} \mid \operatorname{div} \tau \in L^2(\Omega)^k\}$ und $(\tau, v)_0 := (\tau_{ij}, v_{ij})_0$ für $\tau, v \in L^2(\Omega)^{k \times k}$. Ferner ist $(\operatorname{div} \tau)_i = \partial \tau_{ij} / \partial x_j$ für $\tau \in H^1(\Omega)^{3 \times 3}$. Zudem erhält man für $v \in H^1(\Omega)^k$ eine analoge Greensche Formel:

$$(\tau, \nabla v)_0 = (-\operatorname{div} \tau, v)_0 + (\gamma(\tau)n, \gamma(v))_{0, \Gamma}. \quad (\text{II.3})$$

Ist Γ hinreichend glatt, können Normal- und Tangentialräume bzgl. Γ_1 definiert werden. Hierzu beinhalte $t \in L^2(\Gamma)^{k \times (k-1)}$ die Tangentialvektoren bzgl. Γ , so dass $\kappa(x) := (n(x), t(x))$ für $x \in \Gamma$ ein orthonormales System ist. Die Abbildungen $\delta_n : H^1(\Omega, \Gamma_0)^k \rightarrow H_n^{1/2}(\Gamma_1) := \delta_n(H^1(\Omega, \Gamma_0)^k)$ und $\delta_t : H^1(\Omega, \Gamma_0)^k \rightarrow H_t^{1/2}(\Gamma_1) := \delta_t(H^1(\Omega, \Gamma_0)^k)$ seien definiert durch $\delta_n(v) := n^\top \gamma(v)$ und $\delta_t(v) := t^\top \gamma(v)$, und es sei $\delta := (\delta_n, \delta_t)^\top$. Vermöge der Abbildung $\tilde{\delta} : H^{1/2}(\Gamma_1)^k \rightarrow H_{nt}^{1/2}(\Gamma_1)$ mit $\tilde{\delta}(w) := \kappa^\top w$ gilt $H_{nt}^{1/2}(\Gamma_1) := H_n^{1/2}(\Gamma_1) \times H_t^{1/2}(\Gamma_1) \simeq H^{1/2}(\Gamma_1)^k$, wobei durch $\|w\|_{1/2, nt} := \|\tilde{\delta}^{-1}(w)\|_{1/2, \Gamma_1}$ die zugehörige Norm definiert ist. Die Räume $H_{nt}^{1/2}(\Gamma_1)$, $H_n^{1/2}(\Gamma_1)$ und $H_t^{1/2}(\Gamma_1)$ sind Hilberträume, wobei die Normen von $H_n^{1/2}(\Gamma_1)$ und $H_t^{1/2}(\Gamma_1)$ durch $\|\cdot\|_{1/2, n} := \|(\cdot, 0)\|_{1/2, nt}$ und $\|\cdot\|_{1/2, t} := \|(0, \cdot)\|_{1/2, nt}$ gegeben sind, und die Normen ihrer Dualräume $H_n^{-1/2}(\Gamma_1)$ und $H_t^{-1/2}(\Gamma_1)$ mit $\|\cdot\|_{-1/2, n}$ bzw. $\|\cdot\|_{-1/2, t}$ bezeichnet seien. Der Dualraum von $H_{nt}^{1/2}(\Gamma_1)$ sei $H_{nt}^{-1/2}(\Gamma_1)$ mit der Norm $\|\cdot\|_{-1/2, nt}$. Die Räume $H_n^{1/2}(\Gamma_1)$, $H_t^{1/2}(\Gamma_1)$ und $H_{nt}^{1/2}(\Gamma_1)$ sind dicht in $L^2(\Gamma_1)$, $L^2(\Gamma_1)^{k-1}$ bzw. $L^2(\Gamma_1)^k$ enthalten.⁷

II.2 Modellprobleme

Ziel dieses Abschnitts ist die Einführung von einfachen Modellproblemen, die jedoch repräsentativ für eine Vielzahl von elliptischen Minimierungsproblemen, insbesondere für die in Abschnitt II.3 aufgeführten Probleme sind. Hierzu sei $\Omega \subset \mathbb{R}^2$ beschränkt, offen und zusammenhängend. Für den Rand $\Gamma := \partial\Omega$ gelte $\Gamma = \bar{\Gamma}_0 \cup \bar{\Gamma}_1$ und $\int_{\Gamma_0} ds > 0$, wobei $\Gamma_0, \Gamma_1 \subset \Gamma$ disjunkte, offene Teilmengen seien. Ferner sei $f \in L^2(\Omega)$ und $q \in L^2(\Gamma_1)$.

Ein typisches Modellproblem für unrestringierte Minimierungsprobleme ist

$$\min_{v \in H^1(\Omega, \Gamma_0)} E(v) \quad (\text{II.4})$$

mit dem Zielfunktional

$$E(v) := \frac{1}{2}(\nabla v, \nabla v)_0 - (f, v)_0 - (q, \gamma(v))_{0, \Gamma_1}. \quad (\text{II.5})$$

⁷vgl. [64]

Die hier verwendete Bilinearform $a : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ mit $a(v, w) := (\nabla v, \nabla w)_0$ ist offensichtlich symmetrisch und wegen der Cauchy-Schwarzschen Ungleichung auch stetig. Aus der Poincaré-Friederichsschen Ungleichung⁸ folgt, dass a elliptisch ist. Damit ist durch (II.4) ein unrestringiertes elliptisches Minimierungsproblem erster Art gegeben. Nach Satz I.11 besitzt das Minimierungsproblem (II.4) genau eine Lösung u . Für diese ist notwendig und hinreichend:

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla u, \nabla v)_0 = (f, v)_0 + (q, \gamma(v))_{0, \Gamma_1}. \quad (\text{II.6})$$

Das Zielfunktional in (II.5) beschreibt in vereinfachter Form die potentielle Energie bei Durchbiegung einer Membran unter der Belastung f . Dementsprechend ist die Lösung u die Verschiebung im vom System angestrebten, kleinsten energetischen Zustand.

Der zu a assoziierte Operator ist $-\Delta$, wobei der Operator $\Delta := \operatorname{div} \nabla$ Laplace-Operator heißt.⁹ Im Folgenden sei α wie in Satz I.21.¹⁰ Falls $\nabla u \in H(\operatorname{div}, \Omega)$ ist, folgt damit aus Satz I.22: Genau dann ist u Lösung von (II.6), wenn gilt:

$$-\Delta u = f \quad (\text{II.7})$$

$$\alpha_0(u) = q \quad (\text{II.8})$$

Ist $u \in H^2(\Omega)$, folgt aus der Eindeutigkeit von α und der Greenschen Formel II.2, dass $\alpha_0(u) = \partial_n u := \gamma(\nabla u)n$ ist. Damit erhält man aus (II.7)-(II.8):

Satz II.1 *Es sei $u \in H^1(\Omega, \Gamma_0) \cap H^2(\Omega)$. Genau dann ist u Lösung von (II.6), wenn gilt:*

$$-\Delta u = f$$

$$\partial_n u = q.$$

II.2.1 Vereinfachte Signorini-Probleme

Das in diesem Abschnitt behandelte Modellproblem ist ein restringiertes elliptisches Minimierungsproblem erster Art. Das Zielfunktional E ist wie

⁸vgl. z.B. Th.1.5 in [25]

⁹vgl. z.B. Ch.6. Lemma 1.1 in [5]

¹⁰Mit den Bezeichnungen aus Abschnitt I.4 ist $V := H^1(\Omega, \Gamma_0)$, $W := L^2(\Omega)$, $U := H^{1/2}(\Gamma_1) \times \{0\}$ bzw. $U := \{0\} \times H^{1/2}(\Gamma_1)$ und $\beta := (\gamma, 0)$ bzw. $\beta := (0, \gamma)$.

in (II.5) definiert. Im Unterschied zu (II.4) kommen nun Restriktionen auf dem Rand Γ der Form

$$K := \{v \in H^1(\Omega, \Gamma_0) \mid \gamma(v) \geq g\} \quad (\text{II.9})$$

mit $g \in H^{1/2}(\Gamma_1)$ hinzu, wobei die Halbordnungen \leq bzw. \geq als fast überall auf Γ zu verstehen sind. Untersucht wird demnach die Bestimmung von

$$\min_{v \in K} E(v). \quad (\text{II.10})$$

Da K konvex und abgeschlossen ist, existiert nach Satz I.11 genau eine Lösung $u \in K$ von (II.10). Für diese ist notwendig und hinreichend:

$$\forall v \in K : (\nabla u, \nabla(v - u))_0 \geq (f, v - u)_0 + (q, \gamma(v - u))_{0, \Gamma_1}. \quad (\text{II.11})$$

Das Minimierungsproblem (II.10) ist ein vereinfachtes Modellproblem vom sogenannten Signorini-Typ.¹¹ Es beschreibt die Durchbiegung einer Membran bei Vorhandensein eines Hindernisses auf dem Rand.

Gilt $u \in K$ mit $\nabla u \in H(\text{div}, \Omega)$, dann ist u nach Satz I.23 genau dann Lösung von (II.11), wenn

$$-\Delta u = f \quad (\text{II.12})$$

$$q - \alpha_0(u) \in \Lambda_0 \quad (\text{II.13})$$

$$\langle q - \alpha_0(u), \gamma(u) - g \rangle = 0 \quad (\text{II.14})$$

erfüllt ist. Hierbei ist $\Lambda_0 := G'$ der duale Kegel von $G := \{v \in H^{1/2}(\Gamma_1) \mid v \leq 0\}$.

Im Fall, dass $u \in H^2(\Omega)$ ist, können die Bedingungen (II.12)-(II.14) konkretisiert werden. Hierzu sei $v \in L^2(\Gamma_1)$ mit $v \leq 0$. Da $H^{1/2}(\Gamma_1)$ dicht in $L^2(\Gamma_1)$ enthalten ist, existiert eine Folge $\{v_m\} \subset G$ mit $v_m \rightarrow v$. Damit ist $0 \leq \lim_{m \rightarrow \infty} (q - \partial_n u, v_m)_{0, \Gamma_1} = (q - \partial_n u, v)_{0, \Gamma_1}$, woraus $q - \partial_n u \leq 0$ folgt. Aus (II.14) erhält man $0 = (q - \partial_n u, \gamma(u) - g)_{0, \Gamma_1}$ und damit $(q - \partial_n u)(\gamma(u) - g) = 0$.

Satz II.2 *Es sei $u \in K \cap H^2(\Omega)$. Genau dann ist u Lösung von (II.11), wenn gilt:*

$$-\Delta u = f \quad (\text{II.15})$$

$$q - \partial_n u \leq 0 \quad (\text{II.16})$$

$$(q - \partial_n u)(\gamma(u) - g) = 0. \quad (\text{II.17})$$

¹¹vgl. Abschnitt II.3

Da $K = \{v \in H^1(\Omega, \Gamma_0) \mid g - \gamma(v) \in G\}$ und G konvex und abgeschlossen ist, ist das Funktional $\Phi_0 : H^1(\Omega, \Gamma_0) \times G' \rightarrow \mathbb{R}$ mit

$$\Phi_0(v, \mu_0) := \langle \mu_0, \gamma(v) - g \rangle \quad (\text{II.18})$$

nach Satz I.3 ein Multiplikatorfunktional bzgl. K . Damit erhält man mit Hilfe von Satz I.12 eine variationelle Sattelpunktformulierung für die Lösung von (II.10): $u \in H^1(\Omega, \Gamma_0)$ ist Lösung von (II.10), wenn ein $\lambda_0 \in \Lambda_0$ existiert, so dass

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla u, \nabla v)_0 = (f, v)_0 + (q, \gamma(v))_{0, \Gamma_1} - \langle \lambda_0, \gamma(v) \rangle \quad (\text{II.19})$$

$$\forall \mu_0 \in \Lambda_0 : \langle \mu_0 - \lambda_0, \gamma(u) - g \rangle \leq 0 \quad (\text{II.20})$$

gilt. Da $\text{rg}(\gamma|_{H^1(\Omega, \Gamma_0)}) = H^{1/2}(\Gamma_1)$ abgeschlossen in $H^{1/2}(\Gamma_1)$ ist, besitzt das System (II.19)-(II.20) nach Satz I.18 eine eindeutige Lösung.

Mit Hilfe von Satz I.24 kann der Lagrangesche Multiplikator λ_0 explizit angegeben werden: Ist $u \in H^1(\Omega, \Gamma_0) \cap H^2(\Omega)$, dann ist (u, λ) genau dann Lösung von (II.19)-(II.20), wenn u (II.11) oder (II.15)-(II.17) erfüllt und wenn $\lambda_0 = q - \partial_n u$ ist.

Zur Veranschaulichung der in Satz II.2 formulierten Eigenschaften einer Lösung u sei $\Omega := (-1, 1)^2$ und $\Gamma_1 := (-1, 1) \times \{-1\} \cup \{-1\} \times (-1, 1)$. Außerdem sei $f := -1$ sowie $q(x_0, x_1) := -\frac{1}{4}x_1^3$. Die Durchbiegung u einer Membran ist in Abbildung II.1 dargestellt. Links ist die Membran ohne Restriktionen zu sehen, rechts die Membran mit dem Hindernis $g(x_0, x_1) := -x_0^2$ auf dem Rand (untere Linie).



Abb. II.1: Durchbiegung einer Membran ohne Randrestriktionen, mit Hindernis.

Zusätzlich ist für den Hindernisfall der Lagrangesche Multiplikator $\lambda_0 = q - \partial_n u$ (obere Linie) eingetragen. Man erkennt deutlich am rechten Bild,

dass die Membran auf dem Hindernis aufliegt, und dass λ_0 innerhalb dieser Kontaktzone negativ ist und außerhalb verschwindet.

II.2.2 Idealisierte Reibungsprobleme

Ein unrestringiertes elliptisches Minimierungsproblem zweiter Art ist das Modellproblem

$$\min_{v \in H^1(\Omega, \Gamma_0)} H(v) + j(v). \quad (\text{II.21})$$

Hierbei setzt man $H(v) := \frac{1}{2}(\nabla v, \nabla v)_0 - (f, v)_0 - (q, \gamma(v))_{0, \Gamma_1}$ und $j(v) := (s, |\gamma(v)|)_{0, \Gamma_1}$ mit $s \in L^2(\Gamma_1)$ und $s \geq 0$.¹²

Da j offenbar ein konvexes und stetiges Funktional ist, hat das Minimierungsproblem (II.21) nach Satz I.13 genau eine Lösung $u \in H^1(\Omega, \Gamma_0)$. Hierfür ist notwendig und hinreichend:

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma_0) : (\nabla u, \nabla(v - u))_0 + (s, |\gamma(v)|)_{0, \Gamma_1} - (s, |\gamma(u)|)_{0, \Gamma_1} \\ \geq (f, v - u)_0 + (q, \gamma(v - u))_{0, \Gamma_1}. \end{aligned} \quad (\text{II.22})$$

Das Minimierungsproblem (II.21) ist ein stark idealisiertes Reibungsproblem.¹³

Ist $\nabla u \in H(\text{div}, \Omega)$, folgt aus Satz I.25, dass u genau dann Lösung von (II.22) ist, wenn gilt:

$$-\Delta u = f \quad (\text{II.23})$$

$$q - \alpha_1(u) \in \Lambda_1 \quad (\text{II.24})$$

$$\langle q - \alpha_1(u), \gamma(u) \rangle = (s, |\gamma(u)|)_{0, \Gamma_1}, \quad (\text{II.25})$$

wobei $\Lambda_1 := \{\mu_1 \in H^{-1/2}(\Gamma_1) \mid \forall w \in H^{1/2}(\Gamma_1) : |\langle \mu_1, w \rangle| \leq (s, |w|)_{0, \Gamma_1}\}$ ist.

Da $H^{1/2}(\Gamma_1)$ dicht in $L^2(\Gamma_1)$ enthalten ist, ist nach Satz I.29 $\Lambda_1 = \{\mu_1 \in L^2(\Gamma_1) \mid \forall w \in L^2(\Gamma_1) : |\langle \mu_1, w \rangle| \leq (s, |w|)_0\}$. Für $\mu_1 \in \Lambda_1$ und $w \in L^2(\Gamma_1)$ mit $w \geq 0$ folgt damit $(\pm \mu_1 - s, w) \leq 0$. Also ist $|\mu_1| \leq s$. Andererseits folgt aus $|\mu_1| \leq s$ unmittelbar $|\langle \mu_1, w \rangle| \leq (s, |w|)$ für $w \in L^2(\Gamma_1)$. Demnach ist

$$\Lambda_1 = \{\mu_1 \in L^2(\Gamma_1) \mid |\mu_1| \leq s\}. \quad (\text{II.26})$$

¹²Für eine (mehrkomponentige) L^2 -Funktion sind die L^2 -Funktionen $|v|$ und $\text{sign}(v)$ definiert als $|v| := \sqrt{v^\top v}$ und $\text{sign}(v) := |v|^{-1}v$ für $v(x) \neq 0$ und $\text{sign}(v) := 0$ sonst.

¹³vgl. Abschnitt II.3

Ist $u \in H^2(\Omega)$, folgt aus (II.25), dass $(s - \text{sign}(\gamma(u))(q - \partial_n u), |\gamma(u)|)_{0, \Gamma_1} = 0$ ist. Da wegen (II.24) $q - \partial_n u \in \Lambda_1$ ist, folgt aus (II.26) $0 = (s - \text{sign}(\gamma(u))(q - \partial_n u))|\gamma(u)| = s|\gamma(u)| - \gamma(u)(q - \partial_n u)$ und damit

$$(q - \partial_n u)\gamma(u) = s|\gamma(u)|. \quad (\text{II.27})$$

Lemma II.1 *Ist $|q - \partial_n u| \leq s$ und $s > 0$, dann ist die Bedingung (II.27) äquivalent zu*

$$\begin{aligned} |q - \partial_n u| < s &\Rightarrow \gamma(u) = 0 \\ (q - \partial_n u) = s &\Rightarrow \gamma(u) \geq 0 \\ -(q - \partial_n u) = s &\Rightarrow \gamma(u) \leq 0. \end{aligned}$$

Beweis: trivial. □

Zusammen mit Lemma II.1 erhält man schließlich aus (II.23)-(II.25):

Satz II.3 *Es sei $u \in H(\Omega, \Gamma_0) \cap H^2(\Omega)$ und $s > 0$. Genau dann ist u Lösung von (II.22), wenn gilt:*

$$-\Delta u = f \quad (\text{II.28})$$

$$|q - \partial_n u| \leq s \quad (\text{II.29})$$

$$|q - \partial_n u| < s \Rightarrow \gamma(u) = 0 \quad (\text{II.30})$$

$$(q - \partial_n u) = s \Rightarrow \gamma(u) \geq 0 \quad (\text{II.31})$$

$$-(q - \partial_n u) = s \Rightarrow \gamma(u) \leq 0. \quad (\text{II.32})$$

Bemerkung II.1 Wird in Satz II.3 $s = 0$ auf $\Gamma_{1,0} \subset \Gamma_1$ sowie $s > 0$ auf $\Gamma_{1,1} \subset \Gamma_1$ vorausgesetzt, so muss das System (II.28)-(II.32) modifiziert werden: Zu ergänzen ist die Bedingung $\partial_n u = q$ auf $\Gamma_{1,0}$, zudem sind die Bedingungen (II.29)-(II.32) auf $\Gamma_{1,1}$ zu beschränken.

Zur Gewinnung einer variationellen Sattelpunktformulierung ist ein Multiplikatorfunktional bzgl. j herzuleiten: Für $\mu_1 \in \Lambda_1$ ist offenbar auch $-\mu_1 \in \Lambda_1$, woraus für $v \in H^1(\Omega, \Gamma_0)$ unmittelbar folgt, dass $j(v) = (s, |\gamma(v)|)_{0, \Gamma_1} \geq \sup_{\mu_1 \in \Lambda_1} (\mu_1, \gamma(v))_{0, \Gamma_1}$ ist. Weil $|\text{sign}(\gamma(v))s| = s$ ist, gilt andererseits $j(v) = (s, |\gamma(v)|)_{0, \Gamma_1} = (\text{sign}(\gamma(v))s, \gamma(v))_{0, \Gamma_1} \leq \sup_{\mu_1 \in \Lambda_1} (\mu_1, \gamma(v))_{0, \Gamma_1}$. Demnach erfüllt j das Kriterium (I.59), ferner ist $\Phi_1 : H^1(\Omega, \Gamma_0) \times \Lambda_1 \rightarrow \mathbb{R}$ mit $\Phi_1(v, \mu_1) := (\mu_1, \gamma(v))_{0, \Gamma_1}$ ein Multiplikatorfunktional bzgl. j .

Nach Satz I.14 ist $u \in H^1(\Omega, \Gamma_0)$ Lösung des Minimierungsproblems (II.21),

wenn ein $\lambda_1 \in \Lambda_1$ existiert, so dass

$$\forall v \in H^1(\Omega, \Gamma_0) : \quad (\nabla u, \nabla v)_0 = (f, v)_0 + (q, \gamma(v))_{0, \Gamma_1} - (\lambda_1, \gamma(v))_{0, \Gamma_1} \quad (\text{II.33})$$

$$\forall \mu_1 \in \Lambda_1 : (\mu_1 - \lambda_1, \gamma(u))_{0, \Gamma_1} \leq 0 \quad (\text{II.34})$$

erfüllt ist. Da Λ_1 beschränkt ist und γ surjektiv auf $H^{1/2}(\Gamma_1)$ abbildet, existiert nach Satz I.19 eine eindeutige Lösung $(u, \lambda_1) \in H^1(\Omega, \Gamma_0) \times \Lambda_1$ von System (II.33)-(II.34).

Ist $u \in H^1(\Omega, \Gamma_0) \cap H^2(\Omega)$, dann ist (u, λ_1) nach Satz I.26 genau dann Lösung von (II.33)-(II.34), wenn u (II.22) oder (II.28)-(II.32) erfüllt und $\lambda_1 = q - \partial_n u$ ist.

In Abbildung II.2 ist die Durchbiegung u des idealisierten Reibungsproblems mit $s(x_0, x_1) := 1 - x_0^2$ und mit wie in Abschnitt II.2.1 gewählten Größen zu sehen. Abgebildet sind die Funktion s (obere Linie) und der Lagrangesche Multiplikator $\lambda_1 = q - \partial_n u$ (untere Linie). Erkennbar ist die in Satz II.3 (bzw. Bemerkung II.1) geforderte Eigenschaft der Lösung, dass in den Bereichen, in denen $|q - \partial_n u|$ echt kleiner als s ist, die Spur von u verschwinden muss.

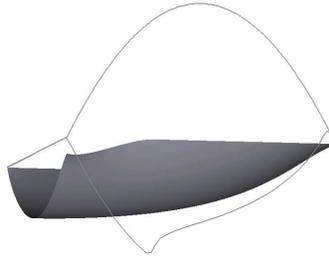


Abb. II.2: Membran beim idealisierten Reibungsproblem.

II.3 Linear-elastische Kontaktprobleme

Für viele Ingenieur-technische Probleme ist die Modellierung über die lineare Elastizitätstheorie eine angemessene Form der Modellbildung. Die lineare Elastizitätstheorie beschreibt das Deformationsverhalten von Festkörper-

kontinua mit Hilfe einer linearisierten Spannungs-Dehnungsbeziehung, die in Abhängigkeit von Verschiebungsgrößen ausgedrückt wird.

Das Ziel dieses Abschnitts besteht darin, Modelle von Festkörperkontakt mit (trockener) Reibung auf der Grundlage der linearen Elastizitätstheorie zu formulieren und in den von Kapitel I geschaffenen Zusammenhang zu stellen.¹⁴ Insbesondere sollen die in Abschnitt I.4 beschriebenen äquivalenten Darstellungsformen genutzt werden, um geeignete physikalische Interpretationen hinsichtlich der im System auftretenden Kräftebilanzen anzugeben. Eine konkrete Anwendung der in diesem Abschnitt dargestellten Modellierungsformen wird in Abschnitt II.4 beschrieben.

Betrachtet wird ein Körper als materielles Kontinuum mit der räumlichen Einbettung $\bar{\Omega} \subset \mathbb{R}^3$, der an einem Randsegment $\Gamma_0 \subset \Gamma$ eingespannt ist und auf den volumen- und oberflächenbezogene Kräfte $f \in L^2(\Omega)^3$ bzw. $q = (q_n, q_t)^\top \in L^2(\Gamma_1) \times L^2(\Gamma_1)^2$ wirken, wobei q_n und q_t die Kraftanteile in normaler und tangentialer Richtung bezeichnen.¹⁵ Die auf den Körper einwirkenden Kräfte verursachen eine Deformation des Körpers, so dass jeder Massepunkt $x \in \Omega$ auf eine neue Position $\phi(x) \in \mathbb{R}^3$ verschoben wird. Die Funktion v mit $v(x) := \phi(x) - x$ heißt in diesem Zusammenhang Verschiebungsfunktion.

Die Abweichung von $\nabla\phi^\top\nabla\phi$ gegenüber der Einheitsmatrix kann als Maß für die Verzerrung oder Dehnung und damit als Maß für die Deformation des Körpers betrachtet werden.

Werden kleine Verschiebungen vorausgesetzt, so können Terme höherer Ordnung vernachlässigt werden, so dass die Dehnungen für eine Verschiebung $v \in H^1(\Omega, \Gamma_0)$ durch den Operator $\varepsilon : H^1(\Omega)^3 \rightarrow L^2(\Omega)_{\text{sym}}^{3 \times 3} := \{\tau \in L^2(\Omega)^{3 \times 3} \mid \tau_{ij} = \tau_{ji}\}$ mit $\varepsilon(v) := \frac{1}{2}(\nabla v + \nabla v^\top)$ wiedergegeben werden. Der Operator ε wird linearisierter Greenscher Dehnungs- oder Verzerrungstensor genannt.

Die potentielle Energie $E(v)$ eines Körpers nach Erleiden einer Deformation um die Verschiebung v berechnet sich aus

$$E(v) := \frac{1}{2}(\sigma(v), \varepsilon(v))_0 - (f, v)_0 - (q, \delta(v))_{0, \Gamma_1},$$

wobei δ der in Abschnitt II.1 beschriebene Operator ist. Das hierdurch definierte Funktional $E : H^1(\Omega, \Gamma_0) \rightarrow \mathbb{R}$ heißt Energiefunktional. Der darin

¹⁴Darstellung und Notation orientieren sich an den in Abschnitt VI.3 aus [25] beschriebenen Ausführungen bezüglich der linearen Elastizitätstheorie.

¹⁵Die Eigenschaften der Mengen Ω , Γ_0 und Γ_1 sind wie in Abschnitt II.1 gewählt.

auf tretende Operator $\sigma : H^1(\Omega)^3 \rightarrow L^2(\Omega)^{3 \times 3}$ ist der sogenannte Spannungstensor. Das wesentliche Merkmal der linearen Elastizitätstheorie besteht in der Annahme, dass die Spannungen in linearer Weise von den Dehnungen abhängen. Dieser Sachverhalt wird durch die Vorgabe

$$\sigma(v)_{ij} := \mathcal{C}_{ijkl} \varepsilon(v)_{kl} \quad (\text{II.35})$$

ausgedrückt, wobei $\mathcal{C}_{ijkl} \in L^\infty(\Omega)$ mit $\mathcal{C}_{ijkl} = \mathcal{C}_{jikl} = \mathcal{C}_{klij}$ gelten soll. Zudem existiere ein $\kappa > 0$, so dass gilt:

$$\forall \tau \in L^2(\Omega)_{\text{sym}}^{3 \times 3} : \mathcal{C}_{ijkl} \tau_{ij} \tau_{kl} \geq \kappa \tau_{ij}^2. \quad (\text{II.36})$$

Die Beziehung (II.35) heißt lineares Materialgesetz. Ein wichtiges lineares Materialgesetz ist das sogenannte Hookesche Gesetz, bei dem

$$\mathcal{C}_{ijkl} := \frac{E}{(1+\nu)(1-2\nu)} \cdot \begin{cases} 1-\nu, & i=j=k=l \\ \nu, & i=j, k=l, k \neq i \\ 1-2\nu, & i=k, j=l \\ 0, & \text{sonst} \end{cases} \quad (\text{II.37})$$

gesetzt wird. Die Materialkenngrößen $E \in \mathbb{R}_+$ und $\nu \in (0, 0.5)$ werden Elastizitätsmodul und Querkontraktions- oder Poissonzahl genannt. Mit den in (II.37) angegebenen Koeffizienten \mathcal{C}_{ijkl} lässt sich der Spannungstensor schreiben als

$$\sigma(v) := \frac{E}{1+\nu} \left(\varepsilon(v) + \frac{\nu}{1-2\nu} \text{tr} \varepsilon(v) I \right).$$

Für weitere Details zur physikalischen Motivation der hier eingeführten Größen sei zum Beispiel auf [25] oder [78] verwiesen.

Setzt man nun die Gültigkeit des Minimalprinzips voraus, so wird die Deformation des Körpers unter der Wirkung der gegebenen volumen- und oberflächenbezogenen Kräfte durch diejenige Verschiebungsfunktion u beschrieben, die den energetisch günstigsten Zustand sichert. Mit anderen Worten: Die Betrachtung der Deformation eines Körpers unter Wirkung gegebener Kräfte führt auf das Minimierungsproblem: Gesucht ist eine Verschiebungsfunktion $u \in H^1(\Omega, \Gamma_0)$, so dass gilt:

$$E(u) = \min_{v \in H^1(\Omega, \Gamma_0)} E(v). \quad (\text{II.38})$$

Offenbar ist $a : H^1(\Omega, \Gamma_0)^3 \times H^1(\Omega, \Gamma_0)^3 \rightarrow \mathbb{R}$ mit $a(v, w) := (\sigma(v), \varepsilon(w))_0$ eine symmetrische und stetige Bilinearform. Die Elliptizität von a folgt aus

(II.36) und aus dem 2. Kornischen Lemma¹⁶. Demnach besitzt das Minimierungsproblem (II.38) nach Satz I.11 genau eine Lösung u . Für diese ist notwendig und hinreichend:

$$\forall v \in H^1(\Omega, \Gamma_0)^3 : (\sigma(u), \varepsilon(v))_0 = (f, v)_0 + (q, \delta(v))_{0, \Gamma_1}. \quad (\text{II.39})$$

Der zu a assoziierte Operator ist $-\operatorname{div} \sigma$. Im Folgenden sei α wie in Satz I.21 definiert.¹⁷ Ist $\sigma(u) \in H(\operatorname{div}, \Omega)$, folgt aus Satz I.22: Genau dann ist u Lösung von (II.39), wenn gilt:

$$-\operatorname{div} \sigma(u) = f \quad (\text{II.40})$$

$$\alpha(u) = q. \quad (\text{II.41})$$

Für $\sigma(u) \in H^1(\Omega)^{3 \times 3}$ sei $\sigma_n(u) := \gamma(\sigma(u))n$. Da für $v \in H^1(\Omega)$

$$\begin{aligned} (\sigma_n(u), \gamma(v))_{0, \Gamma_1} &= (\sigma_n(u), n\delta_n(v) + t\delta_t(v))_{0, \Gamma_1} \\ &= (n^\top \sigma_n(u), \delta_n(v))_{0, \Gamma_1} + (t^\top \sigma_n(u), \delta_t(v))_{0, \Gamma_1} \end{aligned}$$

ist, folgt aus der Eindeutigkeit von α und der Greenschen Formel (II.3), dass $\sigma_{nn}(u) := \alpha_0(u) = n^\top \sigma_n(u)$ und $\sigma_{nt}(u) := \alpha_1(u) = t^\top \sigma_n(u)$ ist. Die Ausdrücke $\sigma_{nn}(u)$ und $\sigma_{nt}(u)$ repräsentieren die im System auftretenden Normal- bzw. Tangentialspannungen. Aus (II.40)-(II.41) erhält man damit:

Satz II.4 *Es sei $u \in H^1(\Omega, \Gamma_0)$ mit $\sigma(u) \in H(\Omega)^{3 \times 3}$. Genau dann ist u Lösung von (II.6), wenn gilt:*

$$-\operatorname{div} \sigma(u) = f \quad (\text{II.42})$$

$$\sigma_{nn}(u) = q_n \quad (\text{II.43})$$

$$\sigma_{nt}(u) = q_t. \quad (\text{II.44})$$

Im Wesentlichen besagt Satz II.4, dass die Reaktionskräfte im Körper mit den auf den Körper einwirkenden Kräften im Gleichgewicht sind. Die Gleichungen (II.42)-(II.44) heißen deshalb auch Gleichgewichtsbedingungen.

Im Gegensatz zu der in diesem Abschnitt dargestellten Vorgehensweise beginnt der Ingenieur-technische Zugang zur Modellierung in der Regel mit der Forderung, dass die Gleichgewichtsbedingungen bezüglich der einwirkenden Kräfte und der Reaktionskräfte erfüllt sind. Anschließend wird über die kinematische Beziehung für ε und über die Materialgesetzbeziehung für σ das Differentialgleichungssystem (II.42)-(II.44) formuliert.

¹⁶vgl. Th 3.3 in [25]

¹⁷Mit den Bezeichnungen aus Abschnitt I.4 ist $V := H^1(\Omega, \Gamma_0)^3$, $W := L^2(\Omega)^3$, $U := H_n^{1/2}(\Gamma_1) \times H_t^{1/2}(\Gamma_1)$ und $\beta := \delta$.

II.3.1 Reibungsfreie Kontaktprobleme

Ist die Deformation des Körpers durch das Vorhandensein eines starren Hindernisses eingeschränkt, so kann dies durch Restriktionen an die Verschiebungsfunktion ausgedrückt werden. Hierzu sei der Einfachheit halber das Oberflächensegment des Hindernisses, das möglicherweise in Kontakt mit dem Körper in Γ_1 geraten könnte, mit $\Upsilon \subset \mathbb{R}^3$ bezeichnet. Ferner seien Υ und Γ_1 durch hinreichend glatte Funktionen $\psi, \varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$ parametrisierbar:

$$\begin{aligned}\Upsilon &= \{(x_1, x_2, \psi(x_1, x_2)) \mid (x_1, x_2) \in \mathbb{R}^2\} \\ \Gamma_1 &= \{(x_1, x_2, \varphi(x_1, x_2)) \mid (x_1, x_2) \in \mathbb{R}^2\}.\end{aligned}$$

Die Hindernisbedingung, die den Körper als stets unterhalb des Hindernisses befindlich verlangt, kann damit für eine Verschiebungsfunktion v in der folgenden Form angegeben werden:

$$\begin{aligned}\varphi(x_1, x_2) + v_3(x_1, x_2, \varphi(x_1, x_2)) \\ \leq \psi(x_1 + v_1(x_1, x_2, \varphi_1(x_1, x_2)), x_2 + v_2(x_1, x_2, \varphi(x_1, x_2)))\end{aligned}\quad (\text{II.45})$$

Soll sich der Körper oberhalb des Hindernisses befinden, ist in (II.45) \geq zu verwenden.¹⁸

Die Bedingung (II.45) ist im Allgemeinen nicht linear. Zur Formulierung einer einfachen Restriktionsmenge K , die insbesondere mit einem Multiplikatorfunktional wie in Abschnitt I.3 erfasst werden kann, ist eine geeignete Linearisierung von (II.45) notwendig. In Ch. 2, [72] wird eine entsprechende Herleitung angegeben, die auf eine näherungsweise Formulierung von (II.45) in der Gestalt $g - \delta_n(v) \geq 0$ führt. Hierbei ist g mit

$$g(x_1, x_2, x_3) = \frac{\psi(x_1, x_2) - \varphi(x_1, x_2)}{\sqrt{1 + \left(\frac{\partial \varphi}{\partial x_1}(x_1, x_2)\right)^2 + \left(\frac{\partial \varphi}{\partial x_2}(x_1, x_2)\right)^2}}$$

die sogenannte Gap-Funktion von Υ und Γ_1 . Die Menge der zulässigen Verschiebungen ist demnach $K := \{v \in H^1(\Omega, \Gamma_0)^3 \mid g - \delta_n(v) \geq 0\}$. Sind φ und ψ hinreichend glatt, kann im Folgenden vorausgesetzt werden, dass $g \in H_n^{1/2}(\Gamma_1)$ ist.

¹⁸Für andere Hinderniskonstellationen sind die Variablen entsprechend zu vertauschen.

Das zu betrachtende Minimierungsproblem lautet: Finde ein $u \in K$, so dass gilt:

$$E(u) = \min_{v \in K} E(v). \quad (\text{II.46})$$

Da K konvex und abgeschlossen ist, besitzt (II.46) nach Satz I.11 genau eine Lösung $u \in K$. Für diese ist notwendig und hinreichend:

$$\forall v \in K : (\sigma(u), \varepsilon(v - u))_0 \geq (f, v - u)_0 + (q, \delta(v - u))_{0, \Gamma_1}. \quad (\text{II.47})$$

Gilt $u \in K$ mit $\sigma(u) \in H(\text{div}, \Omega)$, folgt wegen $\delta_t(\ker \delta_n) = H_t^{1/2}(\Gamma_1)$ aus Satz I.23: Genau dann ist u Lösung von (II.47), wenn gilt:

$$-\text{div } \sigma(u) = f \quad (\text{II.48})$$

$$q_n - \alpha_0(u) \in \Lambda_n \quad (\text{II.49})$$

$$\langle q_n - \alpha_0(u), \delta_n(u) - g \rangle = 0 \quad (\text{II.50})$$

$$\alpha_1(u) = q_t. \quad (\text{II.51})$$

Hierbei ist $\Lambda_n := G'$ der duale Kegel von $G := \{v \in H_n^{1/2}(\Gamma_1) \mid v \geq 0\}$. Analog zur Herleitung von Satz II.2 erhält man damit aus (II.48)-(II.51):

Satz II.5 *Es sei $u \in K$ und $\sigma(u) \in H^1(\Omega)^{3 \times 3}$. Genau dann ist u Lösung von (II.47), wenn gilt:*

$$-\text{div } \sigma(u) = f \quad (\text{II.52})$$

$$q_n - \sigma_{nn}(u) \geq 0 \quad (\text{II.53})$$

$$(q_n - \sigma_{nn}(u))(\delta_n(u) - g) = 0 \quad (\text{II.54})$$

$$\sigma_{nt}(u) = q_t. \quad (\text{II.55})$$

Die Gleichungen (II.52)-(II.55) aus Satz II.5 sind in der folgenden Weise zu interpretieren: Die Gleichungen (II.52) und (II.55) besagen, dass die Volumenkräfte und die tangentialen Oberflächenkräfte mit den inneren Reaktionskräften bzw. mit den Tangentialspannungen im Gleichgewicht sind. Die Bedingung (II.54) bedeutet, dass Kontakt vorhanden ist, oder Gleichgewicht zwischen den normalen Oberflächenkräften und den Normalspannungen besteht. Die Vorzeichenbedingung (II.53) bedeutet, dass die normalen Kontaktkräfte, die hier definiert sind als $\sigma_{nn}(u) - q_n$, höchstens Druckkräfte, aber keine Zugkräfte sind.

Durch die Verwendung des Multiplikatorfunktionals¹⁹ $\Phi_n : H^1(\Omega, \Gamma_0)^3 \times G' \rightarrow \mathbb{R}$, $\Phi_n(v, \mu_n) := \langle \mu_n, \delta_n(v) - g \rangle$ erhält man aus Satz I.12: $u \in$

¹⁹vgl. (II.18)

$H^1(\Omega, \Gamma_0)^3$ ist Lösung von (II.46), wenn ein $\lambda_n \in \Lambda_n$ existiert, so dass

$$\forall v \in H^1(\Omega, \Gamma_0)^3 : \quad (\sigma(u), \varepsilon(v))_0 = (f, v)_0 + (q, \delta(v))_{0, \Gamma_1} - \langle \lambda_n, \delta_n(v) \rangle \quad (\text{II.56})$$

$$\forall \mu_n \in \Lambda_n : \langle \mu_n - \lambda_n, \delta_n(u) - g \rangle \leq 0 \quad (\text{II.57})$$

gilt. Da $\text{rg}(\delta_n) = H_n^{1/2}(\Gamma_1)$ abgeschlossen in $H_n^{1/2}(\Gamma_1)$ ist, besitzt das System (II.56)-(II.57) nach Satz I.18 genau eine Lösung.

Ist $u \in H^1(\Omega, \Gamma_0)$ mit $\sigma(u) \in H^1(\Omega)^{3 \times 3}$, folgt aus Satz I.24, dass (u, λ_n) genau dann Lösung von (II.56)-(II.57) ist, wenn $u \in K$ ist, u (II.47) oder (II.52)-(II.55) erfüllt und wenn $\lambda_n = q_n - \sigma_{nn}(u)$ ist. Damit erhält der Lagrangesche Multiplikator auch eine physikalisch interpretierbare Eigenschaft: Er repräsentiert die negativen normalen Kontaktkräfte.

II.3.2 Reibungsprobleme mit vorgegebener Normalspannung

Das Minimierungsproblem (II.46) ist ein Modell für reibungsfreien Kontakt, da keinerlei Reibungsnebenbedingungen berücksichtigt werden.

Reibung liegt vor, wenn durch einen Reibungswiderstand ein Randpunkt daran gehindert wird, eine andere Position in tangentialer Richtung einzunehmen. Die tangentiale Verschiebung soll erst dann von null verschieden sein, wenn die tangentialen Kontaktkräfte, hier definiert als $\sigma_{nt}(u) - q_t$, vom Betrage her dem Reibungswiderstand entsprechen. Dieser Sachverhalt kann durch das folgende System ausgedrückt werden:

$$|q_t - \sigma_{nt}(u)| \leq s \quad (\text{II.58})$$

$$|q_t - \sigma_{nt}(u)| < s \Rightarrow \delta_t(u) = 0 \quad (\text{II.59})$$

$$|q_t - \sigma_{nt}(u)| = s \Rightarrow \exists \zeta \in \mathbb{R}_{\geq 0} : \delta_t(u) = \zeta(q_t - \sigma_{nt}(u)). \quad (\text{II.60})$$

Die Bedingung (II.58) besagt, dass die tangentialen Kontaktkräfte in einem Randpunkt betragsmäßig durch einen Reibungswiderstand s beschränkt sind. Dass in tangentialer Richtung keine Verschiebung vorhanden ist, wenn die tangentialen Kontaktkräfte vom Betrage her nicht groß genug sind, um den Reibungswiderstand zu überwinden, wird durch die Bedingung (II.59) ausgedrückt. In diesem Fall spricht man von Haftung bzw. von einer Haftzone. Dass eine Randpunkt in tangentialer Richtung $t(q_t - \delta_t(u))$ verschoben wird, wenn die tangentialen Kontaktkräfte genau so groß wie der Reibungswiderstand sind, wird durch die Bedingung (II.60) dargestellt. Hier spricht man von Gleiten bzw. von Gleitzonen.

Reibungsnebenbedingungen, wie sie in (II.58)-(II.60) dargestellt sind, lassen sich nicht in geeigneter Form mit Hilfe einer Restriktionsmenge wie für Hindernisnebenbedingungen formulieren. Stattdessen werden sie in Gestalt eines nicht differenzierbaren Funktionals j in die Modellierung eingearbeitet. Betrachtet wird das Minimierungsproblem: Finde ein $u \in H^1(\Omega, \Gamma_0)$, so dass gilt:

$$H(u) + j(u) = \min_{v \in H^1(\Omega, \Gamma_0)^3} H(v) + j(v). \quad (\text{II.61})$$

Dabei setzt man $H(v) := \frac{1}{2}(\sigma(v), \varepsilon(v))_0 - (f, v)_0 - (q, \delta(v))_{0, \Gamma_1}$ und $j(v) := (s, |\delta_t(v)|)_{0, \Gamma_1}$ mit einer Funktion $s \in L^2(\Gamma_1)$, $s \geq 0$, die im Folgenden den Reibungswiderstand repräsentiert.

Das Funktional j ist konvex und unterhalb stetig, demnach hat das Minimierungsproblem (II.61) nach Satz I.13 genau eine Lösung $u \in H^1(\Omega, \Gamma_0)^3$. Für diese ist notwendig und hinreichend:

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma_0)^3 : (\sigma(u), \varepsilon(v - u))_0 + (s, |\delta_t(v)|)_{0, \Gamma_1} - (s, |\delta_t(u)|)_{0, \Gamma_1} \\ \geq (f, v - u)_0 + (q, \delta(v - u))_{0, \Gamma_1}. \end{aligned} \quad (\text{II.62})$$

Ist $\sigma(u) \in H(\text{div}, \Omega)$, dann ist u nach Satz I.25 Lösung von (II.62) genau dann, wenn gilt:

$$-\text{div } \sigma(u) = f \quad (\text{II.63})$$

$$q_t - \alpha_1(u) \in \Lambda_1 \quad (\text{II.64})$$

$$\langle q_t - \alpha_1(u), \delta_t(u) \rangle = j(u) \quad (\text{II.65})$$

$$\alpha_0(u) = q_n. \quad (\text{II.66})$$

Dabei ist $\Lambda_t := \{\mu_t \in H_t^{-1/2}(\Gamma_1) \mid \forall w \in H_t^{1/2}(\Gamma_1) : |\langle \mu_t, w \rangle| \leq (s, |w|)_{0, \Gamma_1}\}$. Da $H_t^{1/2}(\Gamma_1)$ dicht in $L^2(\Gamma_1)^2$ enthalten ist, folgt mit Hilfe von Lemma I.29 und analog zu Abschnitt II.2.2, dass $\Lambda_t = \{\mu_t \in L^2(\Gamma_1)^2 \mid |\mu_t| \leq s\}$ ist. Wenn $\sigma(u) \in H^1(\Omega)^{3 \times 3}$ ist, erhält man entsprechend zur Herleitung von (II.27) daraus insbesondere

$$(q_t - \sigma_{nt}(u))^\top \delta_t(u) = s |\delta_t(u)|. \quad (\text{II.67})$$

Lemma II.2 *Ist $|q_t - \sigma_{nt}(u)| \leq s$ und $s > 0$, dann sind die Bedingungen (II.67) und (II.59)-(II.60) äquivalent.*

Beweis: Es gelte die Bedingung (II.67). Dann folgt aus der Cauchy-Schwarzschen Ungleichung: $|q_t - \sigma_{nt}(u)| |\delta_t(u)| \geq (q_t - \sigma_{nt}(u))^\top \delta_t(u) = s |\delta_t(u)| \geq |q_t - \sigma_{nt}(u)| |\delta_t(u)|$. Also ist

$$|q_t - \sigma_{nt}(u)| |\delta_t(u)| = (q_t - \sigma_{nt}(u))^\top \delta_t(u) = s |\delta_t(u)|. \quad (\text{II.68})$$

Es sei $|q_t - \sigma_{nt}(u)| < s$. Angenommen, $q_t - \sigma_{nt}(u) \neq 0$ und $\delta_t(u) \neq 0$, dann wäre $s|\delta_t(u)| > |q_t - \sigma_{nt}(u)||\delta_t(u)|$ im Widerspruch zu (II.68). Also ist entweder $q_t - \sigma_{nt}(u) = 0$ oder $\delta_t(u) = 0$. Im Fall $q_t - \sigma_{nt}(u) = 0$ folgt direkt aus (II.67), dass $s|\delta_t(u)| = 0$ ist und damit auch $\delta_t(u) = 0$.

Es sei nun $|q_t - \sigma_{nt}(u)| = s$. Nach (II.68) existiert ein $(\zeta_0, \zeta_1) \in \mathbb{R}^2 \setminus \{0\}$, so dass $\zeta_0(q_t - \sigma_{nt}(u)) + \zeta_1\delta_t(u) = 0$ ist. Ist $\zeta_0 = 0$ und damit $\zeta_1 \neq 0$, folgt $\delta_t(u) = 0$, und man setzt $\zeta := 0$. Angenommen, $\zeta_1 = 0$ und damit $\zeta_0 \neq 0$, dann wäre $q_t - \sigma_{nt}(u) = 0$ im Widerspruch zu $|q_t - \sigma_{nt}(u)| = s > 0$. Demnach bleibt nur $\zeta_0 \neq 0$ und $\zeta_1 \neq 0$, und man setzt $\zeta := -\zeta_0/\zeta_1$. Angenommen, $\zeta < 0$, dann wäre $(q_t - \sigma_{nt}(u))^\top \delta_t(u) = \zeta \delta_t(u)^\top \delta_t(u) < 0$ im Widerspruch zu (II.68).

Es seien die Bedingungen (II.58)-(II.60) erfüllt. Dann sind $\delta_t(u)$ und $q_t - \sigma_{nt}(u)$ linear abhängig. Also gilt $(q_t - \sigma_{nt}(u))^\top \delta_t(u) = |q_t - \sigma_{nt}(u)||\delta_t(u)|$. Die Bedingung (II.67) ist für $\delta_t(u) = 0$ erfüllt. Ist $q_t - \sigma_{nt}(u) = 0$, dann folgt aus (II.59), dass auch $\delta_t(u) = 0$ ist, also ist ebenfalls (II.67) erfüllt. Bleibt noch $\delta_t(u) \neq 0$ und $q_t - \sigma_{nt}(u) \neq 0$. Angenommen, $|q_t - \sigma_{nt}(u)| < s$, dann wäre $\delta_t(u) = 0$. Demnach ist $|q_t - \sigma_{nt}(u)| = s$, woraus (II.67) folgt. \square

Aus (II.63)-(II.66) sowie Lemma II.2 folgt insgesamt:

Satz II.6 *Es sei $u \in H^1(\Omega, \Gamma_0)$ mit $\sigma(u) \in H^1(\Omega)^{3 \times 3}$ und $s > 0$. Genau dann ist u Lösung von (II.62), wenn*

$$-\operatorname{div} \sigma(u) = f \quad (\text{II.69})$$

$$|q_t - \sigma_{nt}(u)| \leq s \quad (\text{II.70})$$

$$|q_t - \sigma_{nt}(u)| < s \Rightarrow \delta_t(u) = 0 \quad (\text{II.71})$$

$$|q_t - \sigma_{nt}(u)| = s \Rightarrow \exists \zeta \in \mathbb{R}_{\geq 0} : \delta_t(u) = \zeta(q_t - \sigma_{nt}(u)) \quad (\text{II.72})$$

$$\sigma_{nn}(u) = q_n. \quad (\text{II.73})$$

Wird in Satz II.6 $s \geq 0$ vorausgesetzt, so sind die Bedingungen (II.70)-(II.72) zu modifizieren. Hierzu sei auf Bemerkung II.1 verwiesen.

Die Gleichungen (II.69) und (II.73) in Satz II.6 beschreiben das Kräftegleichgewicht zwischen den inneren Reaktionskräften bzw. Normalspannungen und den Volumen- bzw. normalen Oberflächenkräften. Die Bedingungen (II.70)-(II.72) implizieren die geforderten Reibungsnebenbedingungen an die tangentialen Kontaktkräfte und an die Verschiebungen in tangentialer Richtung.

Das Funktional j erfüllt die Bedingung (I.59), insbesondere ist das Funktional $\Phi_t : H^1(\Omega, \Gamma_0)^3 \times \Lambda_1 \rightarrow \mathbb{R}$ mit $\Phi_t(v, \mu_t) := (\mu_t, \delta_t(v))_{0, \Gamma_1}$ ein Multi-

plikatorfunktional bzgl. j .²⁰ Man erhält aus Satz I.14: $u \in H^1(\Omega, \Gamma_0)^3$ ist Lösung von (II.61), wenn ein $\lambda_t \in \Lambda_t$ existiert, so dass

$$\forall v \in H^1(\Omega, \Gamma_0)^3 : \quad (\sigma(u), \varepsilon(v))_0 = (f, v)_0 + (q, \delta(v))_{0, \Gamma_1} - (\lambda_t, \delta_t(v))_{0, \Gamma_1} \quad (\text{II.74})$$

$$\forall \mu_t \in \Lambda_t : (\mu_t - \lambda_t, \delta_t(u))_{0, \Gamma_1} \leq 0 \quad (\text{II.75})$$

gilt. Da Λ_t beschränkt ist und δ_t surjektiv auf $H_t^{1/2}(\Gamma_1)$ abbildet, existiert nach Satz I.19 genau eine Lösung $(u, \lambda_t) \in H^1(\Omega, \Gamma_0) \times \Lambda_t$ von System (II.74)-(II.75).

Wenn $u \in H^1(\Omega, \Gamma_0)$ und $\sigma(u) \in H^1(\Omega)^{3 \times 3}$ gilt, ist (u, λ_t) nach Satz I.26 genau dann Lösung von (II.74)-(II.75), wenn u (II.62) oder (II.69)-(II.73) erfüllt und wenn $\lambda_t = q_t - \sigma_{nt}(u)$ ist. Der Lagrangesche Multiplikator λ_t entspricht demnach den negativen tangentialen Kontaktkräften.

II.3.3 Kontaktprobleme mit Reibung

Sollen sowohl Hindernisnebenbedingungen als auch Reibungsnebenbedingungen in die Modellierung eingehen, dann ist das Minimierungsproblem

$$\min_{v \in K} H(v) + j(v) \quad (\text{II.76})$$

zu untersuchen. Nach Satz I.13 besitzt das Minimierungsproblem (II.76) genau eine Lösung $u \in H^1(\Omega, \Gamma_0)^3$, für die die folgende Bedingung notwendig und hinreichend ist:

$$\forall v \in K : (\sigma(u), \varepsilon(v - u))_0 + (s, |\delta_t(v)|)_{0, \Gamma_1} - (s, |\delta_t(u)|)_{0, \Gamma_1} \geq (f, v - u)_0 + \langle q, \delta(v - u) \rangle. \quad (\text{II.77})$$

Aus Satz I.27 folgt für $u \in K$ mit $\sigma(u) \in H(\text{div}, \Omega)$: u ist genau dann Lösung von (II.77), wenn

$$-\text{div } \sigma(u) = f \quad (\text{II.78})$$

$$(q_n - \alpha_0(u), q_t - \alpha_1(u)) \in \Lambda_0 \times \Lambda_1 \quad (\text{II.79})$$

$$\langle q_n - \alpha_0(u), \delta_n(u) - g \rangle = 0 \quad (\text{II.80})$$

$$\langle q_t - \alpha_1(u), \delta_t(u) \rangle = (s, |\delta_t(u)|)_{0, \Gamma_1} \quad (\text{II.81})$$

gilt. Wie in den Sätzen II.5 und II.6 erhält man aus (II.78)-(II.81) und Lemma II.2:

²⁰vgl. Abschnitt II.2.2

Satz II.7 *Es sei $u \in K$ mit $\sigma(u) \in H^1(\Omega)^{3 \times 3}$ und $s > 0$. Genau dann ist u Lösung von (II.77), wenn gilt:*

$$-\operatorname{div} \sigma(u) = f \quad (\text{II.82})$$

$$q_n - \sigma_{nn}(u) \geq 0 \quad (\text{II.83})$$

$$(q_n - \sigma_{nn}(u))(\delta_n(u) - g) = 0 \quad (\text{II.84})$$

$$|q_t - \sigma_{nt}(u)| \leq s \quad (\text{II.85})$$

$$|q_t - \sigma_{nt}(u)| < s \Rightarrow \delta_t(u) = 0 \quad (\text{II.86})$$

$$|q_t - \sigma_{nt}(u)| = s \Rightarrow \exists \zeta \in \mathbb{R}_{\geq 0} : \delta_t(u) = \zeta(q_t - \sigma_{nt}). \quad (\text{II.87})$$

Für den Fall $s \geq 0$ sei auf Bemerkung II.1 verwiesen.

Wie zuvor in den physikalischen Interpretationen von Satz II.5 und Satz II.6 bedeutet die Gleichung (II.82), dass die inneren Reaktionskräfte mit den Volumenkräften im Gleichgewicht sind. Die Gleichung (II.84) besagt, dass die Normalspannungen im Gleichgewicht mit den normalen Oberflächenkräften sind oder Kontakt besteht. Die Bedingung (II.83) ist in der Weise zu interpretieren, dass höchstens Druckkräfte vorhanden sein können. Die Reibungsnebenbedingungen sind in den Bedingungen (II.85)-(II.87) wiederzufinden.

Man erhält aus Satz I.15, dass $u \in H^1(\Omega, \Gamma_0)^3$ Lösung von (II.76) ist, wenn ein $(\lambda_n, \lambda_t) \in \Lambda_n \times \Lambda_t$ existiert, so dass gilt:

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma_0)^3 : \\ (\sigma(u), \varepsilon(v))_0 = (f, v)_0 + (q, \delta(v))_{0, \Gamma_1} - \langle \lambda_n, \delta_n(v) \rangle - \langle \lambda_t, \delta_t(v) \rangle \end{aligned} \quad (\text{II.88})$$

$$\forall (\mu_n, \mu_t) \in \Lambda_n \times \Lambda_t : \langle \mu_n - \lambda_n, \delta_n(v) - g \rangle + (\mu_t - \lambda_t, \delta_t(u))_{0, \Gamma_1} \leq 0. \quad (\text{II.89})$$

Da $\operatorname{rg}(\delta_n)$ abgeschlossen in $H_n^{1/2}(\Gamma_1)$ und Λ_t beschränkt ist und zudem $\delta_n(\ker \delta_t) = H_n^{1/2}(\Gamma_1)$ und $\delta_t(\ker \delta_n) = H_t^{1/2}(\Gamma_1)$ gilt, besitzt das System (II.88)-(II.89) nach Satz I.20 genau eine Lösung $(u, \lambda_n, \lambda_t)$.

Ist $u \in H^1(\Omega, \Gamma_0)$ mit $\sigma(u) \in H(\Omega)^{3 \times 3}$, dann ist $(u, \lambda_n, \lambda_t)$ nach Satz I.28 genau dann Lösung von (II.88)-(II.89), wenn u (II.77) oder (II.82)-(II.87) erfüllt und wenn $\lambda_n = q_n - \sigma_{nn}(u)$ und $\lambda_t = q_t - \sigma_{nt}(u)$ ist.

II.3.4 Das Coulombsche Reibungsgesetz

Bei den bisher vorgestellten Minimierungsproblemen (II.61) und (II.76) wurden Reibungsnebenbedingungen über eine noch nicht spezifizierte Reibungswiderstandsfunktion s eingeführt. Die Frage ist nun, in welcher Weise diese

Funktion anzugeben ist.

Das Coulombsche Reibungsgesetz besagt, dass der Reibungswiderstand proportional zum Betrag der normalen Kontaktkräfte ist. Der zugehörige Proportionalitätsfaktor $\mathcal{F} \geq 0$ heißt Reibungskoeffizient und ist im Wesentlichen abhängig von der Oberflächenbeschaffenheit des elastischen Körpers und des starren Hindernisses.

Für das Minimierungsproblem (II.61) ist nach dem Coulombschen Gesetz demnach $s := \mathcal{F}|q_n|$ zu setzen. Da die normalen Oberflächenkräfte q_n a-priori bekannt sind, ist diese Wahl in Bezug auf die in Abschnitt II.3.2 verwendeten Eindeutigkeits- und Existenzaussagen unproblematisch.

Anders ist die Situation für das Minimierungsproblem (II.76). Hier ist $s := \mathcal{F}|q_n - \sigma_{nn}(u)|$ zu setzen, womit das Funktional j die Gestalt

$$j(v) := (\mathcal{F}|q_n - \sigma_{nn}(u)|, |\delta_t(v)|)_{0, \Gamma_1} \quad (\text{II.90})$$

erhält. In diesem Fall ergeben sich zwei wesentliche Probleme: Zum einen ist der Ausdruck $(s, |\delta_t(v)|)_{0, \Gamma_1}$ nicht sinnvoll definiert, wenn $\sigma_{nn}(u)$ nicht existiert bzw. $\alpha_0(u) \notin L^2(\Gamma_1)$ ist, was bei einer zu geringen Regularität von u der Fall ist. Zum anderen geht mit $\sigma_{nn}(u)$ eine Information in das Funktional j ein, die a-priori nicht bekannt ist. Die Konsequenz aus letzterem ist, dass die Existenz einer Lösung nicht direkt mit Hilfe von Satz I.13 geklärt werden kann.

Stattdessen kann die Frage nach der Existenz einer Lösung im Sinne einer Fixpunkt Betrachtung weiter untersucht werden: Bei vorgegebenem $s \in L^2(\Gamma_1)$ mit $s \geq 0$ sei $u(s)$ die nach Satz I.13 eindeutige, hinreichend reguläre Lösung von (II.77). Der Operator $\mathcal{H} : L^2(\Gamma_1) \rightarrow L^2(\Gamma_1)$ sei definiert als $\mathcal{H}(s) := \mathcal{F}|q_n - \sigma_{nn}(u(s))|$. Besitzt nun \mathcal{H} einen Fixpunkt $\bar{s} \in L^2(\Gamma_1)$, gilt also $\mathcal{H}(\bar{s}) = \bar{s}$, dann ist $u(\bar{s})$ Lösung des Minimierungsproblem (II.76) mit dem in (II.90) angegebenen Funktional j .

In [68] und [82] wird mit Hilfe einer schwachen Version des Fixpunktsatzes von Schauder gezeigt, dass \mathcal{H} einen Fixpunkt besitzt, sofern der Reibungskoeffizient der Beschränkung

$$\max_{x \in \Gamma_1} \mathcal{F}(x) \leq \sqrt{(2\mu)(\lambda + 3\mu)^{-1}}$$

genügt. Außerdem wird in den genannten Arbeiten vorausgesetzt, dass Ω ein unendlich langer Streifen ist und dass $q_n = 0$ ist. Verallgemeinerungen auf beliebige Körper $\Omega \subset \mathbb{R}^3$ sind in [71] enthalten. Abschließend ist jedoch anzumerken, dass durch das Fehlen einer allgemeinen Existenztheorie das Coulombsche Reibungsgesetz nicht unumstritten ist.²¹

²¹vgl. S.271 in [72]

II.4 Eine Anwendung in der Fertigungstechnik

Ein spezielles fertigungstechnisches Problem, das auf reibungsbehaftete Kontaktprobleme führt und das in dieser Arbeit exemplarisch als Anwendungsbereich dargestellt wird, ist das Industrieroboter-gestützte Bandschleifen.

Beim Industrieroboter-gestützten Bandschleifen führt ein Roboterarm ein Werkstück gegen eine schnell rotierende Kontaktscheibe, auf der ein Schleifband gespannt ist. Durch die besondere Oberflächenbeschaffenheit des Schleifbandes kommt es im Kontaktbereich von Kontaktscheibe und Werkstück zu Reibungseffekten mit Materialabtrag am Werkstück.



Abb. II.3: Roboterarm mit Werkstück (Sanitärarmatur), Bandschleifgerät.

In den Abbildungen II.3 und II.4 ist eine typische Konfiguration von Bandschleifapparatur und Werkstück zu sehen.

Durch die flexible Zuführung des Werkstücks mit Hilfe eines mehrgelenkigen Roboterarms und durch Verwenden eines Gummi-elastischen Kontaktscheibenbelags ist man in der Lage, auch komplexe Oberflächengeometrien von Werkstücken zu bearbeiten.

Wichtig für die Zuführung, die einen vordefinierten Abtrag sicherstellen soll, und damit für die Steuerung des Roboterarms ist die genaue Kenntnis der in der Kontaktzone auftretenden Kräfteverhältnisse. Entscheidend dabei ist, dass die normalen Kontaktkräfte und die Lage der Kontaktzone in der Regel mit dem Abtragsergebnis am Werkstück in Beziehung gesetzt werden können. Für einen Überblick zum Industrieroboter-gestützten Bandschleifen sei auf [30], [88] und [108] verwiesen.

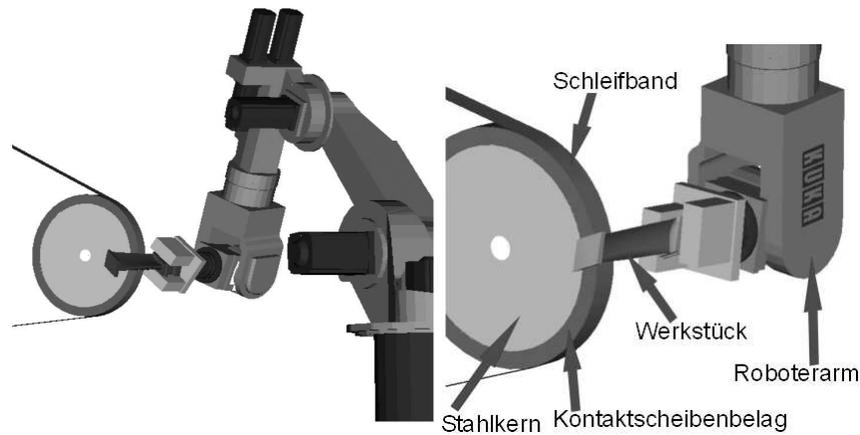


Abb. II.4: Schematische Darstellung des Bandschleifprozesses.

Ziel dieses Abschnitts ist, aufgrund der in Abschnitt II.3 betrachteten Modellformulierungen den reibungsbehafteten Kontaktanteil des Industrieroboter-gestützten Bandschleifens an einer Beispielkonfiguration zu untersuchen. Insbesondere werden die Aussagen und die daraus resultierenden physikalischen Interpretationen der Sätze II.5, II.6 und II.7 veranschaulicht. Hierbei wird im Wesentlichen ausgenutzt, dass man mit Hilfe variationeller Sattelpunktformulierungen die normalen und tangentialen Kontaktkräfte in Gestalt der Lagrangeschen Multiplikatoren erhält.

Mit den zur Verfügung stehenden Modellformulierungen können nur Teilspekte des gesamten Bandschleifprozesses untersucht werden.

So werden etwa Schwingungseffekte bei der Roboterarmzuführung, die im Allgemeinen zu ungenauen Zustellergebnissen führen, nicht berücksichtigt. Auch Beschleunigungseffekte der elastischen Kontaktscheibe, die durch ihre Rotation entstehen, werden nicht berücksichtigt. Insgesamt wird der Bandschleifprozess als statischer Vorgang modelliert, betrachtet wird die Kontaktsituation nur zu einem bestimmten Zeitpunkt.

Die mikroskopische Oberflächenbeschaffenheit des Schleifbandes, d.h. die Verteilung und Eigenschaften der Schleifkörner auf dem Schleifband, wird als homogen vorausgesetzt, so dass die Reibungseigenschaften der Schleifbandoberfläche allein durch einen konstanten Reibungskoeffizienten \mathcal{F} beschrieben werden. Darüber hinaus wird das Schleifband in der Modellierung nicht weiter berücksichtigt.

Die Modellformulierungen aus Abschnitt II.3 setzen das Hindernis als starr

voraus. Dem entsprechend wird das Werkstück als Starrkörper betrachtet. Für eine Modellierung, die die Deformation zweier elastischer Körper einbezieht, sei zum Beispiel auf [38] und [72] verwiesen. Für die Einbeziehung dynamischer Effekte siehe etwa [35], [37], [38] oder [72].

Im Folgenden wird eine konkrete Beispielkonfiguration des Bandschleifprozesses vorgestellt, die mit Hilfe der in Abschnitt II.3 vorgestellten drei Kontaktformulierungen (reibungsfreier Kontakt, reibungsbehafteter Kontakt mit vorgegebener Normalspannung und reibungsbehafteter Kontakt) untersucht werden kann.

Veranschaulicht werden die zugehörigen physikalischen Zusammenhänge anhand von Finite-Elemente-Berechnungen, die hier als Näherungen an die kontinuierlichen Lösungen dienen. Für weitere Ausführungen, die Finite-Elemente-Methode betreffend, sei auf Kapitel III verwiesen. Die bei den Beispielrechnungen verwendete Auflösung der Finite-Elemente-Gitter ist relativ grob gewählt, da sich hiermit unter den eingesetzten Darstellungsmöglichkeiten die Kontaktverhältnisse geeignet demonstrieren lassen. Verfeinerte Techniken zur adaptiven Wahl der Gitterauflösung werden in Abschnitt IV.7 vorgestellt.

Betrachtet wird eine handelsübliche Kontaktscheibe, bestehend aus einem Kontaktscheibenbelag und einem harten Stahlkern, wie sie etwa in den Abbildungen II.5 und II.6 zu sehen ist. Als Beispiel wird eine Kontaktscheibe gewählt, die einen Radius von 1.625 dm und eine Breite von 1.15 dm hat. Der Durchmesser des Kontaktscheibenbelags beträgt 0.33 dm, der Radius des Stahlkerns ist 1.295 dm.

Der Kontaktscheibenbelag wird als profillos vorausgesetzt²² und besteht aus einem Elastomer. Bekanntermaßen setzen sich Elastomere aus dicht verfilzt aneinandergereiht liegenden Fadenmolekülen zusammen, die durch Atombindungen nachträglich weitmaschig vernetzt werden. Sie verhalten sich bis zur Zersetzungstemperatur Gummi-elastisch und können (u.U. um mehrere hundert Prozent) reversibel gedehnt werden. Durch die chemische Netzbindung wird ein Fließen ganz verhindert.²³ Das Gummi-elastische Materialverhalten des Kontaktscheibenbelags kann deshalb angemessen mit dem in Abschnitt II.3 dargestellten Hookeschen Materialgesetz näherungsweise abgebildet werden, da nur kleine Deformationen zu erwarten sind. Der für Elastomere bzw. Gummi-elastische Materialien spezifische Elastizitätsmo-

²²Tatsächlich werden die Kontaktscheibenbeläge häufig mit Nutungen oder Einfräsungen versehen, um die Haltbarkeit zu erhöhen oder die Geräusentwicklung zu senken (vgl. Abbildung II.6).

²³vgl. [45], S. 269

dul E beträgt 0.1mN/dm^2 bis 2mN/dm^2 bei einer Querkontraktionszahl von ca. $\nu = 0.42$.²⁴ Für einen Vergleich mit experimentell gewonnenen Daten sei diesbezüglich auf [31] verwiesen. Bei großen Deformationen ist das Materialverhalten von Elastomeren jedoch im Allgemeinen nichtlinear.²⁵ Da der Stahlkern im Vergleich zum Kontaktscheibenbelag weitaus weniger elastisch ist ($E = 2100\text{mN/dm}^2$), kann angenommen werden, dass die elastischen Verformungen im Stahlkern vernachlässigbar sind und dass somit der Stahlkern als starr vorausgesetzt werden kann. Es reicht also, das Deformationsverhalten des Kontaktscheibenbelags mit Nullrandverschiebungen im Verbindungsbereich von Kontaktscheibenbelag und Stahlkern zu modellieren.

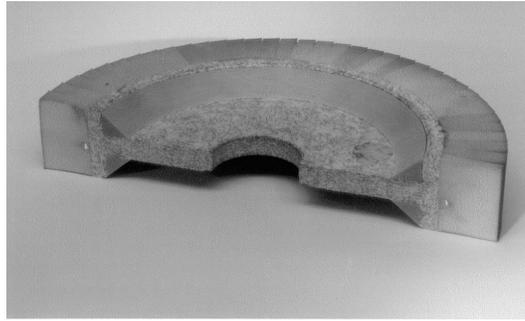


Abb. II.5: Querschnitt einer Kontaktscheibe.



Abb. II.6: Kontaktscheibe mit Nutungen im Kontaktscheibenbelag, Kontaktscheibe mit profillosem Kontaktscheibenbelag.

²⁴vgl. [69], S. 105-107

²⁵vgl. [20], Abschnitt 6.4

Der mögliche Kontaktbereich auf der Außenseite des Kontaktscheibenbelags ist relativ klein und lässt sich insbesondere über die Kenntnis der Werkstückzuführung recht gut im Vornherein eingrenzen. Da Deformationen nur durch die Kontaktsituation hervorgerufen werden, können Deformationen außerhalb des möglichen Kontaktbereichs vernachlässigt werden. Es ist demnach ausreichend, zum Beispiel nur ein Viertel des gesamten Kontaktscheibenbelags zu betrachten und für die Übergangsstellen zum Restkörper entweder Neumannranddaten oder Nullrandverschiebungen zu fordern. Mit den Polarkoordinaten (r, φ) mit dem Ursprung in $(-1.625, 0)$ ist

$$\Omega := \left\{ (x, y, z) \in \mathbb{R}^3 \mid \begin{array}{l} r(x, z) \in (1.295, 1.625), \\ \varphi(x, z) \in [0, \pi/4] \cup (7\pi/4, 0], \\ y \in (-0.575, 0.575) \end{array} \right\}$$

und $\Gamma_0 := \{(x, y, z) \in \bar{\Omega} \mid r(x, z) = 1.295\}$ sowie $\Gamma_1 := \partial\Omega \setminus \Gamma_0$.

Bei dem vorliegenden Bandschleifprozess wird vorausgesetzt, dass die Kontaktscheibe über das Schleifband angetrieben wird.²⁶ Die hierdurch auf einen Teil der Oberfläche des Kontaktscheibenbelags $\tilde{\Gamma}_1 := \{(x, y, z) \in \Gamma_1 \mid r(x, z) = 1.625\}$ induzierten Kräfte gehen als tangentiale Oberflächenkräfte q_t in die Modellierung ein. Volumen-bezogene Kräfte werden, sofern vorhanden, vernachlässigt.

II.4.1 Reibungsfreier Kontakt

Sollen beim Bandschleifprozess nur die geometrischen Kontaktverhältnisse bei einer vordefinierten Zustellung beschrieben und noch keine Reibungsbedingungen berücksichtigt werden, wird die Kontaktsituation durch das Minimierungsproblem (II.46) bzw. durch die variationelle Sattelpunktformulierung (II.56)-(II.57) aus Abschnitt II.3.1 als reibungsfreies Kontaktproblem geeignet modelliert.

Es werden keine auf den Oberflächenbereich $\tilde{\Gamma}_1$ wirkenden Normal- oder Tangentialkräfte einbezogen, da diese für das reibungsfreie Kontaktproblem irrelevant sind. Für die Oberflächenkräfte wird demnach $q := 0$ gesetzt.

²⁶Alternativ kann auch die Kontaktscheibe direkt von einem Motor angetrieben werden.

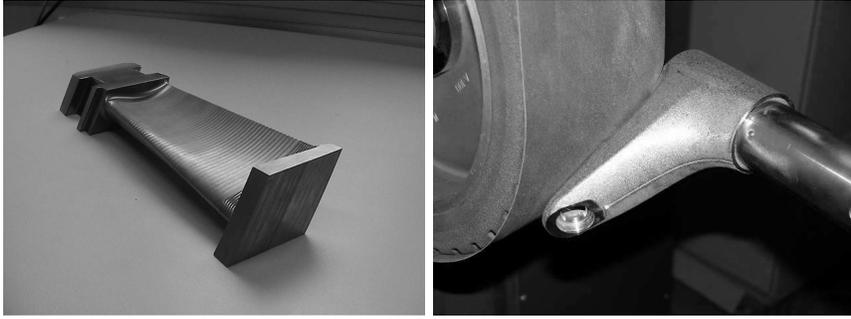


Abb. II.7: Turbinenschaufel, Sanitärarmatur in Kontakt mit Schleifband und Kontaktscheibe.



Abb. II.8: Kontaktscheibenbelag, Werkstückoberfläche.

Als Musterwerkstück wird ein Metallwerkstück mit einer gebogenen Oberfläche (z.B. eine Turbinenschaufel oder eine Sanitärarmatur, vgl. Abbildung II.7) betrachtet. Die Oberfläche wird hier durch die Funktion $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$\psi(y, z) := \begin{cases} d + 1 - \sqrt{1 - (z + 0.5y)^2}, & |z + 0.5y| \leq r \\ d + 1, & |z + 0.5y| > r \end{cases}$$

bzgl. der yz -Koordinatenebene exemplarisch beschrieben. Die Zustellung des Werkstücks entlang der x -Achse wird über den Parameter $d \in \mathbb{R}$ realisiert.



Abb. II.9: Kontaktscheibenbelag in Kontakt mit dem Werkstück bei einer Zustellung von $d := -0.05$ dm.

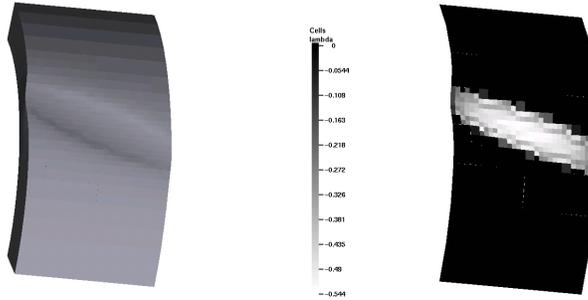


Abb. II.10: Deformierter Kontaktscheibenbelag, normale Kontaktkräfte $\sigma_{nn}(u) = -\lambda_n$.

Die Abbildungen II.9 und II.10 (links) stellen die Deformation des Kontaktscheibenbelags dar, die sich durch die Verschiebung u als Lösung von (II.46) bzw. (II.56)-(II.57) ergibt.

Dass u tatsächlich die Hindernisnebenbedingung $\delta_n(u) - g \leq 0$ erfüllt, wird in den Abbildungen II.9 und II.10 durch das Anpressprofil deutlich: Der Kontaktscheibenbelag wird durch die Zuführung des Werkstücks in einer solchen Weise deformiert, dass ein Eindringen des Werkstücks in den Kontaktscheibenbelag verhindert wird.

Die Forderungen $\sigma_{nn}(u) \leq 0$ und $\sigma_{nn}(u)(\delta_n(u) - g) = 0$ aus Satz II.5 werden in Abbildung II.10 (rechts) veranschaulicht. Die normalen Kontaktkräfte $\sigma_{nn}(u) = -\lambda_n$ sind außerhalb der Kontaktzone null und innerhalb negativ, also Druckkräfte.

II.4.2 Reibungsbehafteter Kontakt mit vorgegebenen Normalspannungen

Sind die Normalspannungen und die Kontaktzone a-priori bekannt, kann die reibungsbehaftete Kontaktsituation beim Bandschleifprozess mit Hilfe des Minimierungsproblems (II.61) bzw. (II.74)-(II.75) modelliert werden. Zum Beispiel sind die Normalspannungen bekannt, wenn sie, wie in Abbildung II.11 zu sehen, mit Hilfe eines speziellen Versuchsaufbaus ermittelt werden²⁷ oder über die Schleifmaschine selbst vorgegeben werden können.²⁸ Die Kontaktzone kann für einfache Werkstückgeometrien, wie zum Beispiel für einen Quader, leicht im Voraus bestimmt werden.

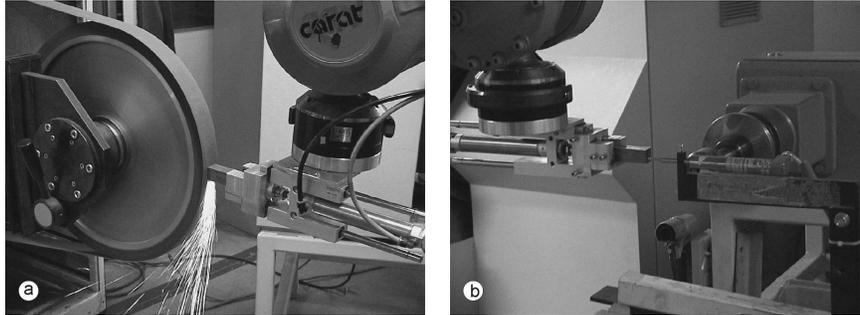


Abb. II.11: Versuchsaufbau zur Messung der Normalspannungen.

Im Gegensatz zum reibungsfreien Kontakt sind die durch das Schleifband induzierten tangentialen Oberflächenkräfte von entscheidender Bedeutung für die Ausbildung von Gleit- bzw. Haftzonen. Die tangentialen Oberflächenkräfte werden hier exemplarisch $q_t := (0, -0.05)^T$ gesetzt. In der nachfolgenden Kontaktsituation ist zu beachten, dass der zweite Tangentialvektor stets in Richtung der positiven z -Achse zeigt, dass also die Verschiebungen auf dem Randsegment $\tilde{\Gamma}_1$ durch die durch q_t induzierten Kräfte tendenziell in Richtung der negativen z -Achse zeigen.

In dem folgenden Beispiel-Problem sind die Normalspannungen durch Setzen der normalen Oberflächenkräfte mit

$$q_n := \begin{cases} -0.2, & |z + 0.5y| \leq r \\ 0, & |z + 0.5y| > r \end{cases}$$

²⁷vgl. [88], S. 73-74

²⁸vgl. [58] oder [59]

vorgegeben. Dementsprechend ist der Bereich $|z+0.5y| \leq r$ die Kontaktzone.

Die Reibungswiderstandsfunktion ist, wie in Abschnitt II.3.4 beschrieben, durch $s := \mathcal{F}|q_n|$ mit $\mathcal{F} := 0.5$ definiert. Insgesamt erhält man die folgende Kontaktsituation:



Abb. II.12: Deformierter Kontaktscheibenbelag in Kontakt mit dem Werkstück. Normalspannungen und Kontaktzone sind a-priori bekannt.

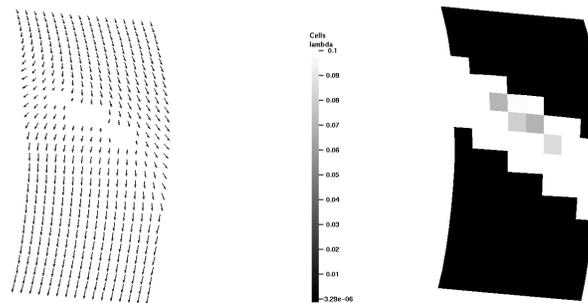


Abb. II.13: Tangentialverschiebungen auf $\tilde{\Gamma}_1$, Absolutbetrag der Tangentialkräfte $|\lambda_t|$.

In Abbildung II.12 ist der durch die vorgegebenen Normalspannungen deformierte Kontaktscheibenbelag zu sehen. Deutlich erkennbar ist die Kontaktzone.

Abbildung II.13 (links) zeigt die Tangentialverschiebungen in $\tilde{\Gamma}_1$. Erkennbar ist, dass die Verschiebungen außerhalb der Kontaktzone im Wesentlichen in Richtung der vorgegebenen Tangentialkräfte verlaufen. Im Innern der Kon-

taktzone bilden sich Gleit- und Haftzonen²⁹ aus, wobei die Haftzonen als weiße Gebiete zu erkennen sind, in denen keine Verschiebungsvektoren abgebildet werden. Die Gleit- und Haftzonen korrespondieren über die in Satz II.6 beschriebenen Bedingungen (II.70)-(II.72) mit dem in Abbildung II.13 dargestellten Absolutbetrag von $\lambda_t = q_t - \sigma_{nt}$. Ist $|\lambda_t| < 0.1 = s$, kommt es zum Haftkontakt, also $\delta_t(u) = 0$. Ist dagegen $|\lambda_t| = 0.1$, wird eine Gleitzzone ausgebildet.

II.4.3 Reibungsbehafteter Kontakt

In der Regel sind beim Bandschleifprozess weder die Normalspannungen noch die Kontaktzone bekannt. Diese allgemeine reibungsbehaftete Kontaktsituation kann über das Minimierungsproblem (II.76) oder über die variationelle Sattelpunktformulierung (II.88)-(II.89) angemessen modelliert werden. Auch hier sind tangentiale Oberflächenkräfte zu berücksichtigen. Wie oben werden diese mit $q_t := (0, -0.05)^\top$ angesetzt. Für die reibungsbehaftete Kontaktsituation ist die Reibungswiderstandsfunktion, wie in Abschnitt II.3.4 angegeben, als $s := \mathcal{F}|q_n - \sigma_{nn}(u)| = \mathcal{F}|\lambda_n|$ zu wählen. Der Reibungskoeffizient wird ebenfalls mit $\mathcal{F} := 0.5$ festgesetzt.

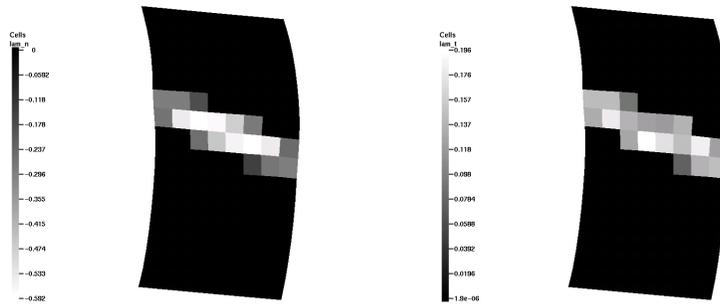


Abb. II.14: Normale Kontaktkräfte $\sigma_{nn}(u) = -\lambda_n$, Absolutbetrag der tangentialen Kontaktkräfte $|\sigma_{nt}(u) - q_t| = |\lambda_t|$.

²⁹Die in Abbildung II.13 (links) abgebildeten Vektoren sind logarithmisch skaliert. Die Länge der Vektoren in der Kontaktzone ist tatsächlich wesentlich kleiner.

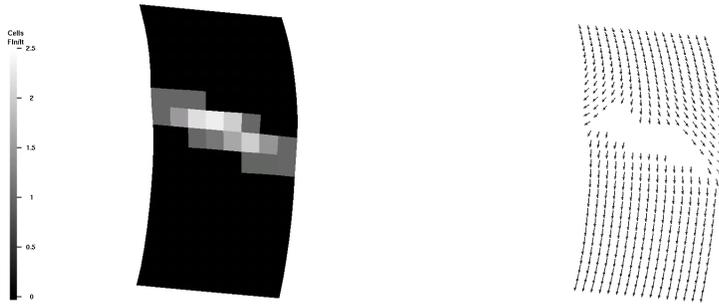


Abb. II.15: Quotient $\mathcal{F}|\lambda_n|/|\lambda_t|$, Randverschiebungen auf $\bar{\Gamma}_1$.

In Abbildung II.14 sind die normalen Kontaktkräfte $\sigma_{nn}(u) = -\lambda_n$ (links) und der Absolutbetrag der tangentialen Kontaktkräfte $\sigma_{nt}(u) - q_t = -\lambda_t$ (rechts) zu sehen. In Abbildung II.15 (links) ist der Quotient $\mathcal{F}|\lambda_n|/|\lambda_t|$ abgebildet. Dabei sind Bereiche der Kontaktzone mit $\mathcal{F}|\lambda_n|/|\lambda_t| = 1$ als Gleitzonen zu erkennen. Diese befinden sich insbesondere am linken und rechten Rand der Kontaktzone. Gilt $|\lambda_t|/(\mathcal{F}|\lambda_n|) > 1$, liegt Haften vor. Der Quotient ist außerhalb der Kontaktzone null.

Die Verteilung der Gleit- und Haftzonen stimmen mit den in Abbildung II.15 (rechts) dargestellten Tangentialverschiebungen auf dem Rand überein. Haftzonen sind dort, wo keine Verschiebungsvektoren abgebildet werden.

II.5 Weitere Modellprobleme

Weitere Beispiele für Probleme mit modellhaftem Charakter, die von ähnlicher Gestalt wie die in Abschnitt II.2 aufgeführten Modellprobleme sind, aber nicht dem Kontext reibungsbehafteter Kontaktprobleme zugeordnet werden, sind modellhafte Vereinfachungen von Hindernisproblemen, elastoplastischen Torsionsproblemen und Typen viskoplastischer Bingham-Fluide. In diesem Abschnitt werden die zugehörigen Modellformulierungen in Form von Variationsungleichungen und variationellen Sattelformulierungen kurz dargestellt. Für weitere Details, insbesondere in Bezug auf Aussagen, die denen aus den Sätzen II.2 und II.3 entsprechen, sei z.B. auf [47], [55] und [56] verwiesen.

II.5.1 Hindernisprobleme

Betrachtet man statt der in (II.9) angegebenen Restriktionsmenge die Menge

$$K := \{v \in H^1(\Omega, \Gamma_0) \mid v \geq g\} \quad (\text{II.91})$$

mit $g \in H^1(\Omega)$, so erhält man ein ähnlich gelagertes Problem wie das oben dargestellte Signorini-Problem.³⁰ Der Unterschied besteht darin, dass nun Restriktionen auf dem ganzen Gebiet Ω , also nicht allein auf dem Randsegment Γ_1 gefordert werden. Das dazugehörige Minimierungsproblem wie in (II.10) beschreibt die Durchbiegung einer Membran bei Vorhandensein eines Hindernisses g bezüglich der gesamten Membranoberfläche. Dieses Minimierungsproblem wird in der Regel Hindernis- oder Obstacle-Problem genannt. Da K konvex und abgeschlossen ist, existiert, sofern $K \neq \emptyset$ ist, nach Satz I.11 genau eine Lösung, die wie in (II.11) variationell charakterisiert ist.

Das Funktional $\Phi : H^1(\Omega, \Gamma_0) \times G' \rightarrow \mathbb{R}$ mit $\Phi(v, \mu) := \langle \mu, v - g \rangle$ ist analog zu (II.18) ein Multiplikatorfunktional bzgl. K , wobei hier der abgeschlossene und konvexe Kegel $G := \{v \in H^1(\Omega, \Gamma_0) \mid v \leq 0\}$ zu wählen ist. Nach Satz I.12 ist u eine Lösung des Hindernisproblems, wenn ein $\lambda_0 \in \Lambda_0 := G'$ existiert, so dass

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla u, \nabla v)_0 = (f, v)_0 + (q, v)_{0, \Gamma_1} - \langle \lambda_0, v \rangle \quad (\text{II.92})$$

$$\forall \mu_0 \in \Lambda_0 : \langle \mu_0 - \lambda_0, u - g \rangle \leq 0 \quad (\text{II.93})$$

gilt. Nach Satz I.18 ist die eindeutige Existenz einer Lösung von (II.92)-(II.93) gesichert.

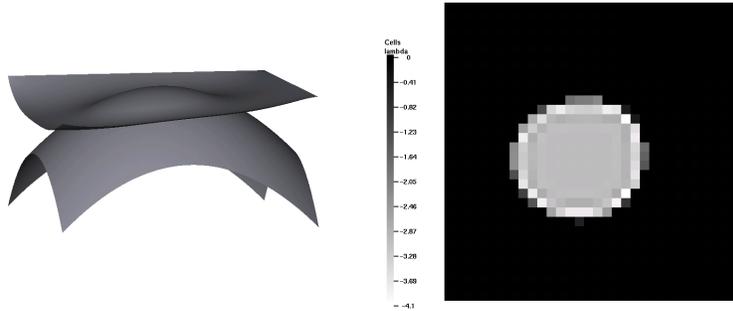


Abb. II.16: Membran auf einem Hindernis, Lagrangescher Multiplikator.

³⁰Hierbei ist \leq bzw. \geq wieder als fast überall gültig zu verstehen.

In Abbildung II.16 (links) ist eine Membran zu sehen, die auf einem Hindernis aufliegt. Das zugrunde liegende Gebiet ist hier $\Omega := (-1, 1)^2$, das Hindernis wird durch die Funktion $g(x_0, x_1) := -\frac{1}{2}(x_0^2 + x_1^2)$ beschrieben. Der zugehörige Lagrangesche Multiplikator λ_0 ist erwartungsgemäß in der Kontaktzone negativ und außerhalb null.

II.5.2 Elastoplastische Torsionsprobleme

Ein weiteres Beispiel für elliptische Minimierungsprobleme erster Art mit Restriktionen, die auf ganz Ω definiert sind, erhält man durch Suchen von

$$\min_{v \in K} \frac{1}{2}(\nabla v, \nabla v)_0 - (f, v)_0 \quad (\text{II.94})$$

$$\text{mit } K := \{v \in H_0^1(\Omega) \mid \nabla v^\top \nabla v \leq 1\}. \quad (\text{II.95})$$

Die Minimierungsaufgabe (II.94)-(II.95) ist ein Modellbeispiel für elasto-plastische Torsion, das etwa in [47], Ch. 5.6 geeignet physikalisch motiviert wird. Auch hier ist K konvex und abgeschlossen³¹, womit nach Satz I.11 genau eine Lösung u existiert, die durch

$$\forall v \in K : (\nabla u, \nabla(v - u))_0 \geq (f, v - u)_0$$

äquivalent gekennzeichnet ist. Zur Herleitung einer variationellen Sattelpunktformulierung ist als Restriktionsmenge in (II.95) die Menge $K = \{v \in H_0^1(\Omega) \mid 1 - \nabla v^\top \nabla v \in G\}$ mit dem abgeschlossenen und konvexen Kegel $G := \{v \in L^2(\Omega) \mid v \geq 0\}$ zu betrachten. Zusammen mit Bemerkung I.3 folgt aus Lemma I.3, dass $\Phi : H^1(\Omega, \Gamma_0) \times G' \rightarrow \mathbb{R}$ mit $\Phi(v, \mu) := \langle \mu, \nabla v^\top \nabla v - 1 \rangle = (\mu, \nabla v^\top \nabla v - 1)_0$ ein Multiplikatorfunktional bzgl. K ist. Wegen $\Phi_\lambda(u + v) - \Phi_\lambda(u) = (\lambda, 2\nabla u^\top \nabla v)_0 + (\lambda, \nabla v^\top \nabla v)_0$ gilt $\langle \Phi'_\lambda(u), v \rangle = 2(\lambda, \nabla u^\top \nabla v)$. Damit folgt nach Satz I.12, dass u eine Lösung des Torsionsproblems ist, wenn ein $\lambda_0 \in \Lambda_0 := G'$ existiert, so dass

$$\forall v \in H_0^1(\Omega) : (\nabla u, \nabla v)_0 + 2(\lambda_0, \nabla u^\top \nabla v)_0 = (f, v)_0 \quad (\text{II.96})$$

$$\forall \mu_0 \in \Lambda_0 : (\mu_0 - \lambda_0, \nabla u^\top \nabla u - 1)_0 \leq 0 \quad (\text{II.97})$$

erfüllt ist. Unter Beachtung, dass der Faktor 2 in (II.96) durch Variation über den Kegel Λ_0 absorbiert werden kann, kann statt (II.96)-(II.97) auch

$$\forall v \in H_0^1(\Omega) : ((1 + \tilde{\lambda}_0)\nabla u, \nabla v)_0 = (f, v)_0$$

$$\forall \mu_0 \in \Lambda_0 : (\mu_0 - \tilde{\lambda}_0, \nabla u^\top \nabla u - 1)_0 \leq 0$$

³¹vgl. Th.3.1 in [56]

betrachtet werden, wobei $\tilde{\lambda}_0 := 2\lambda_0$ ist. Da der Operator $\beta_0 : H_0^1(\Omega) \rightarrow L^2(\Omega)$ mit $\beta_0(v) := \nabla v^\top \nabla v$ nicht in $L(H_0^1(\Omega), L^2(\Omega))$ enthalten ist, können die Aussagen zur Sicherung der Existenz eines Sattelpunktes aus Abschnitt I.2 nicht angewandt werden. Tatsächlich ist der Nachweis eines Sattelpunktes in $H_0^1(\Omega)$ unter den hier aufgeführten Voraussetzungen ein offenes Problem.³²

II.5.3 Bingham-Fluid-Probleme

Neben dem in Abschnitt II.2.2 dargestellten idealisierten Reibungsproblem sind vereinfachte Modelle für Bingham-Fluide ebenfalls Beispiele für unrestringierte elliptische Minimierungsprobleme zweiter Art. Für die physikalischen Hintergründe sei auf [47], Ch. 6 verwiesen. Der wesentliche Unterschied zum Reibungsproblem (II.21) besteht in der Angabe des Funktionals $j : H_0^1(\Omega) \rightarrow \mathbb{R}$, das hier für ein $s \in L^2(\Omega)$ mit $s \geq 0$ als $j(v) := (s, |\nabla v|)_0$ definiert ist. Offenbar ist j konvex, womit nach Satz I.13 das zugehörige Minimierungsproblem, das in der Suche von

$$\min_{v \in H_0^1(\Omega)} \frac{1}{2} (\nabla v, \nabla v) - (f, v)_0 + j(v) \quad (\text{II.98})$$

besteht, genau eine Lösung $u \in H_0^1(\Omega)$ besitzt. Für diese ist die folgende Bedingung hinreichend und notwendig:

$$\forall v \in H_0^1(\Omega) : (\nabla u, \nabla(v - u))_0 + (s, |\nabla v|)_0 - (s, |\nabla u|)_0 \geq (f, v - u)_0.$$

Es sei $\Lambda_1 := \{\mu_1 \in L^2(\Omega)^k \mid |\mu_1| \leq s\}$, dann ist $j(v) \geq (|\mu_1|, |\nabla v|)_0 \geq (\mu_1, \nabla v)_0$ für $\mu_1 \in \Lambda_1$. Also gilt $j(v) \geq \sup_{\mu_1 \in \Lambda_1} (\mu_1, \nabla v)_0$. Andererseits ist $\mu_{1,v} \in \Lambda_1$ mit $\mu_{1,v} := s \nabla v / |\nabla v|$, $|\nabla v| \neq 0$, und damit $j(v) = (\mu_{1,v}, \nabla v)_0 \leq \sup_{\mu_1 \in \Lambda_1} (\mu_1, \nabla v)_0$. Also folgt insgesamt $j(v) = \sup_{\mu_1 \in \Lambda_1} (\mu_1, \nabla v)_0$. Damit erhält man durch den Gebrauch des Multiplikatorfunktionals $\Phi_1 : H_0^1(\Omega) \times \Lambda_1 \rightarrow \mathbb{R}$ mit $\Phi_1(v, \mu_1) := (\mu_1, \nabla v)_0$ eine variationelle Sattelpunktformulierung: Nach Satz I.14 ist $u \in H_0^1(\Omega)$ Lösung des Minimierungsproblems (II.98), wenn ein $\lambda_1 \in \Lambda_1$ existiert, so dass

$$\forall v \in H_0^1(\Omega) : (\nabla u, \nabla v)_0 = (f, v)_0 - (\lambda_1, \nabla v)_0 \quad (\text{II.99})$$

$$\forall \mu_1 \in \Lambda_1 : (\mu_1 - \lambda_1, \nabla u)_0 \leq 0 \quad (\text{II.100})$$

erfüllt ist. Da Λ_1 beschränkt ist, existiert nach Satz I.19 mindestens eine Lösung (u, λ_1) von (II.99)-(II.100), wobei u eindeutig ist.

³²vgl. Remark 9.2 in [55]

Kapitel III

Diskretisierungen elliptischer Minimierungsprobleme

Im Allgemeinen kann die analytische Lösung von elliptischen Minimierungsproblemen erster bzw. zweiter Art nicht direkt bestimmt werden. In der Praxis versucht man, die Lösung u als Grenzwert einer Folge von Approximationslösungen darzustellen. Diese erhält man durch Modifikation des Problems, indem man die Restriktionsmenge K durch konvexe und abgeschlossene Restriktionsmengen $K_h \subset V$ mit $h > 0$ austauscht und die Lösungen des modifizierten Minimierungsproblems bestimmt. Im Fall, dass die Restriktionsmengen K_h Teilmengen endlichdimensionaler Räume sind, heißt dieses Vorgehen im Sinne eines Ritz-Galerkin-Ansatzes auch Diskretisierung des Problems.

Wie für das Minimierungsproblem mit der Restriktionsmenge K sichert Satz I.11 (bzw. Satz I.13) die Existenz einer eindeutigen Folge $\{u_h\}$ mit $u_h \in K_h$, so dass u_h Lösung des elliptischen Minimierungsproblems erster Art (bzw. zweiter Art) mit K_h als Restriktionsmenge ist. Zudem ist hinreichend und notwendig für die Näherungslösungen u_h , dass

$$\forall v_h \in K_h : a(u_h, v_h - u_h) \geq \langle \ell, v_h - u_h \rangle \quad (\text{III.1})$$

$$\forall v_h \in K_h : a(u_h, v_h - u_h) + j(v_h) - j(u_h) \geq \langle \ell, v_h - u_h \rangle \quad (\text{III.2})$$

gilt. Gilt im Fall von elliptischen Minimierungsproblemen erster Art für die Restriktionsmenge $K_h = V_h$ mit einem linearen Unterraum $V_h \subset V$, ist die Variationsgleichung

$$\forall v_h \in K_h : a(u_h, v_h) = \langle \ell, v_h \rangle \quad (\text{III.3})$$

hinreichend und notwendig. Dass für die so gewonnene Folge von Näherungslösungen tatsächlich $\lim_{h \rightarrow 0} u_h = u$ gilt, hängt im Wesentlichen von weiteren Approximationseigenschaften der Mengen K_h ab.¹

Satz III.1 *Es sei V ein reflexiver Raum, $u \in V$ Lösung eines Minimierungsproblems erster Art (bzw. zweiter Art). Die Mengen $K_h \subset V$ mit $h > 0$ seien abgeschlossen und konvex, und es gelte:*

$$(i) \quad \forall v \in K : \forall h > 0 : \exists v_h \in K_h : \lim_{h \rightarrow 0} v_h = v.$$

$$(ii) \quad \forall \{v_h\}_{h>0} \subset V, v_h \in K_h : v_h \rightharpoonup v, h \rightarrow 0 \quad \Rightarrow \quad v \in K.$$

Weiter sei $\{u_h\} \subset V$, so dass u_h die Bedingung (III.1) (bzw. (III.2)) für alle $h > 0$ erfüllt. Dann gilt:

$$\lim_{h \rightarrow 0} u_h = u.$$

Beweis: Es sei $v \in K$ und $\{v_h\}$ gemäß (i). Aus (III.1) (bzw. (III.2)) sowie aus der Stetigkeit von a folgt

$$a(u_h, u_h) \leq a(u_h, v_h) - \langle \ell, v_h - u_h \rangle \leq \nu_1 \|u_h\| \|v_h\| + \|\ell\| (\|v_h\| + \|u_h\|)$$

bzw.

$$\begin{aligned} a(u_h, u_h) &\leq a(u_h, v_h) + j(v_h) - j(u_h) - \langle \ell, v_h - u_h \rangle \\ &\leq \nu_1 \|u_h\| \|v_h\| + j(v_h) + \|\mu\| \|u_h\| + \alpha + \|\ell\| (\|v_h\| + \|u_h\|), \end{aligned}$$

wobei $\mu \in V'$ und $\alpha \in \mathbb{R}$ wie in Lemma I.4 gewählt werden. Da $\{v_h\}$ beschränkt ist, ist wegen der Elliptizität von a auch die Folge $\{u_h\}$ beschränkt. Nach Satz A.3 existiert ein $w \in V$, so dass nach Übergang zu einer Teilfolge $u_h \rightharpoonup w$ für $h \rightarrow 0$ gilt. Aus (ii) erhält man $w \in K$. Aus $a(w, w) \leq \liminf_{h \rightarrow 0} a(u_h, u_h) \leq a(w, v) - \langle \ell, v - w \rangle$ bzw. $a(w, w) \leq$

¹vgl. Ch.1, Th.4.2 in [55]

$\liminf_{h \rightarrow 0} a(u_h, u_h) \leq a(w, v) - (f, v - w) + j(v) - j(w)$ und wegen der Eindeutigkeit der Lösung folgt schließlich $w = u$. Da

$$\begin{aligned} 0 &\leq \limsup_{h \rightarrow 0} a(u - u_h, u - u_h) \\ &\leq \limsup_{h \rightarrow 0} a(u_h, v_h) - (f, v_h - u_h) - a(u_h, u) - a(u_h, u_h - u) \\ &\leq a(u, v) - (f, v - u) - a(u, u) \end{aligned}$$

ist, folgt mit $v = u$, dass sogar $u_h \rightarrow u$ insbesondere ohne Übergang zu einer Teilfolge gilt. \square

Betrachtet man die zu elliptischen Minimierungsproblemen erster und zweiter Art gehörenden Sattelpunktformulierungen, so kann ebenfalls mittels Austausch von V und Λ durch V_h bzw. Λ_H die Lösung $(u, \lambda) \in V \times \Lambda$ als Grenzwert einer Folge von Approximationslösungen $(u_h, \lambda_H) \in V_h \times \Lambda_H$ mit $h, H > 0$ dargestellt werden. Hierbei ist mit den Bezeichnungen aus den Abschnitten I.3.1 und I.3.2 V_h ein reflexiver Raum und Λ_H je nach Typ des Minimierungsproblems entweder ein abgeschlossener und konvexer Kegel $\Lambda_{0,H} \subset U'_0$, eine abgeschlossene und konvexe Teilmenge $\Lambda_{1,H} \subset U'_1$ oder das Produkt von beidem.

Notwendig und hinreichend für die Approximationslösungen $(u_h, \lambda_{0,H})$, $(u_h, \lambda_{1,H})$ bzw. $(u_h, \lambda_{0,H}, \lambda_{1,H})$ ist im Fall von elliptischen Minimierungsproblemen erster Art²

$$\forall v_h \in V_h : a(u_h, v_h) = \langle \ell, v_h \rangle - \langle \lambda_{0,H}, \beta_0(v_h) \rangle \quad (\text{III.4})$$

$$\forall \mu_{0,H} \in \Lambda_{0,H} : \langle \mu_{0,H} - \lambda_{0,H}, \beta_0(u_h) - g \rangle \leq 0 \quad (\text{III.5})$$

und im Fall elliptischer Minimierungsprobleme zweiter Art mit $K_h = V_h$

$$\forall v_h \in V_h : a(u_h, v_h) = \langle \ell, v_h \rangle - \langle \lambda_{1,H}, \beta_1(v_h) \rangle \quad (\text{III.6})$$

$$\forall \mu_{1,H} \in \Lambda_{1,H} : \langle \mu_{1,H} - \lambda_{1,H}, \beta_1(u_h) \rangle \leq 0 \quad (\text{III.7})$$

bzw. mit $K_h \subset V_h$

$$\forall v_h \in V_h : a(u_h, v_h) = \langle \ell, v_h \rangle - \langle \lambda_{0,H}, \beta_0(v_h) \rangle - \langle \lambda_{1,H}, \beta_1(v_h) \rangle \quad (\text{III.8})$$

$$\forall (\mu_{0,H}, \mu_{1,H}) \in \Lambda_{0,H} \times \Lambda_{1,H} :$$

$$\langle \mu_{0,H} - \lambda_{0,H}, \beta_0(u_h) - g \rangle + \langle \mu_{1,H} - \lambda_{1,H}, \beta_1(u_h) \rangle \leq 0. \quad (\text{III.9})$$

²vgl. Satz I.12, Satz I.14 bzw. Satz I.15

Analog zu den Ausgangssystemen kann die Existenz und Eindeutigkeit von Lösungen der Systeme (III.4)-(III.5), (III.6)-(III.7) und (III.8)-(III.9) mit Hilfe der in Abschnitt I.3.3 hergeleiteten Aussagen nachgewiesen werden: Im Fall von System (III.4)-(III.5) gewährleistet Satz I.18 die Existenz einer Approximationsfolge $\{(u_h, \lambda_H)\}$, sofern $\text{rg}(\beta_{0|V_h})$ abgeschlossen in U_0 ist.³ Ist V_h endlichdimensional, ist dies der Fall, da $\beta_0(V_h)$ als endlichdimensionaler Unterraum abgeschlossen ist. Die erste Komponente u_h ist stets eindeutig.⁴ Nach Satz I.9 ist $\Phi_{0,hH} : V_h \times \Lambda_{0,H} \rightarrow \mathbb{R}$ mit $\Phi_{0,hH} := \langle \mu_{0,H}, \beta_0(v_h) - g \rangle$ ein Multiplikatorfunktional bzgl. $K_{hH} := \{v_h \in V_h \mid \forall \mu_{0,H} \in \Lambda_{0,H} : \langle \mu_{0,H}, \beta_0(v_h) \rangle \leq 0\}$. Demnach stimmt u_h mit der Lösung von (III.1) überein, wenn für $K_h := K_{hH}$ gewählt wird. Ist $U'_{0,H}$ mit $\Lambda_{0,H} \subset U'_{0,H} \subset U'_0$ ein reflexiver Raum, dann ist $\lambda_{0,H}$ nach Satz I.16 und Lemma I.7 eindeutig, wenn die Bedingung (I.40) oder (I.37) bzgl. $b_{0,hH} : V_h \times U'_{0,H} \rightarrow \mathbb{R}$ mit $b_{0,hH}(v_h, \mu_{0,H}) := \langle \mu_{0,H}, \beta_0(v_h) \rangle$ gilt.⁵

Für das System (III.6)-(III.7) existiert nach Satz I.19 eine Lösung $(u_h, \lambda_{1,H})$, wenn $\Lambda_{1,H}$ als beschränkt vorausgesetzt wird. Hier ist die erste Komponente u_h ebenfalls eindeutig und stimmt mit der Lösung von (III.2) überein, falls $K_h := V_h$ und $j = j_H : V \rightarrow \mathbb{R}$ mit $j_H(v) := \sup_{\mu_{1,H} \in \Lambda_{1,H}} \langle \mu_{1,H}, \beta_1(v) \rangle$ gewählt wird. Gilt (I.40) oder (I.37) bzgl. $b_{1,hH} : V_h \times U'_{1,H} \rightarrow \mathbb{R}$ mit $b_{1,hH}(v_h, \mu_{1,H}) := \langle \mu_{1,H}, \beta_1(v_h) \rangle$, so ist auch λ_1 eindeutig.

Schließlich besitzt das System (III.8)-(III.9) nach Satz I.17 eine Lösung $(u_h, \lambda_{0,H}, \lambda_{1,H})$, wenn $\text{rg}(\beta_{0|V_h})$ abgeschlossen in U_0 ist, was für endlichdimensionale Räume V_h der Fall ist, und wenn $\Lambda_{1,H}$ beschränkt ist. Die Komponente u_h ist eindeutig und stimmt mit der Lösung von (III.2) mit $K_h := K_{hH}$ und $j := j_H$ überein. Ist (I.40) oder (I.37) bzgl. $b_{hH} : V_h \times U'_{0,H} \times U'_{1,H} \rightarrow \mathbb{R}$ mit $b_{hH}(v_h, \mu_{0,H}, \mu_{1,H}) := \langle \mu_{0,H}, \beta_0(v_h) \rangle + \langle \mu_{1,H}, \beta_1(v_h) \rangle$ erfüllt, dann sind auch $\lambda_{0,H}$ und $\lambda_{1,H}$ eindeutig.

Die Konvergenz der Näherungsfolge $\{(u_h, \lambda_H)\}$ gegen eine Lösung (u, λ) hängt einerseits von Approximationseigenschaften der zugrunde liegenden Mengen V_h und Λ_H ab, die von ähnlicher Gestalt wie in Satz III.1 sind, und andererseits von weiteren Bedingungen, wie zum Beispiel von der Gültigkeit der Bedingung (I.37) bzgl. b_{hH} für ein von h und H unabhängiges α . Für weitere Details sei hier auf [62], [64], [67] und [68] verwiesen.

Die in Satz III.1 angegebenen Voraussetzungen sind Forderungen an die Räume V und V_h bzw. an die Restriktionsmengen K und K_h und beinhal-

³Zu beachten ist, dass $\beta_{0|V_h}$ im Allgemeinen nicht surjektiv ist.

⁴vgl. Satz I.18, Satz I.19 bzw. Satz I.20

⁵Die Bedingung (I.37) heißt in diesem Fall diskrete Babuška-Brezzi-Bedingung.

ten keine weiteren Regularitätsannahmen an die Lösung u . Der Nachteil von Konvergenzaussagen dieses Typs ist jedoch, dass in der Regel keine qualitative Bewertung der Konvergenzgüte z.B. in Form von Konvergenzordnungen möglich ist.

Ist dagegen die Beschaffenheit bzw. die Regularität der Lösung a-priori bekannt, können häufig qualitative Konvergenzaussagen in Form von oberen Fehlerschranken entwickelt werden. Für elliptische Minimierungsprobleme erhält man diese in der Regel über Verallgemeinerungen des sogenannten Céa-Lemmas. Das Céa-Lemma ist in der Literatur⁶ im Wesentlichen für Variationsgleichungen bekannt: Aus der Subtraktion von (I.23) und (III.3) erhält man die sogenannte Galerkin-Orthogonalität

$$\forall v_h \in V_h : a(u - u_h, v_h) = 0, \quad (\text{III.10})$$

aus der unmittelbar zusammen mit der Elliptizität von a folgt:

$$\nu_1 \|u - u_h\|^2 \leq a(u - u_h, u - u_h) = a(u - u_h, u - v_h) \leq \nu_0 \|u - u_h\| \|u_h - v_h\|$$

Demnach gilt:

Lemma III.1 *Es sei $u \in V$ Lösung von (I.23) und $u_h \in V_h$ Lösung von (III.3). Dann existiert ein $C' > 0$, so dass gilt:⁷*

$$\|u - u_h\| \leq C' \inf_{v_h \in V_h} \|u - v_h\|.$$

Um in ähnlicher Weise Fehlerabschätzungen für restringierte elliptische Minimierungsprobleme erster und zweiter Art zu gewinnen, ist eine Verallgemeinerung des Céa-Lemmas erforderlich. Mit den Bezeichnungen aus Abschnitt I.4 folgt, sofern $u \in \text{dom}(A)$ ist, aus (I.22) und (III.1) sowie aus der Youngschen Ungleichung⁸ für $v_h \in K_h$

$$\begin{aligned} & a(u - u_h, u - u_h) \\ &= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\ &\leq a(u - u_h, u - v_h) + a(u, v_h - u_h) - (f, v_h - u_h) - \langle q, \beta(v_h - u_h) \rangle \end{aligned}$$

⁶vgl. Lemma 4.2 in [25]

⁷Es sei bemerkt, dass derartige (generische) Konstanten unabhängig von dem gewählten FE-Gitter bzw. von den nachfolgenden Ausdrücken sind.

⁸Für $a, b \geq 0$ und $\epsilon > 0$ gilt:

$$ab \leq (\epsilon/2)a^2 + (2\epsilon)^{-1}b^2 \quad (\text{III.11})$$

$$\begin{aligned}
&= a(u - u_h, u - v_h) + (f - A(u), u_h - v_h) + \langle q - \alpha(u), u_h - v_h \rangle \\
&\leq \frac{\epsilon}{2} \|u - u_h\|^2 + \frac{\nu_0^2}{2\epsilon} \|u - v_h\|^2 + (f - A(u), u_h - v_h) + \langle q - \alpha(u), u_h - v_h \rangle.
\end{aligned}$$

Aus der Elliptizität von a erhält man hieraus für $0 < \epsilon < 2\nu_1$

$$(\nu_1 - \frac{\epsilon}{2}) \|u - u_h\|^2 \leq \frac{\nu_0^2}{2\epsilon} \|u - v_h\|^2 + (f - A(u), u_h - v_h) + \langle q - \alpha(u), u_h - v_h \rangle$$

und damit für restringierte elliptische Minimierungsprobleme erster Art:

Lemma III.2 *Es sei $u \in \text{dom}(A)$ Lösung von (I.22) und $u_h \in V_h$ Lösung von (III.1). Dann existieren $C' > 0$ und $C'' > 0$, so dass gilt:*

$$\begin{aligned}
\|u - u_h\|^2 \leq \inf_{v_h \in K_h} (C' \|u - v_h\|^2 \\
+ C'' ((f - A(u), u_h - v_h) + \langle q - \alpha(u), \beta(u_h - v_h) \rangle)).
\end{aligned}$$

In analoger Weise zeigt man für (un-)restringierte elliptische Minimierungsprobleme zweiter Art:

Lemma III.3 *Es sei $u \in \text{dom}(A)$ Lösung von (I.31) und $u_h \in V_h$ Lösung von (III.2). Dann existieren $C' > 0$ und $C'' > 0$, so dass gilt:*

$$\begin{aligned}
\|u - u_h\|^2 \leq \inf_{v_h \in K_h} (C' \|u - v_h\|^2 \\
+ C'' ((f - A(u), u_h - v_h) + \langle q - \alpha(u), \beta(u_h - v_h) \rangle + j(v_h) - j(u_h))).
\end{aligned}$$

Ein adäquates Verfahren zur Gewinnung von Approximationslösungen $\{u_h\}$ bzw. $\{(u_h, \lambda_H)\}$ in Bezug auf Variationsprobleme, deren Ansatzräume den in Abschnitt II.1 beschriebenen Sobolevräumen entsprechen, ist die Finite-Elemente-Methode, kurz FEM. Ihre Vorzüge bestehen in einer hohen Flexibilität bezüglich Geometrieadjustierungen und in einer effizienten Implementierbarkeit. Eine Vielzahl von Varianten gewährleistet im hohen Maße einen problemangepassten Einsatz der FEM.

In den Abschnitten III.1 und III.2 wird eine Variante der FEM vorgestellt, die insbesondere für die in Kapitel II beschriebenen Problemklassen geeignet ist. Zudem werden einige teilweise in der Literatur bekannte und im Wesentlichen auf den Lemmata III.1-III.3 beruhende qualitative Konvergenzaussagen aufgeführt.

Der Einsatz von FEM-Diskretisierungen führt in der Regel auf quadratische Minimierungsprobleme im \mathbb{R}^n . In Abschnitt III.3 wird auf Möglichkeiten zur Lösung von Problemen dieses Typs eingegangen.

III.1 Finite-Elemente-Methoden

Der zentrale Ansatz der FEM besteht in der Konstruktion von endlichdimensionalen Teilräumen mit Hilfe von lokal definierten Polynomräumen, mit denen, wie oben beschrieben, eine Diskretisierung des Variationsproblems realisiert wird.

Ausgangspunkt hiervon ist stets die Wahl eines geeigneten Gitters $\mathcal{T} := \{T_0, \dots, T_n\}$ für ein Gebiet $\Omega \subset \mathbb{R}^k$ mit $k \in \mathbb{N}$ und einer darauf definierten Polynomgradverteilung $p : \mathcal{T} \rightarrow \mathbb{N}$. Für \mathcal{T} gelte stets $\bigcup_{T \in \mathcal{T}} \bar{T} = \bar{\Omega}$. Die Elemente des Gitters sind offen, und es gilt, dass $T_i \cap T_j = \emptyset$ für $i \neq j$ ist. Für $k = 2$ sind die Elemente in der Regel Dreiecke oder Vierecke und für $k = 3$ Tetraeder, Hexaeder, Pyramiden oder Prismen. Im Folgenden werden Gitter von Polyedern beliebiger Dimension betrachtet, die anhand des k -dimensionalen Hyperwürfels $[-1, 1]^k$ konstruiert werden. Praxisrelevant sind diesbezüglich Gitter aus Vierecken und Hexaedern.

Das Gitter \mathcal{T} (bzw. eine Gitterfolge) heißt regulär, falls $\bar{T}_i \cap \bar{T}_j$ entweder leer ist oder genau einen Eckpunkt, eine Kante, eine Facette usw. von T_i enthält. Die Paarung (\mathcal{T}, p) wird im Folgenden FE-Gitter genannt.

Wichtige Eigenschaften von Gittern bzw. von Folgen von Gittern ergeben sich aus dem Verhältnis von Maschenweite und Durchmesser der Gitterelemente. Die Maschenweite eines Gitters \mathcal{T} ist definiert als $h := \max_{T \in \mathcal{T}} h_T$, wobei $h_T := \sup_{x, y \in T} \|x - y\|_2$ der maximale Durchmesser des Elements T ist. Eine Folge von Dreiecks-, Tetraeder oder Parallelogramm- bzw. Parallelepipid-Gittern $\{\mathcal{T}_h\}$ heißt nicht-degeneriert⁹, falls ein $\rho_0 > 0$ existiert, so dass

$$\forall h > 0 : \forall T \in \mathcal{T}_h : r_T \geq \rho_0 h_T$$

gilt, wobei $r_T := \sup\{r \in \mathbb{R}_{>0} \mid B_r(x) \subset T, x \in T\}$ ist.¹⁰ Die Folge $\{\mathcal{T}_h\}$ heißt quasi-uniform, falls es ein $\rho_1 > 0$ gibt, so dass gilt:

$$\forall h > 0 : \min\{r_T \mid T \in \mathcal{T}_h\} \geq \rho_1 h.$$

Im Folgenden werden Gitter stets als \mathcal{T}_h mit der Maschenweite h als Index notiert, und es wird zwischen $\{\mathcal{T}_h\}$ und \mathcal{T}_h bzw. $\{(\mathcal{T}_h, p)\}$ und (\mathcal{T}_h, p) nicht mehr unterschieden.

Lokale Polynomräume, basierend auf einem FE-Gitter (\mathcal{T}_h, p) vom oben beschriebenen Typ, werden durch Transformation des Tensorproduktraums

⁹vgl. Def.4.4.13 in [27]. Für eine entsprechende Erweiterung auf Vierecke bzw. Hexaeder, die keine Parallelepipede sind, vgl. [79]

¹⁰ $B_r(x) := \{y \in \mathbb{R}^k \mid \|x - y\|_2 \leq r\}$

$S_k^q := \otimes_{i=1}^k S^q$ mit

$$S^q := \left\{ \hat{v} : [-1, 1] \rightarrow \mathbb{R} \mid \hat{v}(\xi) = \sum_{0 \leq i \leq q} c_i \xi^i, c_i \in \mathbb{R} \right\}$$

konstruiert.¹¹ Einen endlichdimensionalen Teilraum von $H^1(\Omega, \Gamma_0)$ erhält man mit Hilfe bijektiver und hinreichend glatter, für jedes Gitterelement $T \in \mathcal{T}_h$ definierter Abbildungen $\Psi_T : [-1, 1]^k \rightarrow T$ in der folgenden Form:¹²

$$S^p(\mathcal{T}_h) := \{v \in H^1(\Omega, \Gamma_0) \mid \forall T \in \mathcal{T}_h : v|_T \circ \Psi_T \in S_k^{p,T}\}. \quad (\text{III.12})$$

Für eine Diskretisierung bzgl. der Lagrangeschen Multiplikatoren benötigt man einen endlichdimensionalen Teilraum von $L^2(\Gamma_1)$. Hierzu sei \mathcal{T}_1 ein Gitter von Γ_1 und $p_1 : \mathcal{T}_{1,H} \rightarrow \mathbb{N}$ eine Polynomgradverteilung. Für $T \in \mathcal{T}_{1,H}$ sei $\Psi_{1,T} : [-1, 1]^{k-1} \rightarrow T$ bijektiv und hinreichend glatt. Man definiert dann:

$$M^{p_1}(\mathcal{T}_{1,H}) := \{v \in L^2(\Gamma_1) \mid \forall T \in \mathcal{T}_{1,H} : v|_T \circ \Psi_{1,T} \in S_{k-1}^{p_1,T}\}.$$

Offenbar hängt die Güte der Approximation des Raums $H^1(\Omega, \Gamma_0)$ durch den Raum $S^p(\mathcal{T}_h)$ von der Feinheit der Maschenweite h und von der gewählten Polynomgradverteilung p ab. Wird nur die Maschenweite h verringert, um die Approximationseigenschaft des Diskretisierungsansatzes zu verbessern, spricht man von sogenannten h -Methoden. Diskretisierungsansätze heißen p -Methoden, wenn nur der Polynomgrad erhöht wird. Wird sowohl die Maschenweite verkleinert als auch der Polynomgrad erhöht, spricht man von hp -Methoden.

III.1.1 Modellprobleme

Mit den Bezeichnungen aus Abschnitt II.2 sei (\mathcal{T}_h, p) ein FE-Gitter bzgl. des polygonal berandeten Gebiets $\Omega \subset \mathbb{R}^2$. Ferner sei $u_h \in S^p(\mathcal{T}_h)$ die eindeutig bestimmte Lösung von

$$\forall v_h \in S^p(\mathcal{T}_h) : (\nabla u_h, \nabla v_h) = (f, v_h)_0 + (q, \gamma(v_h))_0. \quad (\text{III.13})$$

¹¹Das Tensorprodukt zweier Räume V und W , deren Elemente Funktionen $f : A \rightarrow \mathbb{R}$ bzw. $g : B \rightarrow \mathbb{R}$ sind, ist definiert als $V \otimes W := \{fg : A \times B \rightarrow \mathbb{R} \mid f \in V, g \in W\}$.

¹²Insbesondere sind auch krummlinig begrenzte Gitterelemente etwa mit Hilfe isoparametrischer Transformationen oder Blending-Techniken möglich, vgl. S.111 in [25] bzw. S.107 in [99].

Im Folgenden sei für $m \in \mathbb{N}$ und $0 \leq \theta \leq 1$ der Interpolationsraum¹³ $H^{m+\theta}(\Omega)$ definiert als $H^{m+\theta}(\Omega) := [H^m(\Omega), H^{m+1}(\Omega)]_{\theta,2}$. Im Fall $m \geq 1$ sei $H^{m+\theta}(\Gamma_1) := [H^m(\Gamma_1), H^{m+1}(\Gamma_1)]_{\theta,2}$.

Für quasi-uniforme Parallelogramm-Gitter kann dann gezeigt werden:

Lemma III.4 *Es sei \mathcal{T}_h eine reguläre, quasi-uniforme Gitterfolge aus Parallelogrammen und p konstant sowie $m + \theta > 3/2$. Dann existieren ein Operator $I : H^1(\Omega, \Gamma_0) \cap H^{m+\theta}(\Omega) \rightarrow S^p(\mathcal{T}_h)$ und ein $C_I > 0$, so dass gilt:*

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma_0) \cap H^{m+\theta}(\Omega) : \\ \|v - I(v)\|_1 \leq C_I h^{\min\{p+1, m+\theta\}-1} / p^{m+\theta-1} \|v\|_{m+\theta}. \end{aligned}$$

Gilt zudem $\gamma(u)|_{\Gamma_1} \in H^{m+\theta-k}(\Gamma_1)$ mit $k \leq 1/2$, dann existiert ein $\tilde{C}_I > 0$, so dass für alle $0 \leq t \leq 1$ gilt:

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma_0) \cap H^{m+\theta}(\Omega) : \\ \|\gamma(v) - \gamma(I(v))\|_{t, \Gamma_1} \leq \tilde{C}_I h^{\min\{p+1, m+\theta-k\}-t} / p^{m+\theta-k-t} \|\gamma(v)\|_{m+\theta-k, \Gamma_1}. \end{aligned}$$

Beweis: Th.4.6, Lemma 4.3 in [16]. \square

Zusammen mit dem Céa-Lemma (III.1) erhält man daraus unmittelbar:

Satz III.2 *Es gelten die Voraussetzungen von Lemma III.4, wobei u Lösung von (II.6) sei. Zudem sei u_h Lösung von (III.13). Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 \leq C h^{\min\{p+1, m+\theta\}-1} / p^{m+\theta-1} \|u\|_{m+\theta}.$$

Im Fall einer einspringenden Ecke kann die in Satz III.2 dargestellte Abschätzung verfeinert werden. In diesem Fall besitzt die Lösung u die Darstellung

$$u(\phi, r) = \sum_{i=1}^n \sum_{j=0}^m \sum_{s=0}^t c_{ijs} \psi_{ijs}(\phi) r^{\alpha_i+s} \log^j r + \bar{u}(\phi, r) \quad (\text{III.14})$$

bezüglich der einspringenden Ecke $z \in \partial\Omega$. Hierbei sind (r, ϕ) die Polarkoordinaten mit Ursprung z , die Funktionen ψ_{ijs} sind stückweise analytisch und \bar{u} ist von höherer Regularität als der restliche Ausdruck. Für weitere Details sei auf [17] verwiesen.

¹³vgl. Satz A.9

Satz III.3 *Es sei u Lösung von (II.6) mit der Darstellung (III.14) und u_h Lösung von (III.13). Zudem sei \mathcal{T}_h eine reguläre, quasi-uniforme Gitterfolge aus Parallelogrammen und p konstant. Dann existiert ein $C > 0$, so dass mit $\alpha := \min \alpha_i$ gilt:*

$$\|u - u_h\|_1 \leq C \max(|\log h|^t, |\log p|^t) \min\left(h^\alpha, h^{\min(\alpha, p-\alpha)} / p^{2\alpha}\right)$$

Beweis: Th.5.4 in [16], Th.3.3 in [17]. \square

Aus den Sätzen III.2 und III.3 geht hervor, dass die Konvergenzordnung bei Verwendung quasi-uniformer Gitterfolgen abhängig von der Regularität der Lösung ist. Ohne weitere Maßnahmen kann die Konvergenzordnung $\mathcal{O}(h^p)$ bei h -Methoden nicht erwartet werden, wenn die Lösung $u \notin H^{p+1}(\Omega)$ ist. Speziell bei Lösungen der Form (III.14) können jedoch durch gezielte lokale Verfeinerungen im Bereich der Ecke z Konvergenzverbesserungen erzielt werden, so dass die Konvergenzordnung $\mathcal{O}(h^p)$ trotz mangelnder Regularität erreicht wird.

Für $\tau_0, \tau_1 > 0$, $i = 1, \dots, \lfloor \tau_0/h \rfloor$ und $Z \geq 1$ sei $R_{i,h} := \tau_1 i^Z h^Z$, $d_{i,h} := R_{i+1,h} - R_{i,h}$, und $\mathcal{T}_{h,i} := \{T \in \mathcal{T}_h \mid \exists x \in \bar{T} : \|x - z\|_2 \leq R_{i,h}\}$. Die Mengen $\mathcal{T}_{h,i}$ beschreiben, wie etwa in Abbildung III.1 zu sehen, eine von der Maschenweite abhängige Anzahl von konzentrischen Kreisen um z mit den Radien $R_{i,h}$.

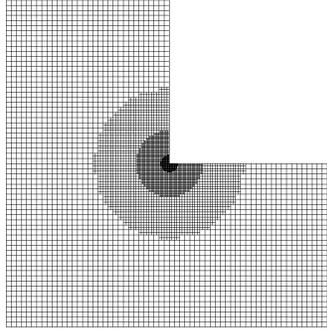


Abb. III.1: Gitter aus einer radialen Gitterfolge mit $Z = 2$, $\tau_0 = 0.07$, $\tau_1 = 100$.

Eine Gitterfolge \mathcal{T}_h heißt radial, wenn ein $\kappa \geq 1$ existiert, so dass gilt:

$$\forall i \in \{1, \dots, \lfloor \tau_0/h \rfloor\} : \forall T \in \mathcal{T}_{h,i} \setminus \mathcal{T}_{h,i-1} : \kappa^{-1} d_{i,h} \leq h_T \leq \kappa d_{i,h}.$$

Satz III.4 *Es sei u Lösung von (II.6) mit der Darstellung (III.14) und u_h Lösung von (III.13). Zudem sei \mathcal{T}_h eine radiale Gitterfolge mit $\min \alpha_i Z > 0$ und p konstant. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 \leq Ch^p$$

Beweis: Th.4.1 in [23]. □

Da die Anzahl der Gitterelemente asymptotisch der Größenordnung $\mathcal{O}(h^{-2})$ entspricht, erhält man durch die Verwendung von radialen Gitterfolgen tatsächlich die gewünschte Konvergenzverbesserung.

Nach Satz III.2 konvergiert die p -Methode für eine analytische Lösung u exponentiell, da das Wachstum von $\|u\|_p$ maximal $\mathcal{O}(p!)$ für $p \rightarrow \infty$ beträgt. Dagegen konvergiert die p -Methode nur noch algebraisch, wenn dies nicht der Fall ist.

Gilt für die Lösung u mit $\alpha \geq 1$ und $0 < \beta < 1$

$$|D^\alpha u| \leq C \left(r^{|\alpha|+\beta-1} \right)^{-1} d^{|\alpha|} \alpha! \quad (\text{III.15})$$

bezüglich der Ecke z , kann durch lokale Verfeinerungen im Bereich der Ecke z und durch eine spezielle Wahl der Polynomgradverteilung exponentielle Konvergenz erzielt werden.¹⁴ Hierzu sei $0 < \sigma < 1$ und $i \in \mathbb{N}$, sowie für $2 \leq j \leq i$

$$\begin{aligned} \widehat{T}_{1,1}^{i,\sigma} &:= (0, \sigma^{i-1})^2, & \widehat{T}_{1,j}^{i,\sigma} &:= (\sigma^{i-j+1}, \sigma^{i-j})^2, \\ \widehat{T}_{2,j}^{i,\sigma} &:= (\sigma^{i-j+1}, \sigma^{i-j}) \times (0, \sigma^{i-j+1}), & \widehat{T}_{3,j}^{i,\sigma} &:= (0, \sigma^{i-j+1}) \times (\sigma^{i-j+1}, \sigma^{i-j}). \end{aligned}$$

Es sei \mathcal{T} ein Gitter aus Vierecken bzgl. Ω und für $T \in \mathcal{T}_z := \{T \in \mathcal{T} \mid z \in \overline{T}\}$ sei $\widehat{\Psi}_T : [0, 1]^2 \rightarrow \overline{T}$ eine bijektive Abbildung mit $\widehat{\Psi}_T(0, 0) = z$. Zudem sei $\mathcal{T}_1^{i,\sigma} := \{\widehat{\Psi}_T(\widehat{T}_{1,1}^{i,\sigma}) \mid T \in \mathcal{T}_z\}$ und $\mathcal{T}_j^{i,\sigma} := \{\widehat{\Psi}_T(\widehat{T}_{r,j}^{i,\sigma}) \mid T \in \mathcal{T}_z, 1 \leq r \leq 3\}$ für $2 \leq j \leq i$. Dann heißt die FE-Gitterfolge (\mathcal{T}_h, p) mit $\mathcal{T}_h := \mathcal{T} \setminus \mathcal{T}_z \cup \bigcup_{1 \leq j \leq i} \mathcal{T}_j^{i,\sigma}$ und $\kappa > 0$ sowie

$$p(T) := \begin{cases} 1, & T \in \mathcal{T}_1^{i,\sigma} \\ \max\{2, \lfloor \kappa j \rfloor\}, & T \in \mathcal{T}_j^{i,\sigma}, 2 \leq j \leq i \\ i, & T \in \mathcal{T} \setminus \mathcal{T}_z \end{cases}$$

für $i = 1, 2, \dots$ geometrisch.

¹⁴Eine solche Darstellung gilt z.B. für Lösungen der Form (III.14), vgl. [17].

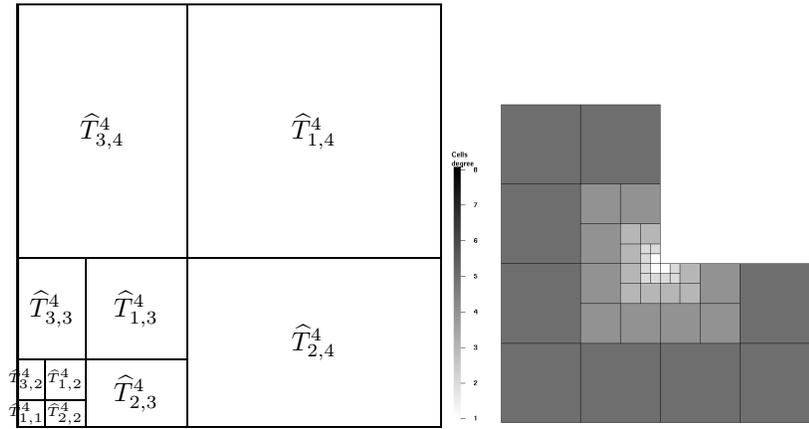


Abb. III.2: $\widehat{T}_{r,j}^{4,\sigma}$ mit $1 \leq r \leq 3$, $1 \leq j \leq 4$ und $\sigma = \frac{2}{5}$, Gitter aus geometrischer Gitterfolge mit $\sigma = 0.5$ und $\kappa = 1$ (hp3 - vgl. Abb. III.5).

Satz III.5 *Es sei u Lösung von (II.6) mit der Darstellung (III.15) und u_h Lösung von (III.13). Zudem sei (\mathcal{T}_h, p) eine geometrische FE-Gitterfolge. Dann existieren $C_0 > 0$ und $C_1 > 0$, so dass gilt:*

$$\|u - u_h\|_1 \leq C_0 \exp\left(-C_1 \sqrt[3]{\dim(S^p(\mathcal{T}_h))}\right)$$

Beweis: Ch.4.4 in [89], Th.3.2 in [17]. □

Die Aussagen der Sätze III.3, III.4 und III.5 können für eine beliebige Anzahl Ecken verallgemeinert werden. Zudem können Lösungen mit den Eigenschaften (III.14) und (III.15) als Funktionen in gewichteten Sobolevräumen betrachtet werden. Diesbezüglich erhält man entsprechende Verallgemeinerungen der genannten Sätze. Hierzu sei auf [10], [11], [13], [16] und [89] verwiesen. In [12] werden schließlich Aussagen des obigen Typs auch für nicht polygonal berandete Gebiete angegeben.

In den Abbildungen III.3 und III.4 sind Gitter aus geometrischen Gitterfolgen mit $\sigma = 0.15$ und $\sigma = 0.5$ sowie $\kappa = 1$ abgebildet. Zur Vermeidung irregulärer Knoten und Kanten sind zusätzliche Gitterelemente eingefügt. Die Verwendung derartiger Abfangelemente ist für den Fall $\sigma \neq 0.5$ notwendig, da im Allgemeinen irreguläre Knoten nur für symmetrische Teilungen

behandelt werden.¹⁵

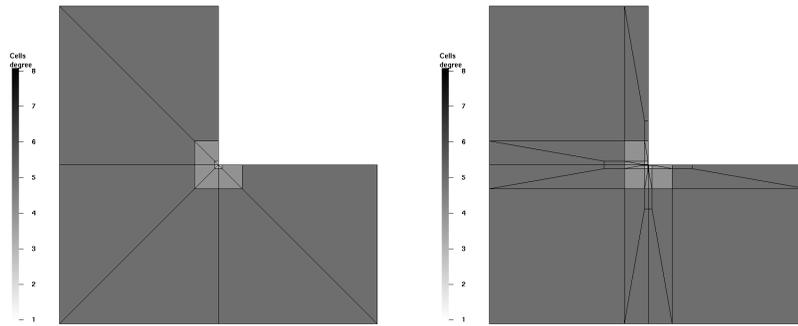


Abb. III.3: Gitter aus geometrischer Gitterfolge mit $\sigma = 0.15$ und $\kappa = 1$ (hp1, hp2 - vgl. Abb. III.5).

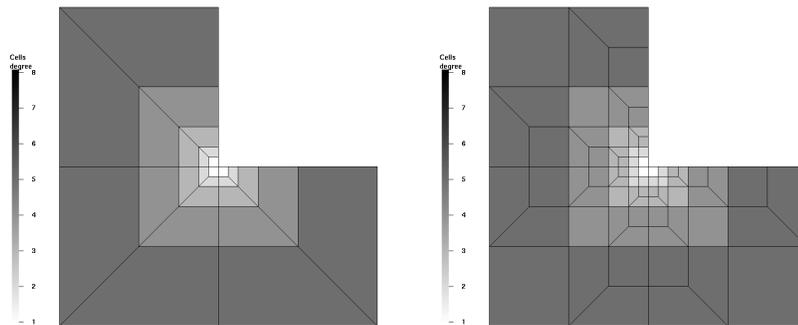


Abb. III.4: Gitter aus geometrischer Gitterfolge mit $\sigma = 0.5$ und $\kappa = 1$ (hp4, hp5 - vgl. Abb. III.5).

In Abbildung III.5 (links) ist das Konvergenzverhalten bzgl. der in den Abbildungen III.2 (rechts), III.3 und III.4 geometrischen FE-Gitterfolgen abgebildet. Die analytische Lösung des Modellproblems ist

$$u(r, \phi) = r^{2/3} \sin((2\phi - \pi)/3) \tag{III.16}$$

auf $\Omega := (-0.5, 0.5)^2 \setminus [0, 0.5]^2$ mit $\Gamma_0 := \{0\} \times [0, 0.5] \cup [0, 0.5] \times \{0\}$. Erkennbar sind vor allem die Auswirkungen bezüglich der Wahl des Grading-

¹⁵vgl. Abschnitt III.2.2

Factors σ bei der Verwendung von geometrischen Gitterfolgen. Die Wahl $\sigma = 0.15$ ist häufig aus praktischer Sicht günstig, was durch Abbildung III.5 bestätigt wird. Tatsächlich kann zumindest für den eindimensionalen Fall nachgewiesen werden, dass $\sigma = (\sqrt{2} - 1)^2 \approx 0.17$ optimal ist.¹⁶

Abbildung III.5 (rechts) zeigt das Konvergenzverhalten bei Verwendung einer quasi-uniformen Gitterfolge mit $p = 1$ (h), eines konstanten Gitters mit $p = 1, 2, 3, \dots$ (p), einer geometrischer FE-Gitterfolge wie in Abbildung III.3 (hp1) und zweier radialer Gitterfolgen mit $p = 1$ und $p = 2$ (rd1, rd2). Deutlich erkennbar ist die hohe Konvergenzrate bei Verwendung geometrischer FE-Gitterfolgen gegenüber den anderen Vorgehensweisen. Die Aussagen der Sätze III.3 und III.4 werden bestätigt.

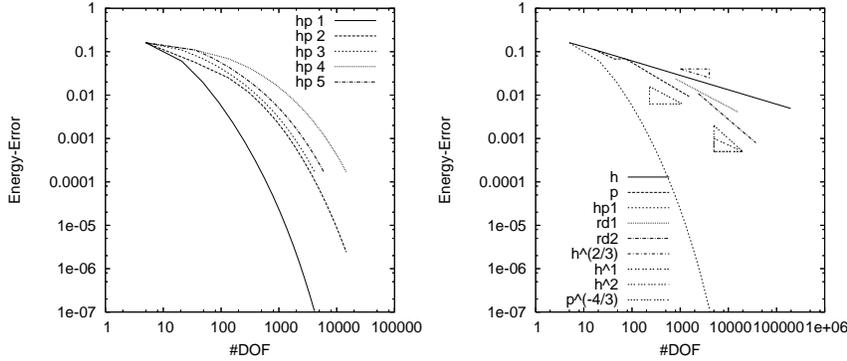


Abb. III.5: H^1 -Fehler bei Verwendung von geometrischen FE-Gitterfolgen, H^1 -Fehler bzgl. h -, p - und hp -Methoden (vgl. Abbildungen III.2, III.3 und III.4).

Im Folgenden wird eine Diskretisierung des vereinfachten Signorini-Problems aus Abschnitt II.2.1 betrachtet. Hierzu seien $K_h \subset S^p(\mathcal{T}_h)$ abgeschlossen und konvex. Außerdem sei $u_h \in K_h$ die eindeutig bestimmte Lösung von

$$\forall v_h \in K_h : (\nabla u_h, \nabla(v_h - u_h))_0 \geq (f, v_h - u_h)_0 + (q, \gamma(v_h - u_h))_{0, \Gamma_1}. \quad (\text{III.17})$$

Zur Sicherung der Konvergenz der Folge $\{u_h\}$ gegen die Lösung u von (II.11) sind die Voraussetzungen von Satz III.1 zu prüfen. Diese sind etwa im Fall $p \equiv 1$ und

$$K_h := \{v_h \in S^p(\mathcal{T}_h) \mid \forall x \in \mathcal{V} \cap \Gamma_1 : v_h(x) \geq g(x)\} \quad (\text{III.18})$$

¹⁶vgl. [9]

und im Fall $p \equiv 2$ und $K_h := \{v_h \in S^p(\mathcal{T}_h) \mid \forall x \in \tilde{\mathcal{V}}_h \cap \Gamma_1 : v_h(x) \geq g(x)\}$ erfüllt.¹⁷ Hierbei bezeichnet \mathcal{V} die Menge aller Knoten in \mathcal{T}_h ohne Γ_0 . Die Menge $\tilde{\mathcal{V}}$ beinhaltet alle Knoten und alle Kantenmittelpunkte.

Mit Hilfe von Lemma III.2 zeigt man für $p \equiv 1$:

Satz III.6 *Es sei $u \in H^1(\Omega, \Gamma_0) \cap H^2(\Omega)$ Lösung von (II.11) und u_h Lösung von (III.17) mit K_h wie in (III.18) sowie $\gamma(u), g \in W^{1,\infty}(\Gamma_1)$.¹⁸ Zudem sei \mathcal{T}_h eine reguläre und quasi-uniforme Gitterfolge aus Parallelogrammen und $p \equiv 1$. Die Menge der Punkte von Γ_1 , an denen die Bedingung $u > g$ zu $u = g$ (und umgekehrt) wechselt, sei endlich. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 \leq Ch.$$

Beweis: Th.6.2 in [28]. □

Die zu (III.17) gehörende variationelle Sattelpunktformulierung lautet: Gesucht ist ein $(u_h, \lambda_{0,H}) \in S^p(\mathcal{T}_h) \times \Lambda_0 \cap M^{p_1}(\mathcal{T}_{1,H})$, so dass gilt:

$$\forall v_h \in S^p(\mathcal{T}_h) :$$

$$(\nabla u_h, \nabla v_h)_0 = (f, v_h)_0 + (q, \gamma(v_h))_{0,\Gamma_1} - (\lambda_{0,H}, \gamma(v_h))_{0,\Gamma_1} \quad (\text{III.19})$$

$$\forall \mu_{0,H} \in \Lambda_0 \cap M^{p_1}(\mathcal{T}_{1,H}) : (\mu_{0,H} - \lambda_{0,H}, \gamma(u_h) - g)_{0,\Gamma_1} \leq 0, \quad (\text{III.20})$$

wobei $\mathcal{T}_{1,H}$ eine Gitterfolge von Γ_1 ist.

Ein wesentlicher Grund für die Einführung von variationellen Sattelpunktformulierungen besteht einerseits darin, dass die Verwendung von K_h in der obigen Form als Approximation von K in der Regel algorithmisch nachteilig ist.¹⁹ Dies gilt insbesondere für höhere Polynomgrade. Andererseits entsprechen die in der Sattelpunktformulierung vorkommenden Lagrangeschen Multiplikatoren physikalisch interpretierbare Größen, an denen man ebenfalls interessiert ist. Zwar ist die Existenz von Diskretisierungslösungen von Sattelpunktformulierungen gemäß den oben stehenden Ausführungen stets gesichert, jedoch geht im Allgemeinen die Interpretierbarkeit der Lagrangeschen Multiplikatoren als physikalische Größen verloren, wenn deren Eindeutigkeit nicht gewährleistet ist. Die Diskretisierung einer variationellen Sattelpunktformulierung heißt in diesem Zusammenhang stabil, wenn die eindeutige Lösbarkeit der Diskretisierungslösung gesichert

¹⁷vgl. Ch.4 Th.3.2 in [55]

¹⁸zur Def. von $W^{p,\infty}(\Gamma_1)$ s. [1]

¹⁹vgl. Abschnitt III.3

ist. Darüber hinaus kann der Stabilitätsbegriff in dem Sinne erweitert werden, dass gleichmäßig stetige Abhängigkeit der Lösungen von den Daten unabhängig vom Diskretisierungsniveau bestehen muss.

Hinreichend für die Eindeutigkeit der Lösung $(u_h, \lambda_{0,H})$ gemäß Lemma I.6 ist die Bedingung (I.40) mit $b_{hH} : S^p(\mathcal{T}_h) \times M^{p_1}(\mathcal{T}_{1,H}) \rightarrow \mathbb{R}$ und $b_{hH}(v_h, \lambda_{0,H}) := (\lambda_{0,H}, \gamma(v_h))_{0,\Gamma_1}$ oder laut Lemma I.7 die schärfere diskrete Babuška-Brezzi-Bedingung

$$\begin{aligned} \exists \alpha \in \mathbb{R}_{>0} : \forall \mu_{0,H} \in M^{p_1}(\mathcal{T}_{1,H}) : \\ \alpha \|\mu_{0,H}\|_{-1/2} \leq \sup_{v_h \in S^p(\mathcal{T}_h) \setminus \{0\}} \frac{(\mu_{0,H}, \gamma(v_h))_{0,\Gamma_1}}{\|v_h\|_1}. \end{aligned} \quad (\text{III.21})$$

In jedem Fall wird deutlich, dass die Eindeutigkeit der Diskretisierungslösung nicht gesichert werden kann, wenn die Räume $S^p(\mathcal{T}_h)$ und $M^{p_1}(\mathcal{T}_{1,H})$ zueinander nicht richtig ausbalanciert sind. Die Balance der Räume $S^p(\mathcal{T}_h)$ und $M^{p_1}(\mathcal{T}_{1,H})$ kann über die Maschenweiten h und H sowie über die Polynomgradverteilungen p und p_1 reguliert werden²⁰, wobei darauf zu achten ist, dass die Approximationseigenschaften der Räume hinreichend gut sind. Kann α in (III.21) unabhängig von h , H , p und p_1 gewählt werden, ist die Diskretisierung im oben genannten Sinn stabil.

Für den Nachweis von (III.21) kann das in [65] (Lemma 3.1) aufgeführte Dualitätsargument verallgemeinert werden: Für ein $\mu_0 \in H^{-1/2}(\Gamma_1)$ sei $v_{\mu_0} \in H^1(\Omega, \Gamma_0)$ Lösung von

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla v_{\mu_0}, \nabla v)_{0,\Omega} + (v_{\mu_0}, v)_{0,\Omega} = \langle \mu_0, \gamma(v) \rangle. \quad (\text{III.22})$$

Das Problem (III.22) heißt regulär, falls ein $0 < \theta < 1$ existiert, so dass $v_{\mu_0} \in H^1(\Omega, \Gamma_0) \cap H^{1+\theta}(\Omega)$ für alle $\mu_0 \in H^{-1/2+\theta}(\Gamma_1) := [H^{-1/2}(\Gamma_1), H^{1/2}(\Gamma_1)]_{\theta,2}$ ist und zusätzlich $\|v_{\mu_0}\|_{1+\theta} \leq C \|\mu_0\|_{-1/2+\theta}$ gilt. Zudem ist eine inverse Abschätzung erforderlich: Für $r \geq 1$ gelte:²¹

$$\forall \mu_{0,H} \in M^{p_1}(\mathcal{T}_{1,H}) : \|\mu_{0,H}\|_{-1/2+\theta} \leq C \max\{1, p_1\}^{r\theta} / H^\theta \|\mu_{0,H}\|_{-1/2}. \quad (\text{III.23})$$

²⁰Weitere Möglichkeiten zur Stabilisierung ergeben sich z.B. durch Hinzufügen oder Entfernen von Funktionen aus $S^p(\mathcal{T}_h)$ bzw. $M^{p_1}(\mathcal{T}_{1,H})$. Stabilisierungstechniken sind vor allem im Bereich der Navier-Stokes Gleichungen und bei Gleichungen zur Beschreibung inkompressibler Materialien bekannt. Diesbezüglich sei z.B. auf Ch.6 in [25] verwiesen.

²¹Mit Blick auf bekannte, approximationstheoretische Resultate bzgl. H^1 - und L^2 -Normen (vgl. z.B. [75] oder Th.4.76 in [89]) kann von $r = 2$ ausgegangen werden.

Satz III.7 *Es seien \mathcal{T}_h und $\mathcal{T}_{1,H}$ reguläre, quasi-uniforme Gitterfolgen sowie p und p_1 konstant. Das Problem (III.22) sei regulär, und es gelte die inverse Abschätzung (III.23). Ist $\max\{1, p_1\}^r/p \cdot h/H$ konstant und hinreichend klein, dann ist (III.21) mit einem von h , H , p und p_1 unabhängigen α erfüllt.*

Beweis: Es sei $\mu_{0,H} \in M^{p_1}(\mathcal{T}_{1,H})$. Dann ist

$$\begin{aligned} \sup_{v_h \in S^p(\mathcal{T}_h) \setminus \{0\}} \frac{(\mu_{0,H}, \gamma(v_h))_{0,\Gamma_1}}{\|v_h\|_1} &\geq \frac{(\mu_{0,H}, \gamma(v_{\mu_{0,H},h}))_{0,\Gamma_1}}{\|v_{\mu_{0,H},h}\|_1} = \|v_{\mu_{0,H},h}\|_1 \\ &\geq \|v_{\mu_{0,H}}\|_1 - \|v_{\mu_{0,H}} - v_{\mu_{0,H},h}\|_1 = \|\mu_{0,H}\|_{-1/2} - \|v_{\mu_{0,H}} - v_{\mu_{0,H},h}\|_1. \end{aligned} \quad (\text{III.24})$$

Aus Satz III.2, der Regularität von (III.22) und der inversen Abschätzung (III.23) folgt

$$\begin{aligned} \|v_{\mu_{0,H}} - v_{\mu_{0,H},h}\|_1 &\leq Ch^\theta/p^\theta \|v_{\mu_{0,H}}\|_{1+\theta} \leq Ch^\theta/p^\theta \|\mu_0\|_{-1/2+\theta} \\ &\leq C(\max\{1, p_1\}^r/p \cdot (h/H))^\theta \|\mu_0\|_{-1/2}. \end{aligned}$$

Zusammen mit (III.24) folgt dann die Behauptung. \square

Die Aussage aus Satz III.7 ist aus praktischer Sicht unbefriedigend, da im Voraus nicht klar ist, wie klein $\max\{1, p_1\}^r/p \cdot h/H$ tatsächlich zu wählen ist. Außerdem ist im Allgemeinen das Erfülltsein der weiteren Voraussetzungen von Satz III.7 ungeklärt.²²

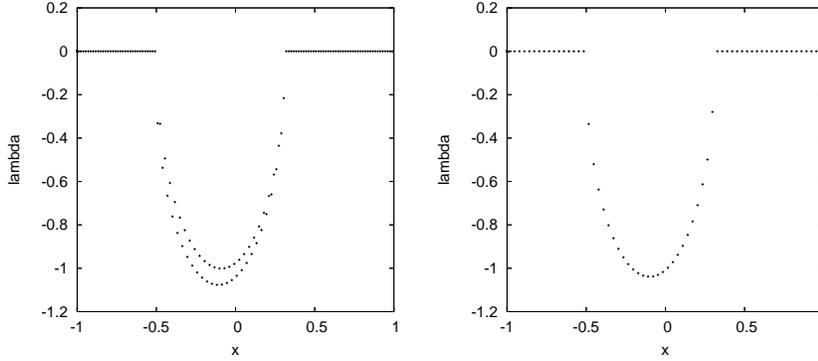


Abb. III.6: $\lambda_{0,H}$ mit $p \equiv 1$, $p_1 \equiv 0$ und $h/H = 1$ bzw. $h/H = 0.5$.

²²vgl. Diskussion der Aussage von Satz III.7 im Ausblick

In Abbildung III.6 ist $\lambda_{0,H}$ bei den Polynomgradverteilungen $p \equiv 1$ und $p_1 \equiv 0$ zu sehen, wobei Ω , f , q und g wie im Beispiel aus Abschnitt II.2 gewählt wird. Bei $h/H = 1$ ist ein für instabile Diskretisierungen typisches Muster zu erkennen, das unter dem Namen Schachbrett-Instabilität²³ bekannt und insbesondere hinsichtlich Lemma I.6 zu deuten ist. Die Lösung $\lambda_{0,H}$ ist in diesem Fall für physikalische Interpretationen im Allgemeinen nicht verwertbar. Zur Stabilisierung kann das Verhältnis h/H gemäß Satz III.7 verkleinert werden. Für $h/H = 0.5$ sind die Instabilitätsmuster nicht mehr zu sehen.

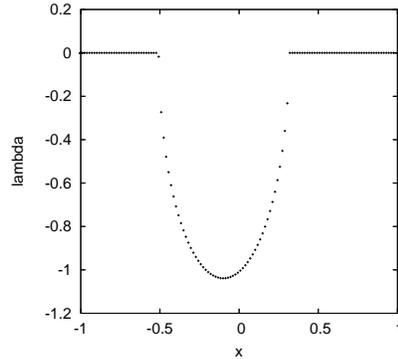


Abb. III.7: $\lambda_{0,H}$ mit $p \equiv 2$, $p_1 \equiv 1$ und $h/H = 1$.

In Abbildung III.7 ist $\lambda_{0,H}$ für die Kombination $p \equiv 2$, $p_1 \equiv 1$ und $h/H = 1$ zu sehen. Im Unterschied zur obigen Konfiguration ist der Lagrangesche Multiplikator $\lambda_{0,H}$ sinnvoll interpretierbar.

Die Kombination $p \equiv 3$, $p_1 \equiv 2$ und $h/H = 1$ führt laut Abbildung III.8 (links) zu einer Lösung $\lambda_{0,H}$ mit dem typischen Instabilitätsmuster. Dagegen hat die Lösung $\lambda_{0,H}$ bei der Wahl $h/H = 0.5$ einen sinnvollen Verlauf (rechts), womit die Diskretisierung als wahrscheinlich stabil anzunehmen ist. Weitere Experimente ergeben, dass die Kombination $h/H = 1$, $p_1 \equiv p - 1$ für gerade Polynomgrade p zu sinnvoll interpretierbaren Lösungen führt, während dies bei ungeraden Polynomgraden erst durch die Wahl $h/H = 0.5$ gelingt.

²³vgl. Abschnitt III.1.2

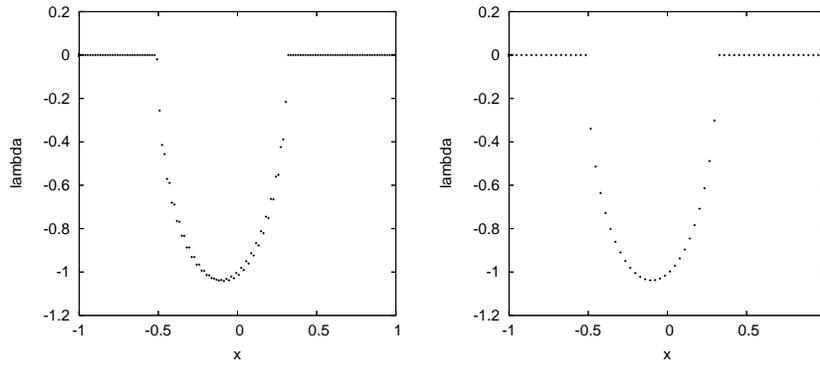


Abb. III.8: $\lambda_{0,H}$ mit $p \equiv 3$, $p_1 \equiv 2$ und $h/H = 1$ bzw. $h/H = 0.5$.

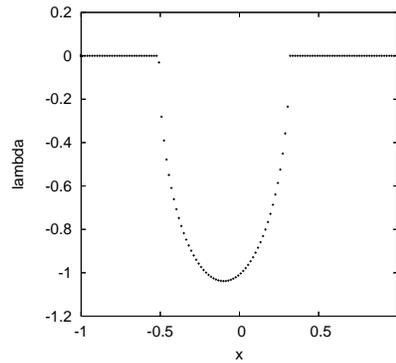


Abb. III.9: $\lambda_{0,H}$ mit $p \equiv 3$, $p_1 \equiv 1$ und $h/H = 1$.

Dass eine Reduktion des Quotienten $\max\{1, p_1\}^r/p$ ebenfalls zu einer Stabilisierung führen kann, wird in Abbildung III.9 deutlich. Hier führt die Kombination $h/H = 1$, $p \equiv 3$ und $p_1 \equiv 1$ anscheinend zu einer stabilen Diskretisierung.

Abschließend ist jedoch festzuhalten, dass das Vorhandensein oder das Fehlen von Schachbrett-Instabilitäten höchstens als ein Indiz für Instabilität bzw. Stabilität zu werten ist.

Für die Kombination $p \equiv 1$ und $p_1 \equiv 0$ kann eine qualitative Konvergenz-

aussage gezeigt werden:

Satz III.8 *Es sei (u, λ_0) Lösung von (II.19)-(II.20) und $(u_h, \lambda_{0,H})$ Lösung von (III.19)-(III.20). Es sei $\gamma(u) \in H^{1,\infty}(T_1)$ für alle $T_1 \in \mathcal{T}_{1,H}$, und die Anzahl der Übergangspunkte zwischen $\gamma(u)$ und g sei endlich.²⁴ Zudem seien $\mathcal{T}_h, \mathcal{T}_H$ regulär und quasi-uniform und h/H konstant und hinreichend klein sowie $p \equiv 1$ und $p_1 \equiv 0$. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 + \|\lambda_0 - \lambda_{0,H}\|_{-1/2} \leq Ch.$$

Beweis: Th.4.1 in [65]. \square

Betrachtet wird nun das idealisierte Reibungsproblem aus Abschnitt II.2.2. Hierzu sei $u_h \in S^p(\mathcal{T}_h)$ Lösung von

$$\begin{aligned} \forall v_h \in S^p(\mathcal{T}_h) : (\nabla u_h, \nabla(v_h - u_h))_0 + (s, |\gamma(v_h)|)_{0,\Gamma_1} - (s, |\gamma(u_h)|)_{0,\Gamma_1} \\ \geq (f, v_h - u_h)_0 + (q, \gamma(v_h - u_h))_{0,\Gamma_1}. \end{aligned} \quad (\text{III.25})$$

Auch hier sind zur Sicherung der Konvergenz der Folge $\{u_h\}$ gegen die Lösung u von (II.22) die Voraussetzungen von Satz III.1 zu prüfen.²⁵

Eine qualitative Konvergenzaussage für konstantes p erhält man unter gewissen Interpolationsvoraussetzungen:²⁶ Es existiere ein Interpolationsoperator $I : H^1(\Omega, \Gamma_0) \cap H^{m+\theta}(\Omega) \rightarrow S^p(\mathcal{T}_h)$ sowie $C_I > 0$ und $\tilde{C}_I > 0$, so dass mit $t, k \in \mathbb{R}$ gilt:

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma_0) \cap H^{m+\theta}(\Omega) : \\ \|v - I(v)\|_1 \leq C_I h^{\min\{p+1, m+\theta\}-1} / p^{m+\theta-1} \|v\|_{m+\theta} \end{aligned} \quad (\text{III.26})$$

$$\|\gamma(v) - \gamma(I(v))\|_{t,\Gamma_1} \leq \tilde{C}_I h^{\min\{p+1, m+\theta-k\}-t} / p^{m+\theta-k-t} \|\gamma(v)\|_{m+\theta-k,\Gamma_1} \quad (\text{III.27})$$

Satz III.9 *Es sei $u \in H^1(\Omega, \Gamma_0) \cap H^{m+\theta}(\Omega)$ Lösung von (II.22) mit $m+\theta \geq 2$ und u_h Lösung von (III.25). Ferner sei $\gamma(u)|_{\Gamma_1} \in H^{m+\theta-k}(\Gamma_1)$ sowie $s|_{\Gamma_1}, (q - \partial_n u)|_{\Gamma_1} \in H^{-t}(\Gamma_1)$, und es gelte (III.26) und (III.27). Zudem sei \mathcal{T}_h eine reguläre und quasi-uniforme Gitterfolge aus Parallelogrammen und p konstant. Dann existiert ein $C > 0$, so dass gilt:*

$$\begin{aligned} \|u - u_h\|_1 \leq C \max\{h^{\min\{p+1, m+\theta\}-1} / p^{m+\theta-1}, \\ h^{0.5(\min\{p+1, m+\theta-k\}-t)} / p^{0.5(m+\theta-k-t)}\}. \end{aligned}$$

²⁴zur Definition von $H^{1,\infty}$ s. [1]

²⁵vgl. Ch.4, Th.3.2 in [55]

²⁶vgl. Th.10.5 in [72]

Beweis: Aus (II.25), (II.29) und (III.27) sowie der Hölderschen Ungleichung folgt

$$\begin{aligned}
& (q - \partial_n u, \gamma(u_h - I(u)))_{0, \Gamma_1} + (s, |\gamma(I(u))|)_{0, \Gamma_1} - (s, |\gamma(u_h)|)_{0, \Gamma_1} \\
& \leq (s, |\gamma(I(u))|)_{0, \Gamma_1} - (q - \partial_n u, \gamma(I(u)))_{0, \Gamma_1} + (|q - \partial_n u| - s, |\gamma(u_h)|)_{0, \Gamma_1} \\
& \leq (s, |\gamma(I(u))|)_{0, \Gamma_1} - (q - \partial_n u, \gamma(I(u)))_{0, \Gamma_1} \\
& = (s, |\gamma(I(u))| - |\gamma(u)|)_{0, \Gamma_1} - (q - \partial_n u, \gamma(I(u) - u))_{0, \Gamma_1} \\
& \leq (\|s\|_{-t, \Gamma_1} + \|q - \partial_n u\|_{-t, \Gamma_1}) \|\gamma(u) - \gamma(I(u))\|_{t, \Gamma_1} \\
& \leq (\|s\|_{-t, \Gamma_1} + \|q - \partial_n u\|_{-t, \Gamma_1}) \\
& \quad \tilde{C}_I h^{\min\{p+1, m+\theta-k\}-t} / p^{m+\theta-k-t} \|\gamma(v)\|_{m+\theta-k, \Gamma_1}
\end{aligned}$$

Aus Lemma III.3, (II.28) und (III.26) erhält man damit

$$\begin{aligned}
& \|u - u_h\|_1^2 \\
& \leq C' \|u - I(u)\|_1^2 + C'' ((q - \partial_n u, \gamma(u_h - I(u)))_{0, \Gamma_1} \\
& \quad + (s, |\gamma(I(u))|)_{0, \Gamma_1} - (s, |\gamma(u_h)|)_{0, \Gamma_1}) \\
& \leq 2 \max \{ C' C_I \|u\|_2^2, C'' \tilde{C}_I (\|s\|_{-t, \Gamma_1} + \|q - \partial_n u\|_{-t, \Gamma_1}) \|\gamma(u)\|_{m+\theta-k, \Gamma_1} \} \\
& \quad \max \{ h^{2(\min\{p+1, m+\theta\}-1)} / p^{2(m+\theta-1)}, h^{\min\{p+1, m+\theta-k\}-t} / p^{m+\theta-k-t} \}.
\end{aligned}$$

□

Korollar III.1 *Es gelten die Voraussetzungen von Satz III.9 mit $k := 1/2$ und $t := -(m + \theta) + 3/2$ sowie $h \leq 1$ und $p \geq m + \theta - 3/2$. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 \leq C h^{\min\{p+1, m+\theta\}-1} / p^{m+\theta-1}.$$

Der Nachweis der in Korollar III.1 geforderten Interpolationsvoraussetzungen ist nicht ohne weiteres gegeben, da Interpolationsabschätzungen insbesondere für negative Normen zu erbringen sind.

Dagegen erhält man mit Hilfe von Lemma III.4 ein u.U. nicht optimales Resultat:

Korollar III.2 *Es gelten die Voraussetzungen von Satz III.9 mit $k := 0$ und $t := 0$ sowie $h^{0.5(\min\{p+1, m+\theta\}-1)} / p^{0.5(m+\theta)-1} \leq 1$. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 \leq C h^{0.5 \min\{p+1, m+\theta\}} / p^{0.5(m+\theta)}.$$

Eine zu (III.25) gehörende variationelle Sattelpunktformulierung lautet: Gesucht ist $(u_h, \lambda_{1,H}) \in S^p(\mathcal{T}_h) \times \Lambda_1 \cap M^{p_1}(\mathcal{T}_{1,H})$, so dass gilt:

$$\forall v_h \in S^p(\mathcal{T}_h) : \quad (\nabla u_h, \nabla v_h)_0 = (f, v_h)_0 + (q, \gamma(v_h))_{0,\Gamma_1} - (\lambda_{1,H}, \gamma(v_h))_{0,\Gamma_1} \quad (\text{III.28})$$

$$\forall \mu_{1,H} \in \Lambda_1 \cap M^{p_1}(\mathcal{T}_{1,H}) : (\mu_{1,H} - \lambda_{1,H}, \gamma(u_h) - g)_{0,\Gamma_1} \leq 0. \quad (\text{III.29})$$

Wie oben ist die Eindeutigkeit der Lösung $(u_h, \lambda_{1,H})$ gesichert, wenn die dazu gehörige diskrete Babuška-Brezzi-Bedingung erfüllt ist, die in diesem Fall mit (III.21) übereinstimmt. Demnach ist die Lösung nach Satz III.7 eindeutig, wenn $\max\{1, p_1\}^r \cdot h/H$ konstant und hinreichend klein ist.

In Abbildung III.10 ist jeweils $\lambda_{1,H}$ bei verschiedenen Quotienten h/H und Polynomgradverteilungen p und p_1 zu sehen. Das hierzu verwendete Beispiel entspricht der Konfiguration aus Abschnitt II.2.2. Da die diskrete Babuška-Brezzi-Bedingung (III.21) für die Systeme (III.19)-(III.20) und (III.28)-(III.29) identisch ist, sind diese Abbildungen in gleicher Weise zu interpretieren.

Auch eine zu Satz III.8 entsprechende Aussage bezüglich des Systems (III.28)-(III.29) kann gewonnen werden. Hierzu sei auf [62] verwiesen.

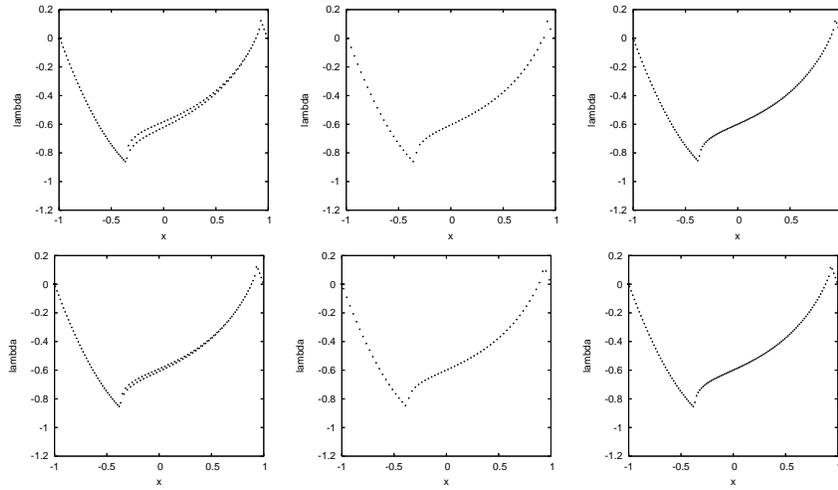


Abb. III.10: $\lambda_{1,H}$ mit $p \equiv 1$, $p_1 \equiv 0$ und $h/H = 1$ bzw. $h/H = 0.5$ sowie mit $p \equiv 2$, $p_1 \equiv 1$ und $h/H = 1$ (oben), $\lambda_{1,H}$ mit $p \equiv 3$, $p_1 \equiv 2$ und $h/H = 1$ bzw. $h/H = 0.5$ sowie mit $p \equiv 3$, $p_1 \equiv 1$ und $h/H = 1$ (unten).

III.1.2 Linear-elastische Probleme

Für Variationsgleichungen, die unrestringierten Problemen der linearen Elastizität zugeordnet sind, lautet die Diskretisierung: Gesucht ist ein $u_h \in S^p(\mathcal{T}_h)^3$, so dass

$$\forall v_h \in S^p(\mathcal{T}_h)^3 : (\sigma(u_h), \varepsilon(v_h))_0 = (f, v_h)_0 + (q, \gamma(v_h))_{0, \Gamma_1} \quad (\text{III.30})$$

erfüllt ist. Konvergenzresultate, die den für das Modellproblem (III.13) genannten Aussagen entsprechen, können auf analoge Weise gewonnen werden. Für einen Überblick sei hierzu auf [17], [18], [99] und [101] sowie auf darin enthaltene Referenzen verwiesen. Unterschiede ergeben sich insbesondere durch eine wesentlich kompliziertere Klassifikation der Regularität von Lösungen²⁷, da das zugrunde liegende Gebiet Ω als dreidimensional vorausgesetzt wird. Werden in der Modellierung Ansätze zur Dimensionsreduktion eingearbeitet, sind in der Regel sogenannte Lockingeffekte zu berücksichtigen. Lockingeffekte können ebenfalls durch die eingehenden Modellparameter bedingt sein.²⁸

Für reibungsfreie Kontaktprobleme aus Abschnitt II.3.1 erhält man eine Diskretisierung in der folgenden Form: Gesucht ist ein $u_h \in K_h$, so dass gilt:

$$\forall v_h \in K_h : (\sigma(u_h), \varepsilon(v_h - u_h))_0 \geq (f, v_h - u_h)_0 + (q, \delta(v_h - u_h))_{0, \Gamma_1}, \quad (\text{III.31})$$

wobei $K_h \subset S^p(\mathcal{T}_h)^3$ eine abgeschlossene und konvexe Menge ist. Wie bei dem entsprechenden Modellproblem (II.11) kann für eine geeignet definierte Menge K_h gezeigt werden, dass die Bedingungen aus Satz III.1 erfüllt sind, so dass $\lim_{h \rightarrow 0} u_h = u$ ist.²⁹ In [72] (Th.6.4) wird analog zu Satz III.6 zudem gezeigt, dass $\|u - u_h\| \leq Ch$ für $p \equiv 1$ gilt.

Die zu (III.31) gehörende variationelle Sattelpunktformulierung lautet: Gesucht ist $(u_h, \lambda_{n,H}) \in S^p(\mathcal{T}_h)^3 \times \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H})$, so dass gilt:

$$\begin{aligned} \forall v_h \in S^p(\mathcal{T}_h)^3 : \\ (\sigma(u_h), \varepsilon(v_h))_0 = (f, v_h)_0 + (q, \delta(v_h))_{0, \Gamma_1} - (\lambda_{n,H}, \delta_n(v_h))_{0, \Gamma_1} \end{aligned} \quad (\text{III.32})$$

$$\forall \mu_{n,H} \in \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H}) : (\mu_{n,H} - \lambda_{n,H}, \delta_n(u_h) - g)_{0, \Gamma_1} \leq 0. \quad (\text{III.33})$$

²⁷vgl. z.B. [39]

²⁸vgl. [17],[25]

²⁹vgl. S. 127 in [72]

Wie zuvor ist die Diskretisierung stabil, wenn eine entsprechende diskrete Babuška-Brezzi-Bedingung mit einer von der Maschenweite und der Polynomgradverteilung unabhängigen Konstanten erfüllt ist. Eine zu Satz III.7 analoge Aussage kann unter Verwendung eines adäquaten Dualitätsarguments³⁰ erzielt werden.

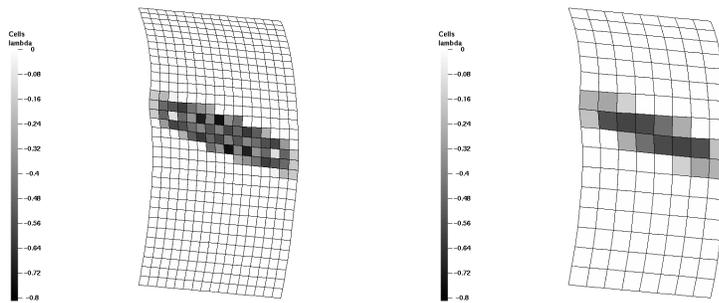


Abb. III.11: $-\lambda_{n,H}$ mit $p \equiv 1$, $p_1 \equiv 0$ und $h/H = 1$ bzw. $h/H = 0.5$.

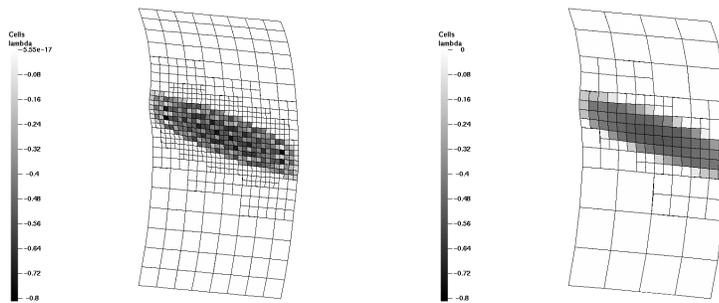


Abb. III.12: Adaptives Gitter, sonst wie in Abb. III.11.

Die Abbildungen III.11 und III.12 zeigen $-\lambda_{n,H}$ im Fall $p \equiv 1$ und $p_1 \equiv 0$. Deutlich erkennbar sind die für Instabilitäten typischen Schachbrettmuster, falls $h/H = 1$ ist. Im Fall $h/H = 0.5$ sind diese Muster nicht mehr zu sehen. Den in Abbildung III.12 dargestellten Lösungen ist ein adaptives

³⁰vgl. Def.4.1 und Lemma 4.3 in [64]

Gitter zugrunde gelegt, um eine höhere Auflösung im Bereich $\lambda_{n,H} > 0$ zu erreichen. Für weitere Details bzgl. Adaptivität sei auf Abschnitt IV.7 verwiesen. Im Fall $p \equiv 2$ und $p_1 \equiv 1$ sind im Gegensatz zum entsprechenden Modellproblem für $h/H = 1$ Muster zu erkennen, die auf eine Instabilität schließen lassen. Bei der Wahl $h/H = 0.5$ sind diese nicht auszumachen (vgl. Abbildung III.13).

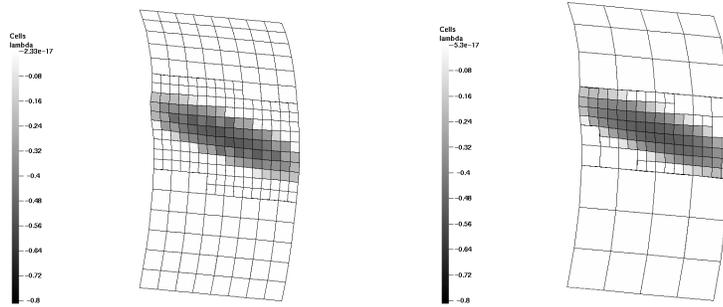


Abb. III.13: Adaptives Gitter, $-\lambda_{n,H}$ mit $p \equiv 2$, $p_1 \equiv 1$ und $h/H = 1$ bzw. $h/H = 0.5$.

Eine Diskretisierung von reibungsbehafteten Kontaktproblemen mit vorgegebenen Normalspannungen ist durch folgende Formulierung gegeben: Gesucht ist ein $u_h \in S^p(\mathcal{T}_h)^3$, so dass

$$\begin{aligned} \forall v_h \in S^p(\mathcal{T}_h)^3 : (\sigma(u_h), \varepsilon(v_h - u_h))_0 + (s, |\delta_t(v_h)|)_{0,\Gamma_1} - (s, |\delta_t(u_h)|)_{0,\Gamma_1} \\ \geq (f, v_h - u_h)_0 + (q, \delta(v_h - u_h))_{0,\Gamma_1} \quad (\text{III.34}) \end{aligned}$$

erfüllt ist. Eine qualitative Konvergenzaussage zu (III.34) ist in [72] (Th. 10.5) enthalten, sie entspricht im Wesentlichen der Aussage von Satz III.9 für das idealisierte Reibungsproblem.

Die Diskretisierung der zugehörigen variationellen Sattelpunktformulierung lautet: Gesucht ist ein $(u_h, \lambda_{1,H}) \in S^p(\mathcal{T}_h)^3 \times \Lambda_t \cap M^{p_1}(\mathcal{T}_{1,H})^2$, so dass gilt:

$$\begin{aligned} \forall v \in S^p(\mathcal{T}_h)^3 : \\ (\sigma(u_h), \varepsilon(v_h))_0 = (f, v_h)_0 + (q, \delta(v_h))_{0,\Gamma_1} - (\lambda_{t,H}, \delta_t(v_h))_{0,\Gamma_1} \quad (\text{III.35}) \end{aligned}$$

$$\forall \mu_{t,H} \in \Lambda_{t,H} \cap M^{p_1}(\mathcal{T}_{1,H})^2 : (\mu_{t,H} - \lambda_{t,H}, \delta_t(u_h))_{0,\Gamma_1} \leq 0. \quad (\text{III.36})$$

Die Frage nach der Stabilität der Diskretisierung kann in ähnlicher Weise wie oben diskutiert werden. Letztlich ist auch hier das Verhältnis von p_1 zu p bzw. h zu H entscheidend.

Für Kontaktprobleme mit Reibung erhält man die folgende Diskretisierung:
Für eine konvexe und abgeschlossene Menge K_h sei $u_h \in K_h$ Lösung von

$$\begin{aligned} \forall v_h \in K_h : (\sigma(u_h), \varepsilon(v_h - u_h))_0 + (s, |\delta_t(v_h)|)_{0,\Gamma_1} - (s, |\delta_t(u_h)|)_{0,\Gamma_1} \\ \geq (f, v_h - u_h)_0 + (q, \delta(v_h - u_h))_{0,\Gamma_1}. \end{aligned} \quad (\text{III.37})$$

Die zugehörige diskretisierte variationelle Sattelpunktformulierung lautet:
Gesucht ist $(u_h, \lambda_{0,H}, \lambda_{1,H}) \in S^p(\mathcal{T}_h)^3 \times \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H}) \times \Lambda_t \cap M^{p_1}(\mathcal{T}_{1,H})^2$,
so dass gilt:

$$\begin{aligned} \forall v_h \in S^p(\mathcal{T}_h)^3 : (\sigma(u_h), \varepsilon(v_h))_0 = (f, v_h)_0 + (q, \delta(v_h))_{0,\Gamma_1} \\ - (\lambda_{n,H}, \delta_n(v_h))_{0,\Gamma_1} - (\lambda_{t,H}, \delta_t(v_h))_{0,\Gamma_1} \end{aligned} \quad (\text{III.38})$$

$$\begin{aligned} \forall (\mu_{n,H}, \mu_{t,H}) \in \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H}) \times \Lambda_t \cap M^{p_1}(\mathcal{T}_{1,H})^2 : \\ (\mu_{n,H} - \lambda_{n,H}, \delta_n(v_h) - g)_{0,\Gamma_1} + (\mu_{t,H} - \lambda_{t,H}, \delta_t(u_h))_{0,\Gamma_1} \leq 0. \end{aligned} \quad (\text{III.39})$$

Für die Stabilität des Verfahrens ist die Gültigkeit der diskreten Babuška-Brezzi-Bedingung hinreichend, die in diesem Fall die Gestalt hat:

$$\begin{aligned} \exists \alpha \in \mathbb{R}_{>0} : \forall (\mu_{0,H}, \mu_{1,H}) \in M^{p_1}(\mathcal{T}_{1,H}) \times M^{p_1}(\mathcal{T}_{1,H})^2 : \\ \alpha \|(\mu_{n,H}, \mu_{t,H})\|_{-1/2} \\ \leq \sup_{v_h \in S^p(\mathcal{T}_h) \setminus \{0\}} \frac{(\mu_{n,H}, \delta_n(v_h))_{0,\Gamma_1} + (\mu_{t,H}, \delta_t(v_h))_{0,\Gamma_1}}{\|v_h\|_1}. \end{aligned} \quad (\text{III.40})$$

Für eine Satz III.7 entsprechende Aussage sei für den Fall $p \equiv 1$ und $p_1 \equiv 0$ auf [64] und [66] verwiesen. Für den hier betrachteten dreidimensionalen Fall wird in [66] gezeigt:

Satz III.10 *Es sei $(u, \lambda_n, \lambda_t)$ Lösung von (II.88)-(II.89), und es sei $(u_h, \lambda_{n,H}, \lambda_{t,H})$ Lösung von (III.38)-(III.39). Zudem seien $\mathcal{T}_h, \mathcal{T}_H$ reguläre und quasi-uniforme Gitterfolgen und h/H konstant. Es sei $p \equiv 1$ und $p_1 \equiv 0$, und es gelte (III.40) mit einem von h und H unabhängigen α . Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 + \|(\lambda_n, \lambda_t) - (\lambda_{n,H}, \lambda_{t,H})\|_{-1/2,nt} \leq Ch^{1/2}.$$

Beweis: Th.5.1 in [66]. □

III.1.3 Weitere Modellprobleme

Eine Diskretisierung des Hindernisproblems (II.91) aus Abschnitt II.5.1 erhält man folgendermaßen: Für einen abgeschlossenen und konvexen Kegel $K_h \subset S^p(\mathcal{T}_h)$ sei $u_h \in K_h$ die eindeutig bestimmte Lösung von

$$\forall v_h \in K_h : (\nabla u_h, \nabla(v_h - u_h))_0 \geq (f, v_h - u_h)_0 + (q, \gamma(v_h - u_h))_{0, \Gamma_1}. \quad (\text{III.41})$$

Werden Dreiecks- statt Vierecksgitter \mathcal{T}_h vorausgesetzt³¹ sind Konvergenzaussagen in der Literatur bekannt: Resultate, die die Konvergenz der Folge $\{u_h\}$ gegen die Lösung u von (II.91) gemäß Satz III.1 sicherstellen, sind z.B. in [56] (Th.2.3) enthalten. Im Fall $p \equiv 1$ und

$$K_h := \{v_h \in S^p(\mathcal{T}_h) \mid \forall x \in \mathcal{V} : v_h(x) \geq g(x)\} \quad (\text{III.42})$$

und im Fall $p \equiv 2$ und

$$K_h := \{v_h \in S^p(\mathcal{T}_h) \mid \forall x \in \tilde{\mathcal{V}} : v_h(x) \geq g(x)\} \quad (\text{III.43})$$

sind die Voraussetzungen von Satz III.1 erfüllt. Eine qualitative Konvergenzaussage besteht in der folgenden Form:

Satz III.11 *Es sei $u \in H^1(\Omega, \Gamma_0) \cap H^2(\Omega)$ Lösung von (II.91) und u_h Lösung von (III.41) mit K_h wie in (III.42) sowie $g \in H^2(\Omega)$. Zudem sei \mathcal{T}_h eine reguläre und quasi-uniforme Dreiecksgitterfolge und $p \equiv 1$. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1 \leq Ch$$

Beweis: Th.4.1 in [28]. □

Für $p \equiv 2$ und K_h wie in (III.43) erhält man unter geeigneten Voraussetzungen an f und g , dass $\|u - u_h\|_1 = \mathcal{O}(h^{3/2-\epsilon})$ für ein beliebiges $\epsilon > 0$ ist. Hierzu sei auf Th.4.4 in [28] verwiesen.

Die zu (III.41) gehörende diskrete variationelle Sattelpunktformulierung lautet: Gesucht ist ein $(u_h, \lambda_H) \in S^p(\mathcal{T}_h) \times \Lambda_0 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)$, so dass gilt:

$$\begin{aligned} \forall v_h \in S^p(\mathcal{T}_h) : \\ (\nabla u_h, \nabla v_h)_0 = (f, v_h)_0 + (q, \gamma(v_h))_{0, \Gamma_1} - (\lambda_{0,H}, v_h)_0 \end{aligned} \quad (\text{III.44})$$

$$\forall \mu_{0,H} \in \Lambda_0 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H) : (\mu_{0,H} - \lambda_{0,H}, u_h - g)_0 \leq 0, \quad (\text{III.45})$$

³¹Für Dreiecksgitter ist in der Konstruktion von $S^p(\mathcal{T}_h)$ in (III.12) der Polynomraum $S_2^p := \{\hat{v} : \hat{T} \rightarrow \mathbb{R} \mid \hat{v}(\xi_0, \xi_1) = \sum_{0 \leq i, j \leq p, i+j \leq p} c_{ij} \xi_0^i \xi_1^j, c_{ij} \in \mathbb{R}\}$ für ein Referenzdreieck \hat{T} zu verwenden.

wobei $\tilde{\mathcal{T}}_H$ eine weitere Gitterfolge von Ω und $\tilde{p} : \tilde{\mathcal{T}}_H \rightarrow \mathbb{N}$ eine Polynomgradverteilung ist. Der Raum $\tilde{M}^{\tilde{p}}(\mathcal{T}_H)$ ist definiert als

$$\tilde{M}^{\tilde{p}}(\mathcal{T}_H) := \{v \in L^2(\Omega) \mid \forall T \in \mathcal{T} : v|_T \circ \Psi_T \in S_k^{\tilde{p}_T}\}.$$

Zu klären bleibt, ob ähnlich wie in Satz III.7 mit Hilfe eines geeigneten Dualitätsarguments und unter der Voraussetzung einer inversen Abschätzung die Gültigkeit der diskreten Babuška-Brezzi-Bedingung für die hier vorgestellte Diskretisierung gezeigt werden kann.

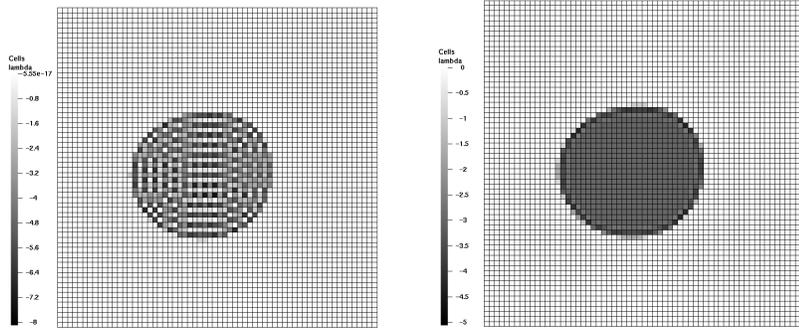


Abb. III.14: $\lambda_{0,H}$ mit $p \equiv 1$, $\tilde{p} \equiv 0$ und $h/H = 1$ bzw. $h/H = 0.5$.

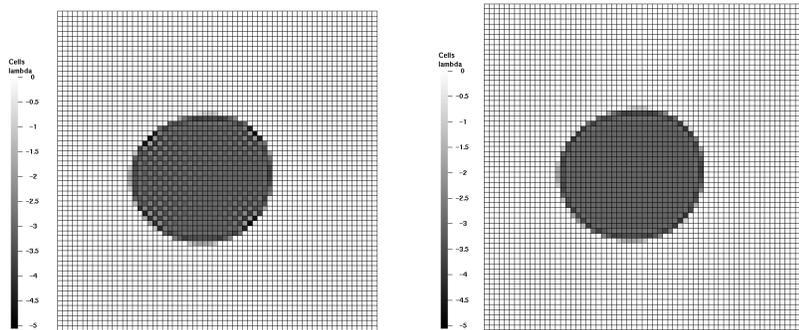


Abb. III.15: $\lambda_{0,H}$ mit $p \equiv 2$, $\tilde{p} \equiv 1$ und $h/H = 1$ bzw. $h/H = 0.5$.

Abbildung III.14 (links) zeigt $\lambda_{0,H}$ bei $h/H = 1$, $p \equiv 1$ und $\tilde{p} \equiv 0$ mit typischem Instabilitätsmuster. Bei Verwendung von $h/H = 0.5$ sind diese

Muster nicht mehr zu erkennen (Abbildung III.14 (rechts)). Die gleichen Beobachtungen ergeben sich ebenfalls für $p \equiv 2$ und $\tilde{p} \equiv 1$ sowie $h/H = 1$ und $h/H = 0.5$ (Abbildung III.15).

Abschließend wird eine Diskretisierung von vereinfachten Bingham-Fluid-Problemen aus Abschnitt II.5.3 betrachtet. Gesucht ist ein $u_h \in S^p(\mathcal{T}_h)$, so dass

$$\forall v_h \in S^p(\mathcal{T}_h) : (\nabla u_h, \nabla(v_h - u_h))_0 + (s, |\nabla v_h|)_0 - (s, |\nabla u_h|)_0 \geq (f, v_h - u_h)_0$$

erfüllt ist. Dass $\lim_{h \rightarrow 0} u_h = u$ gemäß Satz III.1 gilt, ist etwa aus [55] (Th.5.1) zu entnehmen. Eine Diskretisierung der zugehörigen variationellen Sattelpunktformulierung kann in der folgenden Weise angegeben werden: Gesucht ist ein $(u_h, \lambda_{1,H}) \in S^p(\mathcal{T}_h) \times \Lambda_1 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)^k$ als Lösung von

$$\forall v \in S^p(\mathcal{T}_h) : (\nabla u_h, \nabla v_h)_0 = (f, v_h)_0 - (\lambda_{1,H}, \nabla v_h)_0 \quad (\text{III.46})$$

$$\forall \mu_1 \in \Lambda_1 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)^k : (\mu_{1,H} - \lambda_{1,H}, \nabla u_h)_0 \leq 0. \quad (\text{III.47})$$

III.2 h - und hp -Finite-Elemente-Methoden

Zentral bei der Realisierung von Diskretisierungsverfahren ist die Verwendung von geeigneten Basen bezüglich der gewählten Diskretisierungsräume. Ist $\{\varphi_i\} \subset V_h$ eine Basis und $m := \dim V_h$, dann heißt die positiv definite und symmetrische Matrix $A \in \mathbb{R}^{m \times m}$ mit $A_{ij} := a(\varphi_j, \varphi_i)$ Steifigkeitsmatrix, und $L \in \mathbb{R}^m$ mit $L_i := \langle \ell, \varphi_i \rangle$ Lastvektor. Darüber hinaus sei im Folgenden $\{\psi_{r,i}\}$ für $r = 0, 1$ eine Basis von $U'_{r,H}$ mit $m_r := \dim U'_{r,H}$ sowie $B_r \in \mathbb{R}^{m \times m_r}$ mit $B_{r,ij} := \langle \psi_{r,j}, \beta_r(\varphi_i) \rangle$ und $\bar{\Lambda}_r := \{\bar{\mu}_r \in \mathbb{R}^{m_r} \mid \bar{\mu}_{r,i} \psi_{r,i} \in \Lambda_{r,H}\}$, wobei $\Lambda_{r,H}$ die entsprechende Diskretisierung von Λ_0 oder Λ_1 (bzw. Λ_n oder Λ_t) ist.

Hiermit erhält man: $u_h = \bar{u}_i \varphi_i$ ist genau dann Lösung von (III.3), wenn gilt:

$$A\bar{u} = L.$$

Ferner ist u_h genau dann Lösung von (III.1), wenn \bar{u} den Ausdruck

$$\min_{\bar{v} \in \bar{K}} \frac{1}{2} \bar{v}^\top A \bar{v} - \bar{v}^\top L \quad (\text{III.48})$$

minimiert, oder wenn die Variationsungleichung

$$\forall \bar{v} \in \bar{K} : (\bar{u} - \bar{v})^\top A \bar{u} \geq (\bar{u} - \bar{v})^\top L$$

erfüllt ist, wobei $\bar{K} := \{\bar{v} \in \mathbb{R}^m \mid \bar{v}_i \varphi_i \in K_h\}$ ist.

Es sei $\bar{g} \in \mathbb{R}^{m_0}$ mit $\bar{g}_i := \langle \psi_{0,i}, g \rangle$, dann ist $(u_h, \lambda_{0,H})$ mit $\lambda_{0,H} = \bar{\lambda}_{0,i} \psi_{0,i}$ genau dann Lösung von (III.4)-(III.5), wenn gilt:

$$A\bar{u} = L - B_0\bar{\lambda}_0 \quad (\text{III.49})$$

$$\forall \bar{\mu}_0 \in \bar{\Lambda}_0 : (\bar{\mu}_0 - \bar{\lambda}_0)^\top (B_0^\top \bar{u} - \bar{g}) \leq 0. \quad (\text{III.50})$$

Für elliptische Minimierungsprobleme zweiter Art werden an dieser Stelle nur die entsprechenden Aussagen für Sattelpunktformulierungen angegeben:³² $(u_h, \lambda_{1,H})$ mit $\lambda_{1,H} = \bar{\lambda}_{1,i} \psi_{1,i}$ ist genau dann Lösung von (III.6)-(III.7), wenn

$$A\bar{u} = L - B_1\bar{\lambda}_1 \quad (\text{III.51})$$

$$\forall \bar{\mu}_1 \in \bar{\Lambda}_1 : (\bar{\mu}_1 - \bar{\lambda}_1)^\top B_1 \bar{u} \leq 0 \quad (\text{III.52})$$

erfüllt ist. Zudem ist $(u_h, \lambda_{0,H}, \lambda_{1,H})$ genau dann Lösung von (III.8)-(III.9), wenn gilt:

$$A\bar{u} = L - B_0\bar{\lambda}_0 - B_1\bar{\lambda}_1 \quad (\text{III.53})$$

$$\forall (\bar{\mu}_0, \bar{\mu}_1) \in \bar{\Lambda}_0 \times \bar{\Lambda}_1 : (\bar{\mu}_0 - \bar{\lambda}_0)^\top (B_0^\top \bar{u} - \bar{g}) + (\bar{\mu}_1 - \bar{\lambda}_1)^\top B_1 \bar{u} \leq 0. \quad (\text{III.54})$$

Ein wesentlicher Vorteil der Finite-Elemente-Methode ist, dass die Matrizen A und B_r mit Hilfe von Elementmatrizen assembliert werden können, da für $v_h, w_h \in S^p(\mathcal{T}_h)^l$ und $\mu_{r,H} \in M^{p_1}(\mathcal{T}_{1,H})^{l_r}$ typischerweise $a(v_h, w_h) = \sum_T a_T(v_h, w_h)$ und $\langle \mu_{r,H}, \beta_r(v_h) \rangle = \langle \mu_{r,H}, \beta_r(v_h) \rangle = \sum_T \langle \mu_{r,H}, \beta_r(v_h) \rangle_T$ ist. Hierzu sei $\{\xi_{k,i}^{q,\kappa}\}$ eine Basis von $(S_k^q)^\kappa$, dann existieren Matrizen $\pi_T \in \mathbb{R}^{m \times l m_k^{pT}}$ und $\pi_{r,T} \in \mathbb{R}^{m \times l_r m_{k-1}^{p_1,T}}$ mit $m_k^q := \dim S_k^q$, so dass gilt:

$$\forall T \in \mathcal{T}_h : \varphi_{i|T} \circ \Psi_T = \pi_{T,ij} \xi_{k,j}^{pT,l} \quad (\text{III.55})$$

$$\forall T \in \mathcal{T}_{1,H} : \psi_{r,i|T} \circ \Psi_{1,T} = \pi_{r,T,ij} \xi_{k-1,j}^{p_1,T,l_r}. \quad (\text{III.56})$$

Die Matrix $A_T \in \mathbb{R}^{l m_k^{pT} \times l m_k^{pT}}$ mit $A_{T,ij} := a_T(\xi_{k,j}^{pT,l} \circ \Psi_T^{-1}, \xi_{k,i}^{pT,l} \circ \Psi_T^{-1})$ für $T \in \mathcal{T}_h$ heißt Elementmatrix oder lokale Steifigkeitsmatrix. Offenbar gilt:

$$A = \sum_{T \in \mathcal{T}_h} \pi_T A_T \pi_T^\top. \quad (\text{III.57})$$

³²Für weitere Details insbesondere zu Penalisierungsmöglichkeiten von elliptischen Minimierungsproblemen zweiter Art sei auf [55] und [72] verwiesen.

Entsprechende Elementmatrizen erhält man ebenfalls für die Matrizen B_r . Hierzu sei $\tilde{\mathcal{T}}_{1,h} := \{\bar{T} \cap \Gamma_1 \mid T \in \mathcal{T}_h\}$, und für $T \in \tilde{\mathcal{T}}_{1,h}$ sei $\omega(T) \in \mathcal{T}_h$, so dass $\overline{\omega(T)} \cap \Gamma_1 = T$ ist. Ferner gelte

$$\forall T \in \tilde{\mathcal{T}}_{1,h} : \exists c(T) \in \mathcal{T}_{1,H} : T \subset c(T).^{33}$$

Dann definiert man die Elementmatrix $B_{r,T} \in \mathbb{R}^{lm_k^{p_\omega(T)} \times l_r m_{k-1}^{p_{1,c(T)}}}$ mit $B_{r,T,ij} := ((\xi_{k-1,j}^{p_{1,c(T)},l_r} \circ \Psi_{1,c(T)}^{-1})|_T, \beta_r(\xi_{k,i}^{p_\omega(T),l} \circ \Psi_{\omega(T)}^{-1})|_T)_T$ für $T \in \tilde{\mathcal{T}}_{1,h}$. Auch hier ist

$$B_r = \sum_{T \in \tilde{\mathcal{T}}_{1,h}} \pi_{\omega(T)} B_{r,T} \pi_{r,c(T)}^\top. \quad (\text{III.58})$$

In analoger Weise können auch Matrizen \tilde{B}_r sowie $\tilde{B}_{r,T}$ und $\tilde{\pi}_{r,T}$ bezüglich der Räume $\tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)$ bzw. $\tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)^k$ konstruiert werden. Zudem können der Lastvektor L und der Vektor \tilde{g} ebenfalls über Elementvektoren assembliert werden.

Mit Blick auf (III.57) und (III.58) ist die Konstruktion einer geeigneten Basis von $(S_k^q)^k$ und die Angabe der zugehörigen Matrizen π_T bzw. $\pi_{r,T}$ entscheidend für die Assemblierung von A und B_r .³⁴

Unproblematisch ist die Situation hinsichtlich $M^{p_1}(\mathcal{T}_{1,H})^{l_r}$, da in diesem Fall keinerlei Stetigkeitsanforderungen oder Bedingungen an Randwerte gestellt werden. Hier kann eine beliebige Basis von S_k^q gewählt werden, und im Fall $l_r = 1$ setzt man einfach $\pi_{r,T,\zeta_1(T)+i,i} := 1$ für $i = 0, \dots, m_{k-1}^{p_{1,T}} - 1$.³⁵

Die Abbildung $\zeta_1 : \mathcal{T}_{1,H} \rightarrow \mathbb{N}_0$ mit $\zeta_1(T_s) := \sum_{i=0}^{s-1} m_{k-1}^{p_{1,T_i}}$ stellt hier den Zusammenhang von lokaler und globaler Nummerierung her. Für $l_r = 2$ setzt man $\pi_{r,T,\zeta_1(T)+i,i} := \pi_{r,T,m_r/2+\zeta_1(T)+i,i} := 1$ für $i = 0, \dots, m_{k-1}^{p_{1,T}} - 1$. Gemäß (III.56) besteht der Träger der zugehörigen globalen Basis $\{\psi_{r,i}\}$ demnach nur aus jeweils einem Element $T \in \mathcal{T}_{1,H}$.³⁶

Komplizierter ist dagegen die Angabe von π_T bezüglich $S^p(\mathcal{T}_h)^l$ bei Verwendung einer Basis von $(S_k^q)^l$, weil nun Stetigkeitsanforderungen über Kanten bzw. Facetten hinweg sowie Randwerte zu berücksichtigen sind.³⁷ Einen

³³D.h. $\mathcal{T}_{1,H}$ ist ein gröberes Gitter von $\tilde{\mathcal{T}}_{1,h}$.

³⁴Die effiziente Umsetzung von π_T und $\pi_{r,T}$ in Form von sogenannten DoF-Managern (DoF=Degree of Freedom) ist ein wesentlicher Bestandteil moderner Finite-Elemente-Implementierungen. Hierzu sei auf die Finite-Elemente-Software SOFAR [87] verwiesen, mit der auch sämtliche Testrechnungen durchgeführt wurden.

³⁵Hier wie im Folgenden seien die nicht definierten Matrixkomponenten stets 0.

³⁶Die Finite-Elemente-Basisfunktionen werden typischerweise so gewählt, dass ihre Träger möglichst klein sind. Eine wichtige Folge hiervon ist, dass die Matrizen A und B_r dünn besetzt sind, so dass relativ große Systeme behandelt werden können.

³⁷zur Stetigkeit von Funktionen aus $S^p(\mathcal{T}_h)$ vgl. z.B. Satz 5.2 in [25]

wesentlichen Ansatzpunkt liefert der Gebrauch von hierarchischen Basen. Im Folgenden Abschnitt werden hierarchische Basen für die Räume S_2^q und S_3^q mit Hilfe rekursiv definierter Gegenbauerpolynome konstruiert und die zugehörigen Matrizen π_T bzgl. $S^p(\mathcal{T}_h)$ gewonnen. Entsprechende Erweiterungen auf den Raum $S^p(\mathcal{T}_h)^l$ sind dann offensichtlich.

III.2.1 Hierarchische Basen in 2D und 3D

Es sei $\varrho \in \mathbb{R}$, sowie $G_0^\varrho(x) := 1$ und $G_1^\varrho(x) := 2\varrho x$. Dann heißt die Folge der Polynome $\{G_n^\varrho\}_{n \in \mathbb{N}_0}$, die der Rekursionsgleichung

$$(n+1)G_{n+1}^\varrho(x) = 2(n+\varrho)xG_n^\varrho(x) - (n+2\varrho-1)G_{n-1}^\varrho(x) \quad (\text{III.59})$$

genügt, Folge von Gegenbauerpolynomen bezüglich ϱ . Aussagen über Eigenschaften von Gegenbauerpolynomen sind z.B. in [102] enthalten. In dem hier betrachteten Zusammenhang sind vor allem Symmetrie-Eigenschaften und Möglichkeiten zur Berechnung von Ableitungswerten interessant. Für $n, \nu \in \mathbb{N}_0$ und $(\varrho)_\nu := \prod_{i=0}^{\nu-1} (\varrho + i)$ gilt:

$$G_n^\varrho(-x) = (-1)^n G_n^\varrho(x), \quad \partial^\nu G_n^\varrho = 2^\nu (\varrho)_\nu G_{n-\nu}^{\varrho+\nu}. \quad (\text{III.60})$$

Eine Basis von S_2^q erhält man folgendermaßen:

$$\begin{aligned} \xi_{2,0}^q(x_0, x_1) &:= 0.25(1-x_0)(1-x_1), \\ \xi_{2,1}^q(x_0, x_1) &:= 0.25(1+x_0)(1-x_1), \\ \xi_{2,2}^q(x_0, x_1) &:= 0.25(1+x_0)(1+x_1), \\ \xi_{2,3}^q(x_0, x_1) &:= 0.25(1-x_0)(1+x_1). \end{aligned}$$

Diese Basisfunktionen werden auch Knotenmoden genannt. Ist $q > 1$, definiert man die sogenannten Kantenmoden für $i = 0, \dots, q-2$:

$$\begin{aligned} \xi_{2,4+i}^q(x_0, x_1) &:= 0.5(1-x_1)G_{i+2}^{-1/2}(x_0), \\ \xi_{2,4+q-1+i}^q(x_0, x_1) &:= 0.5(1+x_0)G_{i+2}^{-1/2}(x_1), \\ \xi_{2,4+2(q-1)+i}^q(x_0, x_1) &:= 0.5(1+x_1)G_{i+2}^{-1/2}(x_0), \\ \xi_{2,4+3(q-1)+i}^q(x_0, x_1) &:= 0.5(1-x_0)G_{i+2}^{-1/2}(x_1). \end{aligned}$$

Schließlich definiert man für $q > 3 - \tau$ mit $\tau \in \{0, 2\}$, $i = 0, \dots, q-4+\tau$ und $j = 0, \dots, i$ sowie $\alpha(i, j) := i(i+1)/2 + j$ die sogenannten inneren Moden:

$$\xi_{2,4q+\alpha(i,j)}^q(x_0, x_1) := G_{j+2}^{-1/2}(x_0)G_{i-j+2}^{-1/2}(x_1).$$

Knotenmoden haben die Eigenschaft, in genau einem Eckpunkt von $[-1, 1]^2$ den Wert 1 und in den übrigen den Wert 0 zu haben. Kantenmoden sind auf genau einer Kante von $[-1, 1]^2$ von Null verschieden und verschwinden auf den übrigen Kanten. Innere Moden verschwinden auf allen Kanten von $[-1, 1]^2$ und sind nur im Innern von null verschieden.

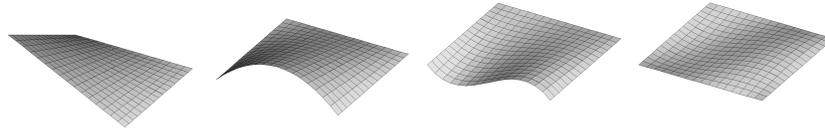


Abb. III.16: $\xi_{2,0}^5$, $\xi_{2,4}^5$, $\xi_{2,5}^5$ und $\xi_{2,22}^5$ bei $\tau = 0$.

Für $\tau = 2$ ist $\{\xi_{2,i}^q\}$ eine Basis von S_2^q . Jedoch können dieser Basis innere Moden entnommen werden, ohne dass die Approximationseigenschaften verloren gehen.³⁸ Die für $\tau = 0$ gewonnene, ausgedünnte Basis heißt auch Basis der Serendipity-Klasse.

Die Basis $\{\xi_{2,i}^q\}$ entspricht (bis auf Normierungsfaktoren) den in [89] und [99] angegebenen Basisfunktionen vom (aufintegrierten) Legendre-Typ. Jedoch ist die obige Darstellung auf der Grundlage von Gegenbauerpolynomen günstiger für Implementierungen und für die Einbeziehung von irregulären Knoten bzw. Kanten.³⁹ Gegenüber Standardbasen vom Lagrange-Typ hat diese Basis mehrere Vorteile: Zum einen können Ableitungen beliebiger Ordnung leicht mit Hilfe der in (III.60) dargestellten Formel ausgedrückt werden. Als Ergebnis erhält man wiederum Gegenbauerpolynome. Zum anderen können Funktions- und Ableitungswerte über die Rekursionsgleichung (III.59) numerisch stabil und mit geringem Aufwand berechnet werden.⁴⁰ Ein weiterer Vorteil ist, dass diese Basis hierarchisch ist, d.h., es gilt $\{\xi_{2,i}^q\} \subset \{\xi_{2,i}^{q+1}\}$. Diese Eigenschaft führt dazu, dass Einträge in π_T , die Kantenmoden mit einem höheren Polynomgrad als bei dem jeweiligen Nachbarlement zugeordnet sind, lediglich auf 0 zu setzen sind, damit die Stetigkeitsanforderung erfüllt ist.

Analysen bzgl. der Performance, insbesondere bzgl. der Konditionszahl der

³⁸vgl. S. 175 in [89]

³⁹vgl. Abschnitt III.2.2

⁴⁰Tatsächlich wird in der Literatur häufig die Implementierung von Basisfunktionen dieses Typs über Monomdarstellungen empfohlen, womit die oben genannten Vorteile verloren gehen. Hierzu sei z.B. auf S.23 in [92] verwiesen.

aufgrund dieser Basis assemblierten Matrizen und bzgl. des Einsatzes von Präkonditionierern sind in [6], [8] und [77] enthalten.

Ein Nachteil der hier angegebenen Basis gegenüber Lagrange-artigen Basen ist jedoch, dass die Verschiebungswerte nur in den zu den Knotenmoden gehörenden Komponenten von \bar{u} abgelesen werden können.

Im zweidimensionalen Fall wird die Verbindung zu der zugehörigen globalen Basis über die Matrix $\pi_T := \pi_T^2$ erzeugt. In diesem Fall stellt die Abbildung $\zeta : \mathcal{V} \cup \mathcal{E} \cup \mathcal{T}_h \rightarrow \mathbb{N}_0$ den Zusammenhang von lokaler und globaler Nummerierung her. Dabei ist $\mathcal{V} = \{V_0, V_1, \dots\}$ die Menge aller Ecken und $\mathcal{E} = \{E_0, E_1, \dots\}$ die Menge aller Kanten von $\mathcal{T}_h = \{T_0, T_1, \dots\}$, die nicht auf Γ_0 liegen. Es sei $p_E := \min\{p_T \mid \bar{T} \cap E \neq \emptyset, T \in \mathcal{T}_h\}$ für eine Kante E , $m_{\mathcal{V}} := |\mathcal{V}|$ und $m_{\mathcal{E}} := \sum_{i=0}^{|\mathcal{E}|-1} (p_{E_i} - 1)$ sowie

$$\begin{aligned} \zeta(V_s) &:= s, & \zeta(E_s) &:= m_{\mathcal{V}} + \sum_{i=0}^{s-1} (p_{E_i} - 1), \\ \zeta(T_s) &:= m_{\mathcal{V}} + m_{\mathcal{E}} + \sum_{i=0}^{s-1} (p_{T_i} - 3 + \tau)(p_{T_i} - 2 + \tau)/2. \end{aligned}$$

Im Folgenden sei \mathcal{T}_h ein reguläres Gitter und $T \in \mathcal{T}_h$, zudem seien V_0, \dots, V_3 die Ecken von T und E_0, \dots, E_3 die Kanten. Wenn $V_\nu \in \mathcal{V}$ bzw. $E_\nu \in \mathcal{E}$ ist, setzt man unter Ausnutzung, dass $\{\xi_{2,i}^q\}$ hierarchisch ist und unter Einbeziehung der Symmetrie-Eigenschaft aus (III.60) mit $\theta_{T,i} \in \{-1, 1\}$:

$$\begin{aligned} \pi_{T, \zeta(V_\nu), \nu}^2 &:= 1, \\ \pi_{T, \zeta(E_\nu) + i, 4 + \nu(p_{E_\nu} - 1) + i}^2 &:= (\theta_{T, \nu})^i, \quad i = 0, \dots, p_{E_i} - 2, \\ \pi_{T, \zeta(T) + \alpha(i, j), 4p_T + \alpha(i, j)}^2 &:= 1, \quad i = 0, \dots, p_T - 4 + \tau, \quad j = 0, \dots, i. \end{aligned}$$

Für $\theta_{T, \nu} \equiv 1$ ist bei ungeradem i nicht notwendigerweise die Stetigkeitsforderung über Kanten hinweg erfüllt. Ein Beispiel: Die Kante $E = \bar{T}_0 \cap \bar{T}_1$ mit $T_0, T_1 \in \mathcal{T}_h$ habe die Eckpunkte P und Q , und es gelte $\Psi_{T_0}(1, -1) = P = \Psi_{T_1}(-1, -1)$ und $\Psi_{T_0}(1, 1) = Q = \Psi_{T_1}(-1, 1)$, dann ist die zugehörige globale Basisfunktion $\varphi_{\zeta_{\mathcal{V}}(E)+1}$ nicht stetig.

Abbildung III.17 zeigt die im obigen Beispiel beschriebene Unstetigkeit. Durch Setzen von $\theta_{T,1} := -1$ wird diese Unstetigkeit behoben. Besser ist jedoch, dieser Problematik mit Hilfe von angepassten Transformationen $\{\Psi_T\}$ und einer angemessenen Wahl von Kantenorientierungen zu begegnen.

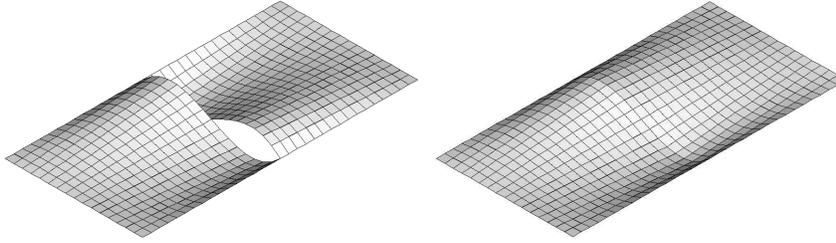


Abb. III.17: $\xi_{2,9}^5 \circ \Psi_{T_0}^{-1}$ und $\xi_{2,17}^5 \circ \Psi_{T_1}^{-1}$ mit $\theta_{T,1} = 1$ und $\theta_{T,1} = -1$.

Definition III.1 Es sei $\mathcal{G}(\mathcal{T}_h)$ die Kantenmenge eines gerichteten Graphen bestehend aus den Kanten von \mathcal{T}_h . Ferner sei \mathcal{Q}^k die Kantenmenge eines gerichteten Graphen bestehend aus den Kanten von $[-1, 1]^k$. Dann heißen die Transformationen $\{\Psi_T\}$ gerichtet, wenn gilt:

$$\forall T \in \mathcal{T}_h : \forall (Q_0, Q_1) \in \mathcal{Q}^k : (\Psi_T(Q_0), \Psi_T(Q_1)) \in \mathcal{G}(\mathcal{T}_h).$$

Für gerichtete Transformationen stimmt demnach die durch $\mathcal{G}(\mathcal{T}_h)$ festgelegte Orientierung der Kanten der Gitterelemente mit der vorgegebenen Orientierung \mathcal{Q}^k des Referenzelements überein.

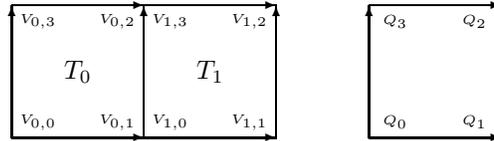


Abb. III.18: Beispiele für $\mathcal{G}(\{T_0, T_1\})$ und \mathcal{Q}^2 .

In Abbildung III.18 ist ein Beispiel für die gerichtete Graphen $\mathcal{G}(\{T_0, T_1\})$ und \mathcal{Q}^2 mit $Q_0 := (-1, -1)$, $Q_1 := (1, -1)$, $Q_2 := (1, 1)$ und $Q_3 := (1, -1)$ dargestellt. Gilt z.B. $\Psi_{T_i}(Q_j) = V_{i,j}$ für $i = 0, 1$ und $j = 0, \dots, 3$, dann sind offenbar die Transformationen $\{\Psi_{T_0}, \Psi_{T_1}\}$ gerichtet. Insgesamt erhält man:

Satz III.12 Es sei \mathcal{T}_h ein reguläres Gitter von $\Omega \subset \mathbb{R}^2$, und die Transformationen $\{\Psi_T\}$ seien bzgl. \mathcal{Q}^2 aus Abb. III.18 gerichtet. Dann existieren Funktionen $\{\varphi_i\} \subset S^p(\mathcal{T}_h)$, so dass $\varphi_i \circ \Psi_T = \pi_{T,ij}^2 \xi_{2,j}^{pT}$ mit $\theta_{T,\nu} \equiv 1$ ist. Im Fall $\tau = 2$ ist $\{\varphi_i\}$ eine Basis von $S^p(\mathcal{T}_h)$.

Im zweidimensionalen Fall ist die Angabe von θ_T auch für ungerichtete Transformationen leicht durchführbar. Hier setzt man $\theta_{T,i} = -1$, wenn für die Kante von T mit dem Index i und den Endpunkten $\Psi_T(Q_{i,0})$ und $\Psi_T(Q_{i,1})$ mit $(Q_{i,0}, Q_{i,1}) \in \mathcal{Q}^2$ gilt, dass $(\Psi_T(Q_{i,0}), \Psi_T(Q_{i,1})) \notin \mathcal{G}(\mathcal{T}_h)$ ist. Wesentlich komplizierter ist dagegen die Situation für drei Dimensionen, da hier nicht nur die Orientierung der Kanten, sondern auch die Orientierung der Facetten zu beachten ist. Im Finite-Element Paket SOFAR [87] sind hierzu spezielle Datenstrukturen verfügbar, die die Verwendung von hierarchischen Basen von $S^p(\mathcal{T}_h)$ auch für ungerichtete Transformationen ermöglichen.⁴¹

Im Folgenden wird eine Basis für S_3^q angegeben, der im Fall von gerichteten Transformationen eine Basis von $S^p(\mathcal{T}_h)$ zugeordnet werden kann.

Die Knotenmoden sind hier:

$$\begin{aligned}\xi_{3,0}^q(x_0, x_1, x_2) &:= 0.125(1-x_0)(1-x_1)(1-x_2), \\ \xi_{3,1}^q(x_0, x_1, x_2) &:= 0.125(1+x_0)(1-x_1)(1-x_2), \\ \xi_{3,2}^q(x_0, x_1, x_2) &:= 0.125(1+x_0)(1+x_1)(1-x_2), \\ \xi_{3,3}^q(x_0, x_1, x_2) &:= 0.125(1-x_0)(1+x_1)(1-x_2), \\ \xi_{3,4}^q(x_0, x_1, x_2) &:= 0.125(1-x_0)(1-x_1)(1+x_2), \\ \xi_{3,5}^q(x_0, x_1, x_2) &:= 0.125(1+x_0)(1-x_1)(1+x_2), \\ \xi_{3,6}^q(x_0, x_1, x_2) &:= 0.125(1+x_0)(1+x_1)(1+x_2), \\ \xi_{3,7}^q(x_0, x_1, x_2) &:= 0.125(1-x_0)(1+x_1)(1+x_2).\end{aligned}$$

Für $q > 1$ und $i = 0, \dots, q-2$ erhält man die folgenden Kantenmoden:

$$\begin{aligned}\xi_{3,8+i}^q(x_0, x_1, x_2) &:= 0.25(1-x_1)(1-x_2)G_{i+2}^{-1/2}(x_0), \\ \xi_{3,8+q-1+i}^q(x_0, x_1, x_2) &:= 0.25(1+x_0)(1-x_2)G_{i+2}^{-1/2}(x_1), \\ \xi_{3,8+2(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1+x_1)(1-x_2)G_{i+2}^{-1/2}(x_0), \\ \xi_{3,8+3(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1-x_0)(1-x_2)G_{i+2}^{-1/2}(x_1), \\ \xi_{3,8+4(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1-x_0)(1-x_1)G_{i+2}^{-1/2}(x_2),\end{aligned}$$

⁴¹Zu bedenken ist, dass sich die Transformationen $\{\Psi_T\}$ in der Regel aus einer zufälligen Reihenfolge der Eckpunkte der Gitterelemente T ergeben. Das ist z.B. der Fall, wenn Gitter aus CAD-Daten konstruiert werden. Die resultierenden Transformationen sind im Allgemeinen ungerichtet. Transformationen können z.B. folgendermaßen angegeben werden: Sind etwa $V_{T,0}, \dots, V_{T,3}$ die Eckpunkte des Gitterelements T im Uhrzeigersinn oder entgegengesetzt, erhält man mit $\Psi_T(x_0, x_1) := \xi_{2,i}^1(x_0, x_1)V_{T,i}$ eine Transformation von $[-1, 1]^2$ auf T .

$$\begin{aligned}
\xi_{3,8+5(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1+x_0)(1-x_1)G_{i+2}^{-1/2}(x_2), \\
\xi_{3,8+6(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1+x_0)(1+x_1)G_{i+2}^{-1/2}(x_2), \\
\xi_{3,8+7(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1-x_0)(1+x_1)G_{i+2}^{-1/2}(x_2), \\
\xi_{3,8+8(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1-x_1)(1+x_2)G_{i+2}^{-1/2}(x_0), \\
\xi_{3,8+9(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1+x_0)(1+x_2)G_{i+2}^{-1/2}(x_1), \\
\xi_{3,8+10(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1+x_1)(1+x_2)G_{i+2}^{-1/2}(x_0), \\
\xi_{3,8+11(q-1)+i}^q(x_0, x_1, x_2) &:= 0.25(1-x_0)(1+x_2)G_{i+2}^{-1/2}(x_1).
\end{aligned}$$

Für $q > 3 - \tau_0$ sowie $i = 0, \dots, q - 4 + \tau_0$ und $j = 0, \dots, i$ erhält man mit $r_0 := 8 + 12(q - 1)$ und $r_1 := ((p - 3 + \tau_0)(p - 2 + \tau_0))/2$ die Facettenmoden

$$\begin{aligned}
\xi_{3,r_0+\alpha(i,j)}^q(x_0, x_1, x_2) &:= 0.5(1-x_2)G_{j+2}^{-1/2}(x_0)G_{i-j+2}^{-1/2}(x_1), \\
\xi_{3,r_0+r_1+\alpha(i,j)}^q(x_0, x_1, x_2) &:= 0.5(1+x_0)G_{j+2}^{-1/2}(x_1)G_{i-j+2}^{-1/2}(x_2), \\
\xi_{3,r_0+2r_1+\alpha(i,j)}^q(x_0, x_1, x_2) &:= 0.5(1+x_1)G_{j+2}^{-1/2}(x_2)G_{i-j+2}^{-1/2}(x_0), \\
\xi_{3,r_0+3r_1+\alpha(i,j)}^q(x_0, x_1, x_2) &:= 0.5(1+x_2)G_{j+2}^{-1/2}(x_0)G_{i-j+2}^{-1/2}(x_1), \\
\xi_{3,r_0+4r_1+\alpha(i,j)}^q(x_0, x_1, x_2) &:= 0.5(1-x_0)G_{j+2}^{-1/2}(x_1)G_{i-j+2}^{-1/2}(x_2), \\
\xi_{3,r_0+5r_1+\alpha(i,j)}^q(x_0, x_1, x_2) &:= 0.5(1-x_1)G_{j+2}^{-1/2}(x_2)G_{i-j+2}^{-1/2}(x_0).
\end{aligned}$$

Die inneren Moden sind für $q > 5 - \tau_1$ und $i = 0, \dots, q - 6 + \tau_1$ und $j = 0, \dots, i$ und $l = 0, \dots, j$ sowie $r_2 := 8 + 12(p - 1) + 3(p - 3 + \tau_0)(p - 2 + \tau_0)$ und $\beta(i, j, l) := i(i + 1)(2i + 4)/12 + \alpha(j, l)$ definiert als

$$\xi_{3,r_2+\beta(i,j,l)}^q(x_0, x_1, x_2) := G_{l+2}^{-1/2}(x_0)G_{j-l+2}^{-1/2}(x_1)G_{i-j+2}^{-1/2}(x_2).$$

Für $\tau_0 = \tau_1 = 0$ erhält man eine Basis der Serendipity-Klasse. Für $\tau_0 = 2$ und $\tau_1 = 4$ spannt $\{\xi_{3,i}^q\}$ den Raum S_3^q auf.

Zur Definition der Matrix $\pi_T := \pi_T^3$ ist die Index-Abbildung $\zeta : \mathcal{V} \cup \mathcal{E} \cup \mathcal{F} \cup \mathcal{T}_h \rightarrow \mathbb{N}_0$ erforderlich, wobei $\mathcal{F} := \{F_0, F_1, \dots\}$ die Menge aller Facetten von \mathcal{T}_h ist, die nicht auf Γ_0 liegen. Bezüglich Elementen aus \mathcal{V} und \mathcal{E} ist ζ wie

im zweidimensionalen Fall definiert. Darüber hinaus setzt man:

$$\begin{aligned}\zeta(F_s) &:= m_{\mathcal{V}} + m_{\mathcal{E}} + \sum_{i=0}^{s-1} (p_{F_i} - 3 + \tau_0)(p_{F_i} - 2 + \tau_0)/2, \\ \zeta(T_s) &:= m_{\mathcal{V}} + m_{\mathcal{E}} + m_{\mathcal{F}} \\ &\quad + \sum_{i=0}^{s-1} ((p_{T_i} - 3 + \tau_0)(p_{T_i} - 4 + \tau_1)(p_{T_i} - 5 + \tau_1))/6,\end{aligned}$$

wobei $m_{\mathcal{F}} := \sum_{i=0}^{|\mathcal{F}|-1} (p_{F_i} - 3 + \tau_0)(p_{F_i} - 2 + \tau_0)/2$ und $p_F := \min\{p_T \mid \bar{T} \cap F \neq \emptyset, T \in \mathcal{T}_h\}$ für eine Facette F definiert sind.

Für $T \in \mathcal{T}_h$ seien V_0, \dots, V_7 die Ecken, E_0, \dots, E_{11} die Kanten und F_0, \dots, F_5 die Facetten von T . Sofern $V_\nu \in \mathcal{V}$, $E_\nu \in \mathcal{E}$ bzw. $F_\nu \in \mathcal{F}$ ist, setzt man mit $\theta_{T,\nu}^1, \theta_{T,\nu,ij}^2 \in \{-1, 1\}$:

$$\begin{aligned}\pi_{T,\zeta(V_\nu),\nu}^3 &:= 1 \\ \pi_{T,\zeta(E_\nu)+j,8+\nu(p_{E_\nu}^\mathcal{E}-1)+i}^3 &:= (\theta_{T,\nu}^1)^i, \quad i = 0, \dots, p_{E_\nu} - 2 \\ \pi_{T,\zeta(F_\nu)+\alpha(i,j),r_0+\nu r_1+\alpha(i,j)}^3 &:= \theta_{T,\nu,ij}^2, \quad i = 0, \dots, p_{F_\nu} - 4 + \tau_0, \quad j = 0, \dots, i \\ \pi_{T,\zeta(T)+\beta(i,j,l),r_2+\beta(i,j,l)}^3 &:= 1, \quad i = 0, \dots, q - 6 + \tau_1, \\ &\quad j = 0, \dots, i, \quad l = 0, \dots, j\end{aligned}$$

Für den dreidimensionalen Fall ist eine geeignete Kantenorientierung des Referenzelements anhand des folgenden Beispiels mit den Kantenmengen $\mathcal{G}(\{T_0, T_1\})$ und \mathcal{Q}^3 gegeben; hier kann $\theta_{T,\nu}^1 = \theta_{T,\nu,ij}^2 = 1$ gesetzt werden:

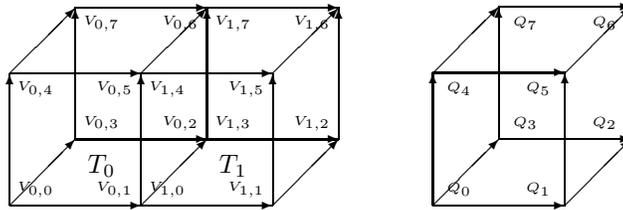


Abb. III.19: Beispiele für $\mathcal{G}(\{T_0, T_1\})$ und \mathcal{Q}^3 .

In Abbildung III.19 ist $Q_0 := (-1, -1, -1), \dots, Q_3 := (-1, 1, -1)$ und $Q_4 := (-1, -1, 1), \dots, Q_7 := (-1, 1, 1)$. Analog zum zweidimensionalen Fall folgt aus $\Psi_{T_i}(Q_j) = V_{i,j}$ für $i = 0, 1$ und $j = 0, \dots, 7$, dass die Transformationen $\{\Psi_{T_0}, \Psi_{T_1}\}$ gerichtet sind.

Satz III.13 *Es sei \mathcal{T}_h ein reguläres Gitter von $\Omega \subset \mathbb{R}^3$, und die Transformationen $\{\Psi_T\}$ seien bzgl. \mathcal{Q}^3 aus Abb. III.19 gerichtet. Dann existieren Funktionen $\{\varphi_i\} \subset S^p(\mathcal{T}_h)$, so dass $\varphi_i \circ \Psi_T = \pi_{T,ij}^3 \mathcal{E}_{3,j}^{pT}$ mit $\theta_T^1 \equiv \theta_{T,ij,\nu}^2 \equiv 1$ ist. Im Fall $\tau_0 = 2$ und $\tau_1 = 4$ ist $\{\varphi_i\}$ eine Basis von $S^p(\mathcal{T}_h)$.*

III.2.2 Irreguläre Knoten, Kanten, Facetten

Lokale Verfeinerungsprozesse, so wie sie z.B. bei Gitteradaptierungsmaßnahmen⁴² zur Reduktion der lokalen Maschenweiten eingesetzt werden, führen in der Regel auf irreguläre Gitter mit sogenannten irregulären Knoten, Kanten bzw. Facetten. So sind zum Beispiel die in Abbildung III.1 und III.2 dargestellten Gitter ebenso wie die Gitter aus dem Abschnitt IV.7.3 und dem Anhang A.1 durch lokale Verfeinerungen entstanden und irregulär. Werden keine weitere Zerlegungsmaßnahmen ergriffen, um die Irregularität des Gitters (wie z.B. in den Abbildungen III.3 und III.4 geschehen) aufzulösen, kann die im vorherigen Abschnitt definierte Matrix π_T nicht verwendet werden. Im Folgenden wird eine Modifikation dieser Matrix vorgestellt, so dass auch irreguläre Gitter genutzt werden können. Zentral hierbei ist die Einteilung der Knoten, Kanten und Facetten des zugrunde liegenden Gitters in reguläre und irreguläre Elemente.

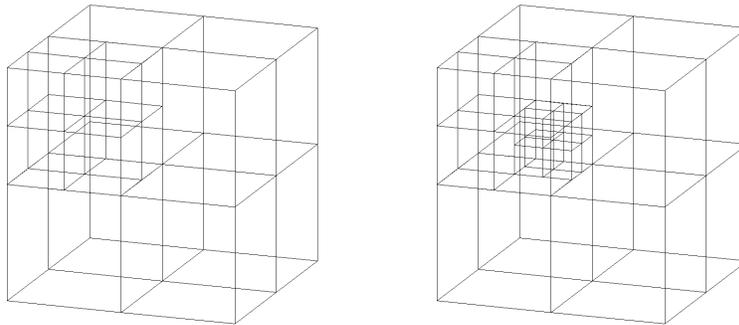


Abb. III.20: 1- und 2-irreguläre Gitter.

Vorzugsweise werden Viereckselemente symmetrisch in 4 Vierecke, sowie Hexaederelemente in 8 Hexaeder unterteilt.⁴³ Dabei werden die Kanten geteilt

⁴²vgl. Abschnitt III.1.1 und IV.7.3

⁴³Denkbar sind auch andere Zerlegungsformen. Diesbezüglich können in dem FEM-

und die Facetten geviertelt. Die Orientierung der neuen Kanten wird dabei von den geteilten Kanten übernommen.

In einem so lokal verfeinerten Gitter \mathcal{T}_h heißt eine Kante E regulär oder 0-irregulär, wenn sie keine echte Teilmenge einer anderen Kante von \mathcal{T}_h ist. Analog heißt eine Facette regulär oder 0-irregulär, wenn sie keine echte Teilmenge einer anderen Facette von \mathcal{T}_h ist. Eine Kante E heißt k -Kanten-irregulär mit dem Index $r \in \{0, 1\}$ bzw. k -Facetten-irregulär mit dem Index $r \in \{0, \dots, 3\}$, wenn sie die r -te Teilungskante einer $(k-1)$ -irregulären Kante $\mathcal{I}^1(E)$ bzw. die r -te Teilungsfacette einer $(k-1)$ -irregulären Facette $\mathcal{I}^2(E)$ ist. Ebenso heißt eine Facette F k -irregulär mit dem Index $r \in \{0, \dots, 3\}$, wenn sie die r -te Teilungsfacette einer $(k-1)$ -irregulären Facette $\mathcal{I}^2(F)$ ist. Schließlich heißt ein Knoten V k -Kanten-irregulär bzw. k -Facetten-irregulär, wenn er eine $(k-1)$ -irreguläre Kante $\mathcal{I}^1(V)$ bzw. Facette $\mathcal{I}^2(V)$ teilt bzw. viertelt. Nicht alle irregulären Kanten oder Facetten müssen dabei Kanten von Gitterelementen von \mathcal{T}_h sein.

Das Problem bei der Verwendung von irregulären Gittern ist die Gewährleistung der Stetigkeitsanforderung über irreguläre Kanten oder Facetten hinweg. Die Lösung besteht darin, die lokalen Basisfunktionen, die auf Gitterelementen mit k -irregulären Kanten oder Facetten definiert sind und auf diesen nicht verschwinden, über Kopplungsgewichte mit lokalen Basisfunktionen zu verknüpfen. Diese sind den entsprechenden $(k-1)$ -irregulären Kanten bzw. Facetten zugeordnet, so dass insgesamt stetige Basisfunktionen erzeugt werden. Dieser Kopplungsprozess wird im Folgenden mit Hilfe rekursiv definierter Mengen beschrieben, die die Kopplungen und die dazugehörigen Kopplungsgewichte enthalten. Entscheidend bei dem hier gewählten Ansatz ist, dass beliebige Konstellationen k -irregulärer Knoten, Kanten oder Facetten dabei berücksichtigt werden.⁴⁴

Für einen regulären Knoten V , eine reguläre Kante E und eine reguläre Facette F definiert man für $i = 0, \dots, p_E - 2$ und $\kappa = 0, \dots, r_{1,F} - 1$ mit $r_{1,F} := (p_F - 3 + \tau_0)(p_F - 2 + \tau_0)/2$:

$$\mathcal{B}(V, 0) := \{(V, 0, 1)\}, \quad \mathcal{B}(E, i) := \{(E, i, 1)\}, \quad \mathcal{B}(F, \kappa) := \{(F, \kappa, 1)\}.$$

Ein Tripel (B, b, β) der Mengen $\mathcal{B}(\cdot, \cdot)$ besteht aus einem Freiheitsgradträger

Paket `sofAR` [87] beliebige Zerlegungsmuster definiert werden. So sind auch Zerlegungen, wie sie in Abbildung III.3 und III.4 dargestellt werden, möglich.

⁴⁴In der Literatur wird in der Regel ein entsprechendes Vorgehen lediglich für 1-irreguläre Gitter beschrieben. Hierzu sei z.B. auf Abschnitt 3.6 in [41] oder [92] verwiesen. In den benannten Arbeiten führt man im Fall k -irregulärer Gitter mit $k > 1$ zusätzliche lokale Verfeinerungen durch, um 1-irreguläre Gitter zu erzeugen.

B (Ecke, Kante oder Facette), einem lokalen Freiheitsgradindex b und einem Kopplungsgewicht β .

Es seien E eine Kante und F eine Facette, dann bezeichnen $V_{E,\nu}$ bzw. $V_{F,\nu}$ die ν -te Ecke von E bzw. von F sowie $E_{F,\nu}$ die ν -te Kante von F . Für einen k -Kanten-irregulären bzw. k -Facetten-irregulären Knoten V sei:

$$\begin{aligned} \mathcal{B}(V, 0) &:= \{(V_{\mathcal{I}^1(V),\nu}, \nu, \gamma_{1,0,\nu,0}^{0,0,0} \mid \nu = 0, 1\} \\ &\quad \cup \{(\mathcal{I}^1(V), i', \gamma_{1,1,0,i'}^{1,0,0} \mid i' = 0, \dots, p_{\mathcal{I}^1(V)} - 2\}, \\ \mathcal{B}(V, 0) &:= \{(V_{\mathcal{I}^2(V),\nu}, \nu, \gamma_{2,0,\nu,0}^{0,0,0} \mid \nu = 0, \dots, 3\} \\ &\quad \cup \{(E_{\mathcal{I}^2(V),\nu}, i', \gamma_{2,1,\nu,i'}^{0,0,0} \mid i' = 0, \dots, p_{E_{\mathcal{I}^2(V)}} - 2\} \\ &\quad \cup \{(\mathcal{I}^2(V), \alpha(i', j'), \gamma_{2,2,0,\alpha(i',j')}^{0,0,0} \mid i' = 0, \dots, p_{\mathcal{I}^2(V)} - 4 + \tau_0; \\ &\quad \quad \quad j' = 0, \dots, i'\} \end{aligned}$$

Für eine k -Kanten-irreguläre bzw. k -Facetten-irreguläre Kante E mit dem Index r setzt man für $i = 0, \dots, p_E - 2$:

$$\begin{aligned} \mathcal{B}(E, i) &:= \{(\mathcal{I}^1(E), i', \gamma_{1,1,0,i'}^{1,r,i} \mid i' = 0, \dots, p_{\mathcal{I}^1(E)} - 2\} \\ \mathcal{B}(E, i) &:= \{(E_{\mathcal{I}^2(E),\nu}, i', \gamma_{2,1,\nu,i'}^{1,r,i} \mid i' = 0, \dots, p_{E_{\mathcal{I}^2(E),\nu}} - 2\} \\ &\quad \cup \{(\mathcal{I}^2(E), \alpha(i, j), \gamma_{2,2,0,\alpha(i,j)}^{1,r,i} \mid i' = 0, \dots, p_{\mathcal{I}^2(E)} - 4 + \tau_0; \\ &\quad \quad \quad j' = 0, \dots, i'\} \end{aligned}$$

Für eine k -irreguläre Facette F sei für $\kappa = 0, \dots, r_{1,F} - 1$:

$$\mathcal{B}(F, \kappa) := \{(\mathcal{I}^2(F), \alpha(i, j), \gamma_{2,2,0,\alpha(i,j)}^{2,r,\kappa} \mid i' = 0, \dots, p_{\mathcal{I}^2(F)} - 4 + \tau_0; \\ j' = 0, \dots, i'\}$$

Hiermit definiert man rekursiv für einen Freiheitsgradträger C und einem lokalen Freiheitsgradindex κ :

$$\mathcal{C}(C, \kappa) := \{(B, b, \beta) \mid (B, b, \beta) \in \mathcal{B}(C, \kappa), B \text{ ist regulär}, B \notin \Gamma_0\} \bigoplus_{\substack{(B, b, \beta) \in \mathcal{B}(C, \kappa), \\ B \text{ ist irregulär}, B \notin \Gamma_0}} \beta \mathcal{C}(B, b)$$

mit $\beta\{(B_0, b_0, \beta_0), (B_1, b_1, \beta_1), \dots\} := \{(B_0, b_0, \beta\beta_0), (B_1, b_1, \beta\beta_1), \dots\}$ und

$$\begin{aligned} \mathcal{C}_0 \uplus \mathcal{C}_1 &:= \{(B, b, \beta) \mid (B, b, \beta) \in \mathcal{C}_0, \nexists \beta' : (B, b, \beta') \in \mathcal{C}_1\} \\ &\cup \{(B, b, \beta) \mid (B, b, \beta) \in \mathcal{C}_1, \nexists \beta' : (B, b, \beta') \in \mathcal{C}_0\} \\ &\cup \{(B, b, \beta + \beta') \mid (B, b, \beta) \in \mathcal{C}_0, \exists \beta' : (B, b, \beta') \in \mathcal{C}_1\}. \end{aligned}$$

Die Indizierung der Kopplungsgewichte $\gamma_{d, c_0, \nu, c_1}^{h_0, r, h_1} \in \mathbb{R}$ ist folgendermaßen zu interpretieren: Die Indizes h_0 und c_0 bezeichnen die Dimension des irregulären Elements B bzw. des zugehörigen Kopplungselements $V_{\mathcal{I}^d(B), \nu}$, $E_{\mathcal{I}^d(B), \nu}$ oder $\mathcal{I}^d(B)$. Die Indizes h_1 und c_1 bezeichnen den Index der lokalen Basisfunktionen, die bzgl. des irregulären Elements bzw. des Kopplungselements definiert sind. Die Kopplungsgewichte können für die hier aufgeführten Basen $\{\xi_{2,i}^q\}$ und $\{\xi_{3,i}^q\}$ rekursiv ermittelt werden. Hierzu sei auf Anhang A.2 verwiesen.

Anhand des folgenden Beispiels wird die Bedeutung der Mengen $\mathcal{B}(\cdot, \cdot)$ und $\mathcal{C}(\cdot, \cdot)$ deutlich:

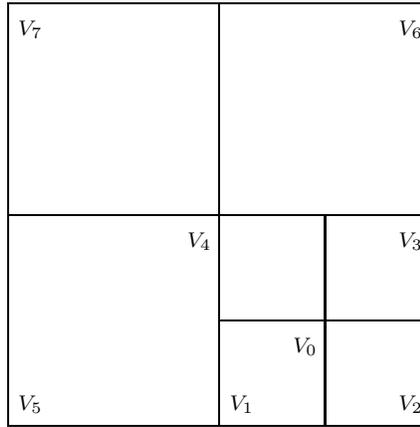


Abb. III.21: Kopplungen des 2-Facetten-irregulären Knotens V_0 (vgl. Abbildung III.20).

Abbildung III.21 zeigt die Kopplungen des 2-Facetten-irregulären Knotens V_0 bei $p \equiv 1$. Laut Anhang A.2 erhält man die knotenbezogenen Kopplungsgewichte $\gamma_{1,0,\nu,0}^{0,0,0} = 0.5$ und $\gamma_{2,0,\nu,0}^{0,0,0} = 0.25$. Hiermit gewinnt man:

$$\begin{aligned} \mathcal{B}(V_0, 0) &= \{(V_1, 0, 0.25), (V_2, 0, 0.25), (V_3, 0, 0.25), (V_4, 0, 0.25)\} \\ \mathcal{B}(V_1, 0) &= \{(V_5, 0, 0.5), (V_2, 0, 0.5)\}, \end{aligned}$$

$$\begin{aligned}
\mathcal{B}(V_2, 0) &= \{(V_2, 0, 1)\}, \\
\mathcal{B}(V_3, 0) &= \{(V_2, 0, 0.5), (V_6, 0, 0.5)\}, \\
\mathcal{B}(V_4, 0) &= \{(V_5, 0, 0.25), (V_2, 0, 0.25), (V_6, 0, 0.25), (V_7, 0, 0.25)\} \\
\mathcal{B}(V_5, 0) &= \{(V_5, 0, 1)\} \\
\mathcal{B}(V_6, 0) &= \{(V_6, 0, 1)\} \\
\mathcal{B}(V_7, 0) &= \{(V_7, 0, 1)\}
\end{aligned}$$

Daraus folgt die Kopplungsmenge:

$$\begin{aligned}
\mathcal{C}(V_0, 0) &= \{(V_2, 0, 0.25)\} \uplus ((0.25\mathcal{C}(V_1, 0) \uplus 0.25\mathcal{C}(V_3, 0)) \uplus 0.25\mathcal{C}(V_4, 0)) \\
&= \{(V_2, 0, 0.25)\} \uplus (\{(V_5, 0, 0.125), (V_2, 0, 0.125)\} \\
&\quad \uplus \{(V_2, 0, 0.125), (V_6, 0, 0.125)\}) \uplus 0.25\mathcal{C}(V_4, 0) \\
&= \{(V_2, 0, 0.25)\} \uplus (\{(V_5, 0, 0.125), (V_2, 0, 0.25), (V_6, 0, 0.125)\} \\
&\quad \uplus \{(V_5, 0, 0.0625), (V_2, 0, 0.0625), (V_6, 0, 0.0625), (V_7, 0, 0.0625)\}) \\
&= \{(V_2, 0, 0.25)\} \\
&\quad \uplus \{(V_5, 0, 0.1875), (V_2, 0, 0.3125), (V_6, 0, 0.1875), (V_7, 0, 0.0625)\} \\
&= \{(V_5, 0, 0.1875), (V_2, 0, 0.5625), (V_6, 0, 0.1875), (V_7, 0, 0.0625)\}
\end{aligned}$$

Die Abbildungen $\zeta : \tilde{\mathcal{V}} \cup \tilde{\mathcal{E}} \cup \mathcal{T}_h \rightarrow \mathbb{N}_0$ bzw. $\zeta : \tilde{\mathcal{V}} \cup \tilde{\mathcal{E}} \cup \tilde{\mathcal{F}} \cup \mathcal{T}_h \rightarrow \mathbb{N}_0$ für die globale Nummerierung seien wie in Abschnitt III.2.1 bzgl. der Mengen $\tilde{\mathcal{V}} \subset \mathcal{V}$, $\tilde{\mathcal{E}} \subset \mathcal{E}$ und $\tilde{\mathcal{F}} \subset \mathcal{F}$ definiert, wobei diese Mengen nur die regulären Knoten, Kanten bzw. Facetten von \mathcal{T}_h enthalten.

Für den zweidimensionalen Fall kann hiermit π_T^2 folgendermaßen definiert werden: Es seien $T \in \mathcal{T}_h$ sowie V_0, \dots, V_3 die Ecken von T und E_0, \dots, E_3 die Kanten. Wenn $V_\nu \in \mathcal{V}$ bzw. $E_\nu \in \mathcal{E}$ ist, setzt man:

$$\begin{aligned}
\pi_{T, \zeta(B)+b, \nu}^2 &:= \beta, \quad \nu = 0, \dots, 3; (B, b, \beta) \in \mathcal{C}(V_\nu, 0) \\
\pi_{T, \zeta(B)+b, 4+\nu(p_{E_\nu}-1)+i}^2 &:= (\theta_{T, \nu})^i \beta, \quad i = 0, \dots, p_{E_\nu} - 2; (B, b, \beta) \in \mathcal{C}(E_\nu, i) \\
\pi_{T, \zeta(T)+\alpha(i, j), 4p_T+\alpha(i, j)}^2 &:= 1, \quad i = 0, \dots, p_T - 4 + \tau; j = 0, \dots, i.
\end{aligned}$$

Im dreidimensionalen Fall erhält man für π_T^3 : Es seien $T \in \mathcal{T}_h$ sowie V_0, \dots, V_7 die Ecken, E_0, \dots, E_{11} die Kanten und F_0, \dots, F_5 die Facetten von T .

Falls $V_\nu \in \mathcal{V}$, $E_\nu \in \mathcal{E}$ bzw. $F_\nu \in \mathcal{F}$ ist, setzt man:

$$\begin{aligned} \pi_{T,\zeta(B)+b,\nu}^3 &:= \beta, \quad \nu = 0, \dots, 3; (B, b, \beta) \in \mathcal{C}(V_\nu, 0) \\ \pi_{T,\zeta(B)+b,8+\nu(p_{E_\nu}^\varepsilon - 1)+i}^3 &:= (\theta_{T,\nu}^1)^i \beta, \quad i = 0, \dots, p_{E_\nu} - 2; (B, b, \beta) \in \mathcal{C}(E_\nu, i) \\ \pi_{T,\zeta(B)+b,r_0+\nu r_1+\alpha(i,j)}^3 &:= \theta_{T,\nu}^2 \beta, \quad i = 0, \dots, p_{F_\nu} - 4 + \tau_0; j = 0, \dots, i; \\ &\quad (B, b, \beta) \in \mathcal{C}(F_\nu, \alpha(i, j)) \\ \pi_{T,\zeta(T)+\beta(i,j,l),r_2+\beta(i,j,l)}^3 &:= 1, \quad i = 0, \dots, q - 6 + \tau_1, j = 0, \dots, i, l = 0, \dots, j \end{aligned}$$

Satz III.14 *Es sei \mathcal{T}_h ein reguläres oder irreguläres Gitter von $\Omega \subset \mathbb{R}^k$, $k \in \{2, 3\}$, und die Transformationen $\{\Psi_T\}$ seien bzgl. \mathcal{Q}^k aus Abb. III.18 bzw. III.19 gerichtet. Dann existieren Funktionen $\{\varphi_i\} \subset S^p(\mathcal{T}_h)$, so dass $\varphi_i \circ \Psi_T = \pi_{T,ij}^k \mathcal{S}_{k,j}^{pT}$ mit $\theta_{T,\nu}^1 \equiv \theta_{T,ij,\nu}^2 \equiv 1$ ist. Im Fall $\tau = 2$ bzw. $\tau_0 = 2$ und $\tau_1 = 4$ ist $\{\varphi_i\}$ eine Basis von $S^p(\mathcal{T}_h)$.*

Auf einen Beweis wird an dieser Stelle verzichtet, da er sich im Wesentlichen aus den zuvor definierten (Daten-)Strukturen ergibt.

III.3 Lösungsverfahren für quadratische Minimierungsprobleme

Die in Abschnitt III.2 aufgeführten Probleme (III.48), (III.49)-(III.50), (III.51)-(III.52) und (III.53)-(III.54) sind bzw. führen auf quadratische Minimierungsprobleme mit linearen oder nichtlinearen Nebenbedingungen. In der Literatur existiert hierzu eine Fülle von Lösungsverfahren, die unter dem Begriff der nichtlinearen Optimierung und insbesondere des Quadratic-Programmings zusammengefasst werden. Für einen Überblick sei auf [50], [54], [70] oder [83] verwiesen. Speziell für das Problem (III.48) wird in der Finite-Elemente-Literatur häufig ein projektives SOR-Verfahren vorgeschlagen⁴⁵, das in Abschnitt III.3.1 auch für Minimierungsprobleme formuliert wird, die den diskretisierten reibungsfreien Kontaktproblemen aus Abschnitt III.1.2 zugeordnet sind.

Verfahren zur Lösung der Sattelpunktprobleme (III.49)-(III.50), (III.51)-(III.52) und (III.53)-(III.54) werden in Abschnitt III.3.2 skizziert. Insbesondere wird mit Hilfe einer Umformulierung der Sattelpunktprobleme in quadratische Minimierungsprobleme der Zugang zu Standard-Optimierungsverfahren ermöglicht.

⁴⁵vgl. S.67 in [55], S.40 in [56], S.20 in [68], S.128 in [72]

III.3.1 Das SOR-Verfahren mit Projektion

Ein einfaches, leicht implementierbares Verfahren zum Lösen von Minimierungsproblemen vom Typ (III.48) ist das SOR-Verfahren mit Projektion. Es handelt sich dabei um eine Modifikation des bekannten Successive-Overrelaxation-Verfahrens mit einer zusätzlichen Projektion der Iterierten auf die Restriktionsmenge \bar{K} in jedem Iterationsschritt.

Betrachtet wird zunächst das Minimierungsproblem (III.48) bezüglich des diskretisierten vereinfachten Signorini-Problems (III.17) mit K_h wie in (III.18) sowie bezüglich des diskretisierten Hindernisproblems (III.41) mit K_h wie in (III.42).

Für $V_h := S^p(\mathcal{T}_h)$ mit $p \equiv 1$ und $\{\varphi_i\}$ wie in Abschnitt III.2.1 gewinnt man:

$$\bar{K} = \{\bar{v} \in \mathbb{R}^m \mid \forall V \in \mathcal{V}' : \bar{v}_{\zeta(V)} \geq g(V)\}.$$

Im Fall (III.18) ist $\mathcal{V}' := \mathcal{V} \cap \Gamma_1$, und im Fall (III.42) ist $\mathcal{V}' := \mathcal{V}$. Damit ist die Abbildung $P : \mathbb{R}^m \rightarrow \mathbb{R}^m$ mit $P(\bar{v}) := (P_0(\bar{v}_0), \dots, P_{m-1}(\bar{v}_{m-1}))$ für $0 \leq i < m$ und

$$P_i(\bar{v}_i) := \begin{cases} g(V), & \exists V \in \mathcal{V}' : i = \zeta(V), \bar{v}_i \leq g(V) \\ \bar{v}_i, & \text{sonst} \end{cases}$$

eine Projektion auf \bar{K} .

Ein Iterationsschritt des SOR-Verfahrens mit Projektion kann für einen Relaxationsparameter $0 < \omega < 2$ durch die folgende Vorschrift formuliert werden: Für $i = 0, \dots, m-1$ sei

$$\hat{u} \leftarrow \bar{u}^n, \quad \hat{u}_i \leftarrow 0 \tag{III.61}$$

$$\bar{u}_i^{n+1} \leftarrow P_i(\bar{u}_i^n + \omega((L_i - A_{i,\cdot}\hat{u})/A_{ii} - \bar{u}_i^n)), \tag{III.62}$$

wobei $A_{i,\cdot}$ den i -ten Zeilenvektor von A bezeichnet.

Satz III.15 *Es sei \bar{u} Lösung von (III.48) mit $V_h := S^p(\mathcal{T}_h)$ und $p \equiv 1$. Die Folge $\{\bar{u}^n\}$ sei gegeben durch (III.61)-(III.62). Dann gilt:*

$$\forall \bar{u}^0 \in \bar{K} : \lim_{n \rightarrow \infty} \bar{u}^n = \bar{u}.$$

Beweis: Prop.II.2.3 in [56], Ch.2, Th.1.3 in [55]. □

Die Konvergenzgeschwindigkeit des SOR-Verfahrens mit Projektion ist erwartungsgemäß nicht besser als die des reinen SOR-Verfahrens. Mit Hilfe eines Zusatzschrittes kann das Verfahren jedoch wesentlich beschleunigt werden, wobei das Verfahren zu einem SSOR-Verfahren erweitert wird. Der Ansatz besteht darin, eine verbesserte Lösung in dem durch die Suchrichtungen $r := \bar{u}^n - \bar{u}^{n-1}$ und $s := \bar{u}^{n+1} - \bar{u}^n$ aufgespannten affinen Unterraum $\bar{u}^{n+1} + \text{span}\{r, s\}$ zu suchen, wobei r und s zusätzlich bezüglich der Restriktion geeignet angepasst werden. Hierzu ist lediglich das zweidimensionale Problem

$$\min_{(\alpha, \beta) \in \mathbb{R}^2} \frac{1}{2} (\bar{u}^{n+1} + \alpha r + \beta s)^\top A (\bar{u}^{n+1} + \alpha r + \beta s) - (\bar{u}^{n+1} + \alpha r + \beta s)^\top L$$

zu lösen. Ein Iterationsschritt der beschleunigten Variante lautet:

$$r \leftarrow \bar{u}^n - \bar{u}^{n-1},$$

$$\bar{u}^{n+1} \leftarrow \left\{ \begin{array}{l} i = 0, \dots, m-1 \\ i = m-1, \dots, 0 \end{array} \right\} \text{ (III.61)-(III.62)},$$

$$s \leftarrow \bar{u}^{n+1} - \bar{u}^n,$$

$$\exists V \in \mathcal{V}' : \bar{u}_{\zeta(V)}^{n+1} + r_{\zeta(V)} < g(V) \vee \bar{u}_{\zeta(V)}^{n+1} + s_{\zeta(V)} < g(V) :$$

$$r_{\zeta(V)} \leftarrow s_{\zeta(V)} \leftarrow 0,$$

$$d_0 \leftarrow r^\top (L - A\bar{u}^{n+1}), \quad d_1 \leftarrow r^\top A s, \quad d_2 \leftarrow r^\top A r,$$

$$\beta \leftarrow ((s^\top (L - A\bar{u}^{n+1}))d_2 - d_0 d_1) / ((s^\top A s)d_2 - d_1^2), \quad \alpha \leftarrow (d_0 - \beta d_1) / d_2,$$

$$\bar{u}^{n+1} \leftarrow \bar{u}^{n+1} + \alpha r + \beta s.$$

Für weitere Details sei auf [24] verwiesen. In dieser Arbeit wird insbesondere die Einbettung des Algorithmus in ein kaskadisches Mehrgitterverfahren diskutiert.

m	SOR mit Projektion			
	ZFW	Ref-ZFW	Tol.	n
16	-0.26241	0.02329	1.0E-1	5
64	-0.24365	0.00452	1.0E-2	8
256	-0.24050	0.00138	1.0E-3	36
1024	-0.23946	3.35E-4	1.0E-4	179
4096	-0.23920	7.84E-5	1.0E-4	649
16384	-0.23914	1.55E-5	1.0E-5	3163
65536	-0.23912	3.16E-8	1.0E-5	11627

Tab. III.1: Iterationsverhalten des SOR-Verfahrens mit Projektion.

SSOR mit Projektion (beschleunigt)				
m	ZFW	Ref-ZFW	Tol.	n
16	-0.26262	0.02349	1.0E-1	3
64	-0.24366	0.00453	1.0E-2	6
256	-0.24050	0.00138	1.0E-2	10
1024	-0.23946	3.35E-4	1.0E-3	22
4096	-0.23920	7.85E-5	1.0E-3	50
16384	-0.23914	1.55E-5	1.0E-4	133
65536	-0.23912	2.35E-12	1.0E-5	335

Tab. III.2: Iterationsverhalten des beschleunigten SSOR-Verfahrens mit Projektion.

Die Tabellen III.1 und III.2 zeigen das Iterationsverhalten des SOR-Verfahrens mit Projektion und des beschleunigten SSOR-Verfahrens. Als Beispielkonfiguration wurde das Hindernisproblem aus Abschnitt II.5.1 gewählt. Für die Diskretisierungen wurde $p \equiv 1$ gesetzt, und ein reguläres Rechtecksgitter global verfeinert. Für jede Verfeinerungsstufe wurden n Iterationen benötigt, bis die vorgegebene Toleranz (Tol.) den Wert von $\|\bar{u}^n - \bar{u}^{n-1}\|_2$ unterschritten hat. Die Toleranz ist als die größte Zehnerpotenz gewählt, so dass sich die Differenz von dem Zielfunktionswert ZFW und dem Referenzzielfunktionswert Ref = $-0.2391286\dots$ (gerechnet mit $m = 65536$ und einer Toleranz von 10^{-7}) bei Erreichen der Toleranz stabilisiert. Der durch das Verfahren bedingte Fehler liegt damit unterhalb des Diskretisierungsfehlers. Erkennbar ist, dass das beschleunigte SSOR-Verfahren mit Projektion deutlich weniger Iterationen als das SOR-Verfahren benötigt, auch wenn berücksichtigt wird, dass die beschleunigte Variante doppelt so viele SOR-Schritte durchführt. Zudem reicht für das beschleunigte Verfahren die Vorgabe einer geringeren Toleranz aus, um den Diskretisierungsfehler zu erreichen.

SQOPT				
m	ZFW	Ref-ZFW	MOT	Fkt.
16	-0.26266	0.02353	1.0E-1	39
64	-0.24366	0.00453	1.0E-2	138
256	-0.24051	0.00138	1.0E-2	559
1024	-0.23946	3.35E-4	1.0E-2	2206
4096	-0.23920	7.86E-5	1.0E-3	8812

Tab. III.3: Iterationsverhalten von SQOPT.

Tabelle III.3 zeigt das Iterationsverhalten des kommerziellen Optimierungspakets SQOPT [51], das speziell für großdimensionierte konvexe, quadratische Minimierungsprobleme mit Gleichheits- und Ungleichheitsnebenbedingungen entwickelt wurde.

Die angegebene Toleranz (MOT) entspricht dem Programmparameter MINOR OPTIMALITY TOLERANCE von SQOPT.⁴⁶ Zu sehen ist insbesondere, dass die Anzahl der Funktionsauswertungen (Fkt.), die in etwa jeweils einem SOR-Schritt entsprechen, deutlich höher als die Anzahl der Iterationsschritte der SOR-Verfahren ist. Letztendlich macht dieser Vergleich deutlich, dass der Einsatz kommerzieller Pakete mit Black-Box Charakter nicht immer zu den erhofften Effizienzsteigerungen führt, und deshalb mit großer Vorsicht durchgeführt werden sollte. Zu bedenken ist, dass kommerzielle Optimierungsverfahren in der Regel für wesentlich größere Problemklassen konzipiert sind, so dass Effizienzeinbußen häufig nicht zu vermeiden sind. Außerdem führen Finite-Elemente-Diskretisierungen im Allgemeinen zu Gleichungssystemen mit sehr großen Dimensionierungen. Kommerzielle Verfahren können jedoch häufig nur Systeme mit eher moderater Größe effizient behandeln.⁴⁷

Zum Abschluss dieses Abschnitts wird geklärt, in welcher Weise das SOR-Verfahren mit Projektion auch für das Minimierungsproblem (III.48) bezüglich des diskretisierten reibungsfreien Kontaktproblems (III.31) mit $K_h := \{v_h \in S^p(\mathcal{T}_h)^3 \mid \delta_n(v_h) \leq g\}$ eingesetzt werden kann. Hier erhält man:⁴⁸

$$\bar{K} = \{\bar{v} \in \mathbb{R}^m \mid \forall V \in \mathcal{V}' : n_{V,\iota} \cdot \bar{v}_{\zeta(V)+\iota} \leq g(V)\},$$

wobei $\mathcal{V}' := \mathcal{V} \cap \Gamma_1$ ist, und n_V den äußeren Normalenvektor in $V \in \mathcal{V}'$ mit Länge 1 bezeichnet.

Eine Projektion P der Form $P(\bar{v}) = (P_0(\bar{v}_0), \dots, P_{m-1}(\bar{v}_{m-1}))$ ist für die Restriktion \bar{K} nicht unmittelbar angebbbar. Um dennoch das SOR-Verfahren mit Projektion (bzw. die beschleunigte Variante) zu nutzen, ist eine geeignete Variablentransformation erforderlich. Es sei $\bar{y} := (I + M)\bar{v}$, wobei $I \in \mathbb{R}^{m \times m}$ die Einheitsmatrix ist, und die Matrix $M \in \mathbb{R}^{m \times m}$ sei durch

$$\begin{aligned} M_{\zeta(V)+\rho(V),\zeta(V)+\rho(V)} &:= n_{\rho(V)} - 1 \\ M_{\zeta(V)+\rho(V),\zeta(V)+(\rho(V)+1)\%3} &:= n_{V,(\rho(V)+1)\%3} \\ M_{\zeta(V)+\rho(V),\zeta(V)+(\rho(V)+2)\%3} &:= n_{V,(\rho(V)+2)\%3} \end{aligned}$$

definiert.⁴⁹ Hierbei bezeichnet $\rho(V) \in \{0, 1, 2\}$ den Index der betragsgrößten

⁴⁶vgl. S.29 in [51]

⁴⁷Das SQOPT-Verfahren war bei den durchgeführten Testrechnungen nicht in der Lage, Verfeinerungsstufen mit mehr als 16000 Freiheitsgraden mit den dem Autor zur Verfügung stehenden Rechnerressourcen zu bewältigen. Das hierfür erforderliche zusätzliche Speicheraufkommen (Real-Storage, S.14 in [51]) betrug wesentlich mehr als 2GB.

⁴⁸im Sinne der Einsteinschen Summenkonvention

⁴⁹Hier bezeichnet $i\%k$ den ganzzahligen Rest bei ganzzahliger Division von i durch k .

Komponente von n_V , es ist demnach $n_{V,\rho(V)} := \max\{|n_{V,0}|, |n_{V,1}|, |n_{V,2}|\}$. Offenbar gilt:

$$\bar{v} \in \bar{K} \Leftrightarrow \forall V \in \mathcal{V}' : \bar{y}_{\zeta(V)+\rho(V)} \leq g(V).$$

Also ist $\bar{u} := (I + M)^{-1}\bar{x} \in \bar{K}$ genau dann Lösung von (III.48), wenn \bar{x} den Ausdruck

$$\min_{\bar{y} \in \bar{K}^*} \frac{1}{2} \bar{y}^\top A^* \bar{y} - \bar{y}^\top L^* \quad (\text{III.63})$$

mit

$$A^* := ((I + M)^{-1})^\top A (I + M)^{-1}, \quad L^* := ((I + M)^{-1})^\top L \\ \bar{K}^* := \{\bar{y} \in \mathbb{R}^m \mid \mathcal{V}' : \bar{y}_{\zeta(V)+\rho(V)} \leq g(V)\}$$

minimiert. Zur Lösung von (III.63) kann das SOR-Verfahren mit der Projektion $P^* : \mathbb{R}^m \rightarrow \mathbb{R}^m$, $P^*(\bar{y}) := (P_0^*(\bar{y}_0), \dots, P_{m-1}^*(\bar{y}_{m-1}))$ und

$$P_i^*(\bar{y}_i) := \begin{cases} g(V), & \exists V \in \mathcal{V}' : i = \zeta(V) + \rho(V), \bar{y}_i \geq g(V) \\ \bar{y}_i, & \text{sonst} \end{cases}$$

eingesetzt werden. Es ist leicht einzusehen, dass $(I + M)^{-1} = (I + \tilde{M})$ mit

$$\tilde{M}_{\zeta(V)+\rho(V), \zeta(V)+\rho(V)} := (1 - n_{V,\rho(V)})/n_{V,\rho(V)} \\ \tilde{M}_{\zeta(V)+\rho(V), \zeta(V)+(\rho(V)+1)\%3} := -n_{V,(\rho(V)+1)\%3}/n_{V,\rho(V)} \\ \tilde{M}_{\zeta(V)+\rho(V), \zeta(V)+(\rho(V)+2)\%3} := -n_{V,(\rho(V)+2)\%3}/n_{V,\rho(V)}$$

gilt. Damit erhält für $i = \zeta(V) + (\rho(V) + \iota)\%3$ mit $V \in \mathcal{V}'$ und für $\iota \in \{0, 1, 2\}$:

$$A_{i,\cdot}^* \bar{y} = ((A + \tilde{M}^\top A)(I + \tilde{M}))_{i,\cdot} \bar{y} \\ = A_{i,\cdot} (I + \tilde{M}) \hat{z} + (\tilde{M}^\top A)_{i,\cdot} (I + \tilde{M}) \bar{y} \\ = A_{i,\cdot} (I + \tilde{M}) \bar{y} + \tilde{M}_{\zeta(V)+\rho(V),i} A_{\zeta(V)+\rho(V),\cdot} (I + \tilde{M}) \bar{y}, \\ L_i^* = ((I + \tilde{M})^\top L)_i \\ = L_i + \tilde{M}_{\zeta(V)+\rho(V),i} L_{\zeta(V)+\rho(V)}, \\ A_{ii}^* = A_{ii} + 2(\tilde{M}A)_{ii} + (\tilde{M}^\top A \tilde{M})_{ii} \\ = A_{ii} + 2\tilde{M}_{\zeta(V)+\rho(V),i} A_{i,\zeta(V)+\rho(V)} \\ + \tilde{M}_{\zeta(V)+\rho(V),i}^2 A_{\zeta(V)+\rho(V),\zeta(V)+\rho(V)}.$$

In allen anderen Fällen gilt:

$$A_{i,\cdot}^* \bar{y} = A_{i,\cdot} (I + \tilde{M}) \bar{y}, \quad L_i^* = L_i, \quad A_{ii}^* = A_{ii}$$

Ein Iterationsschritt des SOR-Verfahrens mit Projektion lässt sich hiermit für einen Relaxationsparameter $0 < \omega < 2$ in der folgenden Form angeben:

Für $i = 0, \dots, m-1$ sei

$$\begin{aligned} \hat{x} &\leftarrow \bar{x}_i^n, \quad \hat{x}_i \leftarrow 0, \\ \hat{x}_{\zeta(V)+\rho(V)} &\leftarrow \hat{x}_{\zeta(V)+\rho(V)} + \sum_{\iota=0}^2 \tilde{M}_{\zeta(V)+\rho(V), \zeta(V)+\rho(V)+\iota} \cdot \hat{x}_{\zeta(V)+\iota}, \\ d_0 &\leftarrow A_{i,\cdot} \hat{x}, \quad d_1 \leftarrow L_i, \quad d_2 \leftarrow A_{ii}, \\ \exists V \in \mathcal{V}', \iota \in \{0, 1, 2\} : i &= \zeta(V) + (\rho(V) + \iota) \% 3 : \\ d_0 &\leftarrow d_0 + \tilde{M}_{\zeta(V)+\rho(V), i} A_{\zeta(V)+\rho(V), \cdot} \hat{x}, \\ d_1 &\leftarrow d_1 + \tilde{M}_{\zeta(V)+\rho(V), i} L_{\zeta(V)+\rho(V)}, \\ d_2 &\leftarrow d_2 + 2\tilde{M}_{\zeta(V)+\rho(V), i} A_{i, \zeta(V)+\rho(V)}, \\ &\quad + \tilde{M}_{\zeta(V)+\rho(V), i}^2 A_{\zeta(V)+\rho(V), \zeta(V)+\rho(V)}, \\ \bar{x}_i^{n+1} &\leftarrow P_i^* (\bar{x}_i^n + \omega((d_1 - d_0)/d_2 - \bar{x}_i^n)). \end{aligned}$$

SOR mit Projektion				
m	ZFW	Ref-ZFW	Tol.	n
18	2.17E-4	0.00261	1.0E-2	3
90	0.00243	3.97E-4	1.0E-3	18
540	0.00198	8.48E-4	1.0E-3	24
3672	0.00263	2.05E-4	1.0E-3	45
26928	0.00283	0.0	1.0E-4	325

Tab. III.4: Iterationsverhalten des SOR-Verfahrens mit Projektion.

SQOPT				
m	ZFW	Ref-ZFW	MOT	n
18	2.15E-4	0.00262	1.0E-3	22
90	0.00244	3.92E-4	1.0E-3	68
540	0.00198	8.49E-4	1.0E-4	854
3672	0.00267	1.65E-4	4.5E-14	3402

Tab. III.5: Iterationsverhalten von SQOPT.

In den Tabellen III.4 und III.5 ist das Iterationsverhalten des SOR-Verfahrens mit Projektion und Variablentransformation sowie das Iterationsverhalten des SQOPT-Verfahrens dargestellt. Das zugrunde liegende Problem

entspricht dem reibungsfreien Kontaktproblem aus Abschnitt II.4.1. Auch hier zeigt sich, dass das kommerzielle Programmpaket SQOPT für dieses Beispiel eher eingeschränkt einsetzbar ist. Auffällig ist insbesondere die sehr kleine Toleranzvorgabe in Zeile 4.

Entscheidend für die in diesem Abschnitt vorgestellten SOR-Verfahren ist die Angabe einer Projektion $P : \mathbb{R}^m \rightarrow \mathbb{R}^m$ der Form $P(\bar{v}) = (P_0(\bar{v}_0), \dots, P_{m-1}(\bar{v}_{m-1}))$. Werden die hierarchischen Basisfunktionen aus Abschnitt III.2.1 für $p > 1$ verwendet, ist die Angabe einer Projektion dieser Form nicht möglich, da die Komponenten des Lösungsvektors nicht als Funktionswerte von u_h in bestimmten Punkten interpretierbar sind.

Ein Ausweg besteht in der Verwendung von Verfahren, die auf Sattelpunktformulierungen basieren, die im Folgenden erläutert werden sollen.

III.3.2 Lösungsverfahren für Sattelpunktprobleme

Zur Lösung von Sattelpunktproblemen der Form (III.49)-(III.50), (III.51)-(III.52) und (III.53)-(III.54) können ebenfalls projektive Verfahren formuliert werden. In der Literatur häufig anzutreffen sind diesbezüglich projektive Verfahren vom Uzawa- bzw. Arrow-Hurwicz-Typ.⁵⁰

Unter Verwendung der Notation aus Abschnitt III.2 seien $P : \mathbb{R}^{\tilde{m}} \rightarrow \mathbb{R}^{\tilde{m}}$ eine Projektion auf $\bar{\Lambda}$, $B \in \mathbb{R}^{m \times \tilde{m}}$ und $\tilde{g} \in \mathbb{R}^{\tilde{m}}$. Im Fall (III.49)-(III.50) bzw. (III.51)-(III.52) setzt man $\tilde{m} := m_r$, $\bar{\Lambda} := \bar{\Lambda}_r$, $B := B_r$ und $\tilde{g} := \bar{g}$ bzw. $\tilde{g} := 0$ sowie im Fall (III.53)-(III.54) $\tilde{m} := m_0 + m_1$, $\bar{\Lambda} := \bar{\Lambda}_0 \times \bar{\Lambda}_1$, $B := (B_0, B_1)$ und $\tilde{g} := (\bar{g}, 0)$.

Ein Iterationsschritt des projektiven Uzawa-Verfahrens hat für $0 < \hat{\varrho}_0 \leq \varrho_n \leq \hat{\varrho}_1$ und $\bar{\lambda}^0 \in \bar{\Lambda}$ die Form:

$$\bar{u}^n \leftarrow A^{-1}(L - B\lambda^n) \quad (\text{III.64})$$

$$\lambda^{n+1} \leftarrow P(\lambda^n - \varrho_n(B^\top \bar{u}^n - \tilde{g})). \quad (\text{III.65})$$

Ein Iterationsschritt des projektiven Arrow-Hurwicz- (oder Gradienten-) Verfahrens lautet für $\hat{\varrho}_0, \hat{\varrho}_1 > 0$ und $\bar{\lambda}^0 \in \bar{\Lambda}$:

$$\bar{u}^{n+1} \leftarrow \bar{u}^n - \hat{\varrho}_0 S^{-1}(A\bar{u}^n - L + B\lambda^n) \quad (\text{III.66})$$

$$\lambda^{n+1} \leftarrow P(\lambda^n - \hat{\varrho}_1(B^\top \bar{u}^{n+1} - \tilde{g})), \quad (\text{III.67})$$

wobei S eine beliebige, positiv definite Matrix ist, für die man üblicherweise die Einheitsmatrix wählt.

⁵⁰vgl. z.B. Ch.2, Sec.4.3 in [55]

Satz III.16 *Es sei $(\bar{u}, \bar{\lambda})$ Lösung von (III.49)-(III.50), (III.51)-(III.52) bzw. (III.53)-(III.54). Dann existieren $\hat{\rho}_0, \hat{\rho}_1 > 0$, so dass für die Folge $\{(\bar{u}^n, \bar{\lambda}^n)\}$, die durch (III.64)-(III.65) bzw. (III.66)-(III.67) definiert ist, gilt:*

$$\forall \bar{\lambda}^0 \in \bar{\Lambda} : \lim_{n \rightarrow \infty} (\bar{u}^n, \bar{\lambda}^n) = (\bar{u}, \bar{\lambda}).$$

Beweis: Ch.2,Th.4.1 bzw. Th.4.2 in [55]. □

Eine andere Möglichkeit, Näherungslösungen der hier untersuchten Sattelpunktprobleme zu gewinnen, besteht in der Umformulierung in Minimierungsprobleme bezüglich der Lagrangeschen Parameter und in der Anwendung von Standard-Optimierungsverfahren aus dem Bereich des Quadratic-Programmings oder der nichtlinearen Optimierung, wie sie in den oben angegebenen Referenzen zu finden sind. Insbesondere kann auch hier versucht werden, kommerzielle Programmpakete wie etwa das in Abschnitt III.3.1 bereits erwähnte Quadratic-Programming Paket SQOPT oder das für nichtlineare Optimierungsaufgaben konzipierte SQP-Optimierungspaket SNOPT [52], [53] einzusetzen.⁵¹

Mit der oben eingeführten Notation lautet das Sattelpunktproblem: Gesucht ist ein $(\bar{u}, \bar{\lambda}) \in \mathbb{R}^m \times \mathbb{R}^{\bar{m}}$, so dass gilt:

$$A\bar{u} = L - B\bar{\lambda} \tag{III.68}$$

$$\forall \bar{\mu} \in \bar{\Lambda} : (\bar{\mu} - \bar{\lambda})^\top (B^\top \bar{u} - \tilde{g}) \leq 0. \tag{III.69}$$

Auflösen von (III.68) nach \bar{u} und Einsetzen in (III.69) liefert:

$$\forall \bar{\mu} \in \bar{\Lambda} : (\bar{\mu} - \bar{\lambda})^\top (B^\top A^{-1}(L - B\bar{\lambda}) - \tilde{g}) \leq 0.$$

Also gilt:

$$\forall \bar{\mu} \in \bar{\Lambda} : (B^\top A^{-1}B\bar{\lambda} - (B^\top A^{-1}L - \tilde{g}))^\top (\bar{\mu} - \bar{\lambda}) \geq 0.$$

Da $B^\top A^{-1}B$ positiv semidefinit ist, folgt insgesamt aus Satz I.1:

⁵¹Für die oben skizzierten Verfahren vom Uzawa- bzw. Arrow-Hurwicz-Typ ist die Angabe einer geeigneten Projektion häufig sehr kompliziert. Die Umformulierung des Sattelpunktproblems in ein Minimierungsproblem und die Einbeziehung von Optimierungsverfahren (u.U. in Form kommerzieller Programmpakete) ist aus praktischer Sicht häufig vorteilhafter.

Satz III.17 *Das Paar $(\bar{u}, \bar{\lambda})$ ist genau dann Lösung von (III.68)-(III.69), wenn $\bar{u} = A^{-1}(L - B\bar{\lambda})$ ist, und wenn für $\bar{\lambda}$ gilt:*

$$Z(\bar{\lambda}) = \min_{\bar{\mu} \in \bar{\Lambda}} Z(\bar{\mu}), \quad Z(\bar{\mu}) := \frac{1}{2} \bar{\mu}^\top B^\top A^{-1} B \bar{\mu} - \bar{\mu}^\top (B^\top A^{-1} L - \tilde{g}). \quad (\text{III.70})$$

Für die Anwendung von Optimierungsverfahren ist die Formulierung einer geeigneten Optimierungsaufgabe erforderlich, die etwa als Input für ein kommerzielles Optimierungspaket verwendet werden kann. Hierzu sei $\{\psi_{r,i}\}$ die gemäß Abschnitt III.2 konstruierte Basis von $U'_{r,H} = M^{p_1}(\mathcal{T}_{1,H})^{l_r}$. Die zugehörige lokale Basis von S_{k-1}^q sei mit $\{\tilde{\xi}_{k-1,i}^q\}$ bezeichnet. Es ist offenbar $\bar{\mu}_0 \in \bar{\Lambda}_0$ genau dann, wenn

$$\forall T \in \mathcal{T}_{1,H}, \forall x \in [-1, 1]^{k-1} : \sum_{j=0}^{m_{k-1}^{p_1,T} - 1} \bar{\mu}_{0, \zeta_1(T)+j} \tilde{\xi}_{k-1,j}^{p_1,T}(x) \geq 0 \quad (\text{III.71})$$

ist.⁵² Zudem ist $\bar{\mu}_1 \in \bar{\Lambda}_1$ genau dann, wenn gilt:

$$\forall T \in \mathcal{T}_{1,H}, \forall x \in [-1, 1] : \left| \sum_{j=0}^{m_1^{p_1,T} - 1} \bar{\mu}_{1, \zeta_1(T)+j} \tilde{\xi}_{1,j}^{p_1,T}(x) \right| \leq |s(\Psi_{1,T}(x))| \quad (\text{III.72})$$

bzw. $\forall T \in \mathcal{T}_{1,H}, \forall x \in [-1, 1]^2 :$

$$\begin{aligned} & \left(\sum_{j=0}^{m_2^{p_1,T} - 1} \bar{\mu}_{1, \zeta_1(T)+j} \tilde{\xi}_{2,j}^{p_1,T}(x) \right)^2 + \left(\sum_{j=0}^{m_2^{p_1,T} - 1} \bar{\mu}_{1, m_r/2 + \zeta_1(T)+j} \tilde{\xi}_{2,j}^{p_1,T}(x) \right)^2 \\ & \leq s(\Psi_{1,T}(x))^2. \end{aligned} \quad (\text{III.73})$$

Die Bedingungen (III.71) sowie (III.72) und (III.73) können in dieser Form nicht für eine diskrete Optimierungsaufgabe verwendet werden, da für alle $x \in [-1, 1]^{k-1}$ getestet wird. Eine Diskretisierung erhält man, indem man (III.71), (III.72) bzw. (III.73) nur für eine endliche Anzahl von Testpunkten $\{\hat{x}_{p_1,T,i}\}_{i=0,\dots,\hat{n}_{p_1,T}-1}$ formuliert, wobei $\hat{n}_{p_1,T}$ hinreichend groß gewählt

⁵²Der Ausdruck (III.71) bezieht sich auf linear-elastische Kontaktprobleme. Im Fall des Modellproblems vom Sginorini-Typ ist " \leq " in (III.71) zu verwenden.

werden sollte, so dass der Diskretisierungsfehler unbeeinflusst bleibt.⁵³
 Es seien hierzu $\zeta' : \mathcal{T}_{1,H} \rightarrow \mathbb{N}_0$ mit $\zeta'(T_s) := \sum_{i=0}^{s-1} \hat{n}_{p_1, T_i}$ und $\hat{n} := \sum_{T \in \mathcal{T}_{1,H}} \hat{n}_{p_1, T}$ sowie $C \in \mathbb{R}^{\hat{n} \times m_0}$ mit

$$C_{\zeta'(T)+i, \zeta_1(T)+j} := \tilde{\zeta}_{k-1, j}^q(\hat{x}_{T, i})$$

für $T \in \mathcal{T}_{1,H}$ sowie $i = 0, \dots, \hat{n}_{p_1, T} - 1$ und $j = 0, \dots, m_{k-1}^{p_1, T} - 1$.
 Die Bedingungen (III.71) und (III.72) können hiermit durch $C\bar{\mu}_0 \geq 0$ bzw. $-\bar{s} \leq C\bar{\mu}_1 \leq \bar{s}$ approximiert werden, wobei $\bar{s} \in \mathbb{R}^{\hat{n}}$ durch $\bar{s}_{\zeta'(T)+i} := s(\Psi_T(\hat{x}_{p_1, T, i}))$ für $T \in \mathcal{T}_{1,H}$ und $i = 0, \dots, \hat{n}_{p_1, T} - 1$ definiert ist. Für die Bedingung (III.73) erhält man $R_{\zeta'(T)+i}(\bar{\mu}) \leq 0$, wobei $R : \mathbb{R}^{m_1} \rightarrow \mathbb{R}^{\hat{n}}$ als $R_{\zeta'(T)+i}(\bar{\mu}) := (C_{\zeta'(T)+i, \cdot} \bar{\mu}_1^0)^2 + (C_{\zeta'(T)+i, \cdot} \bar{\mu}_1^1)^2 - \bar{s}_{\zeta'(T)+i}^2$ definiert ist, und $\bar{\mu}_1^0 := (\bar{\mu}_0, \dots, \bar{\mu}_{m_1/2-1})$ und $\bar{\mu}_1^1 := (\bar{\mu}_{m_1/2}, \dots, \bar{\mu}_{m_1-1})$ ist.
 Insgesamt kann (III.70) durch das folgende Optimierungsproblem näherungsweise dargestellt werden: Gesucht ist $\bar{\lambda} \in \bar{\Lambda}'$, so dass gilt:

$$Z(\bar{\lambda}) = \min_{\bar{\mu} \in \bar{\Lambda}'} Z(\bar{\mu}), \quad (\text{III.74})$$

wobei im Fall (III.49)-(III.50) $\bar{\Lambda}' := \{\bar{\mu}_0 \in \mathbb{R}^{m_0} \mid C\bar{\mu}_0 \geq 0\}$ ist. Im Fall (III.51)-(III.52) und $l_r = 1$ ist $\bar{\Lambda}' := \{\bar{\mu}_1 \in \mathbb{R}^{m_1} \mid -\bar{s} \leq C\bar{\mu}_1 \leq \bar{s}\}$. Das Problem (III.74) ist demnach ein quadratisches Optimierungsproblem mit linearen Nebenbedingungen. Hierfür können effiziente Verfahren des Quadratic-Programmings eingesetzt werden.

Im Fall (III.51)-(III.52) und $l_r = 2$ sowie im Fall (III.53)-(III.54) setzt man $\bar{\Lambda}' := \{\bar{\mu}_1 \in \mathbb{R}^{m_1} \mid R(\bar{\mu}_1) \leq 0\}$ bzw. $\bar{\Lambda}' := \{(\bar{\mu}_0, \bar{\mu}_1) \in \mathbb{R}^{m_0} \times \mathbb{R}^{m_1} \mid C\bar{\mu}_0 \geq 0, R(\bar{\mu}_1) \leq 0\}$. Für diese Fälle sind die Nebenbedingungen nichtlinear. Zur numerischen Lösung müssen deshalb Verfahren der nichtlinearen Optimierung eingesetzt werden.

Die Tabellen III.6 und III.7 zeigen das Iterationsverhalten des Uzawa-Verfahrens und des Verfahrens, das durch Umformulierung des Sattelpunktproblems in ein Minimierungsproblem gewonnen wird. Der Vergleich beider Verfahren ist angemessen, da der Aufwand eines Uzawa-Schritts in etwa einer Funktionsauswertung des eingesetzten Optimierungsverfahrens entspricht.

⁵³Hier erweist sich eine Tschebyscheff-Einteilung als vorteilhaft. Gegenüber einer Einteilung in äquidistanten Gitterpunkten sind in diesem Fall wesentlich weniger Testpunkte erforderlich. Für Abschätzungen bezüglich Schwankungen von Polynomen in Gitterpunkten des Einheitsintervalls sei auf [48] verwiesen.

		Uzawa				
m	m_0	ZFW	Ref-ZFW	Tol.	ϱ_n	n
16	4	-0.64052	0.09120	1.0E-2	1.0	3
64	8	-0.54819	0.00112	1.0E-5	4.0	20
256	16	-0.54653	0.00277	1.0E-5	8.0	28
1024	32	-0.54870	6.16E-4	1.0E-5	17.0	68
4096	64	-0.54913	1.80E-4	1.0E-6	37.0	91
16384	128	-0.54927	4.40E-5	1.0E-6	70.0	111
65536	256	-0.54930	8.99E-6	1.0E-6	140.0	177
262144	512	-0.54931	5.97E-9	1.0E-7	280.0	480

Tab. III.6: Iterationsverhalten des Uzawa-Verfahrens.

		SQOPT			
m	m_0	ZFW	Ref-ZFW	MOT	Fkt.
16	4	-0.64043	0.09111	1.0E-1	4
64	8	-0.54819	0.00112	1.0E-2	6
256	16	-0.54654	0.00277	1.0E-2	8
1024	32	-0.54870	6.17E-4	1.0E-4	16
4096	64	-0.54913	1.80E-4	1.0E-4	28
16384	128	-0.54927	4.41E-5	1.0E-5	57
65536	256	-0.54930	8.89E-6	1.0E-5	109
262144	512	-0.54931	0	1.0E-6	212

Tab. III.7: Iterationsverhalten von SQOPT.

Das Minimierungsproblem (III.74) wird mit Hilfe des kommerziellen Programmpakets SQOPT numerisch gelöst, es kann jedoch ein beliebiges Optimierungsverfahren, das diesem Problem angepasst ist, an dieser Stelle verwendet werden. Das zugrunde liegende Beispiel entspricht dem vereinfachten Signorini-Problem aus Abschnitt II.2.1.

Neben den bisher betrachteten Größen sind in den Tabellen III.6 und III.7 zusätzlich die Dimension m_0 von $U'_{0,H}$ und der Parameter ϱ_n , der jeweils möglichst optimal gewählt wurde, eingetragen.

Erkennbar ist, dass das Uzawa-Verfahren etwa doppelt so viele Iterationen benötigt wie das Verfahren, das durch Umformulierung gewonnen wird, sofern man Iterationen und Funktionsauswertungen ungefähr als gleichwertig betrachtet. Im Gegensatz zum vorherigen Abschnitt kann für dieses Beispiel das kommerzielle Optimierungspaket SQOPT wirksam eingesetzt werden. Der Grund hierfür ist vor allem die moderate Anzahl an Optimierungsvariablen, die für die feinste hier betrachtete Verfeinerungsstufe bei 512 liegt. Diese Größenordnung kann von kommerziellen Optimierungsverfahren, insbesondere von SQOPT, äußerst effizient bewältigt werden. Eine kleine Anzahl an

Optimierungsvariablen ist insbesondere bei allen Problemen vom Signorini-Typ zu erwarten, da hier der Kontaktbereich auf dem Rand liegt.

Abschließend ist anzumerken, dass für die Einbettung in Optimierungsverfahren im Allgemeinen effiziente Auswertungen der Zielfunktion Z und ihres Gradienten anzugeben sind. Da die Steifigkeitsmatrix A während des Optimierungsprozesses nicht geändert wird, kann eine Faktorisierung einmalig vorgenommen werden, so dass die Auswertung der Zielfunktion und des zugehörigen Gradienten durch ein effizientes Vorwärts-Rückwärtseinsetzen vollzogen werden kann. Auch für das Uzawa-Verfahren reicht eine einmalige Faktorisierung der Matrix A . Alternativ können auch iterative Verfahren, wie etwa vorkonditionierte CG-Verfahren, eingesetzt werden, wobei die Lösung beim jeweils vorhergehenden Iterationsschritt bzw. bei der letzten Funktionsauswertung als Startwert verwendet werden kann.⁵⁴

⁵⁴An dieser Stelle lassen sich vor allem kommerzielle Programmpakete vorteilhaft nutzen. In Bezug auf Faktorisierungen sei etwa auf das Paket `UMFPACK` [40] und für CG-Verfahren z.B. auf `SLAP` [91] verwiesen.

Kapitel IV

Fehlerkontrolle und Adaptivität

Ein wichtiger Bestandteil moderner Diskretisierungsverfahren ist die Kontrolle des Diskretisierungsfehlers $u - u_h$ bzw. $\lambda - \lambda_H$, gemessen in einer Norm oder bezüglich eines anderen, beliebigen Fehlerfunktionals. Unter einer a-posteriori Fehlerkontrolle versteht man im Allgemeinen die Angabe einer oberen Schranke für den Diskretisierungsfehler, die (u.U. bis auf eine generische Konstante) aus der Approximationslösung berechenbar ist und die außerdem die richtige Fehlerasymptotik wiedergibt. Der Nutzen einer derartigen Schranke ist offensichtlich: Unterschreitet die Schranke eine vorgegebene Fehlertoleranz, so gilt dies auch für den Diskretisierungsfehler. Das Verfahren kann in diesem Fall abgebrochen werden.

Ein weiterer Nutzen ergibt sich, wenn die Fehlerkontrolle Aussagen über den lokalen Diskretisierungsfehler ermöglicht. Sind Bereiche im Rechengebiet mit relativ großen Fehleranteilen bekannt, können dort spezielle Adaptierungsmaßnahmen im Diskretisierungsprozess vorgenommen werden, so dass es insgesamt zu einer deutlich größeren Fehlerreduktion als bei globalen Verfeinerungen kommt.

In diesem Kapitel werden zunächst normbezogene Fehlerschranken für elliptische Minimierungsprobleme erster und zweiter Art hergeleitet. Insbesondere kann hierfür der in Abschnitt I.3 formulierte allgemeine funktionalanalytische Rahmen genutzt werden. Anschließend werden darauf basierend Finite-Elemente-Fehlerkontrollen bezüglich der H^1 -Norm für die in den Ab-

schnitten II.2, II.3 und II.5 dargestellten Problemklassen konkretisiert und anhand von numerischen Experimenten geprüft.

Ansätze zur Kontrolle von Diskretisierungsfehlern, die in beliebigen linearen Fehlermaßen bewertet werden sollen, werden abschließend diskutiert.

Der Hauptnutzen der hier betrachteten Fehlerkontrollen besteht vor allem in der Anwendung adaptiver Verfahren zur Konvergenzsteigerung. So werden sowohl h - als auch hp -adaptive Verfeinerungsstrategien vorgestellt und getestet. Insbesondere werden das Konvergenzverhalten und die resultierenden adaptiven FE-Gitter mit den a-priori Ergebnissen aus den Abschnitten III.1.1 und III.1.2 verglichen.

IV.1 Ein allgemeiner Zugang für elliptische Minimierungsprobleme

Das Ziel dieses Abschnitts besteht in der Formulierung eines allgemeinen Zugangs für die a-posteriori Fehlerkontrolle des Diskretisierungsfehlers elliptischer Minimierungsprobleme erster und zweiter Art.

Grundlegend ist hierbei die in Kapitel I eingeführte Beschreibung von Minimierungsproblemen in Form von variationellen Sattelpunktproblemen auf normierten Räumen. Bei der folgenden Darstellung ist mit $\|\cdot\|$ die jeweils zum Raum des eingesetzten Elements gehörende Norm gemeint.

Variationelle Sattelpunktprobleme für elliptische Minimierungsprobleme erster und zweiter Art können in der folgenden Weise zusammengefasst werden: Es seien V und U normierte Räume sowie $\Lambda \subset U'$ und $\beta \in L(V, U)$ surjektiv. Ferner seien $a : V \times V \rightarrow \mathbb{R}$ eine stetige und elliptische Bilinearform, $\ell \in V'$ und $\tilde{g} \in U$. Für $(u, \lambda) \in V \times \Lambda$ gelte:

$$\forall v \in V : a(u, v) = \langle \ell, v \rangle - \langle \lambda, \beta(v) \rangle \quad (\text{IV.1})$$

$$\forall \mu \in \Lambda : \langle \mu - \lambda, \beta(u) - \tilde{g} \rangle \leq 0. \quad (\text{IV.2})$$

Für einen Diskretisierungsansatz, wie in Kapitel III beschrieben, wird der folgende allgemeine Rahmen gewählt: Es seien $V_h \subset V$ ein Unterraum und $\Lambda_H \subset \Lambda$. Für $(u_h, \lambda_H) \in V_h \times \Lambda_H$ gelte:

$$\forall v \in V_h : a(u_h, v_h) = \langle \ell, v_h \rangle - \langle \lambda_H, \beta(v_h) \rangle \quad (\text{IV.3})$$

$$\forall \mu_H \in \Lambda_H : \langle \mu_H - \lambda_H, \beta(u_h) - \tilde{g} \rangle \leq 0. \quad (\text{IV.4})$$

Als Fehlerkontrolle wird in diesem Zusammenhang die Gewinnung einer Zahl $\eta \in \mathbb{R}_{\geq 0}$ bezeichnet, so dass gilt:

$$\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \leq C\eta^2 \quad (\text{IV.5})$$

Wichtig ist, dass die Konstante $C > 0$ weder von V_h oder Λ_H noch von (u, λ) abhängt. In die reelle Zahl η geht nur das Paar (u_h, λ_H) ein, nicht aber (u, λ) . Da in der Regel (u_h, λ_H) eine diskrete (FE-)Lösung ist, ist η eine a-posteriori berechenbare Größe und wird im Folgenden Fehlerschätzer genannt. Für Fehlerschätzer ist die Forderung zu stellen, dass sie bezüglich der Wahl $V_h := V$ und $\Lambda_H := \Lambda$ konsistent sind: Für $u_h = u$ und $\lambda_H = \lambda$ soll $\eta = 0$ sein.¹ Im Gegensatz zu η wird C als generische Konstante in der Regel nicht näher bestimmt.²

Im Folgenden werden Fehlerschätzer für elliptische Minimierungsprobleme erster Art sowie für unrestringierte und restringierte elliptische Minimierungsprobleme zweiter Art hergeleitet.³ Die Kernidee der folgenden Ausführungen besteht darin, das Problem (IV.5) für Fehlerschätzer von Variationsgleichungen zugänglich zu machen. Für die Konkretisierung letzterer kann dann auf etablierte Techniken zur Fehlerkontrolle zurückgegriffen werden. Der wesentliche Vorteil hierbei ist, dass der allgemeine funktionalanalytische Kontext nicht verlassen werden muss.

Entscheidend für die Herleitung der Fehlerschätzer ist die Idee, über das folgende Hilfsproblem eine Verbindung zu etablierten Fehlerschätzern für Variationsgleichungen herzustellen: Für λ_H aus (IV.3)-(IV.4) ist ein $u_* \in V$ gesucht, so dass gilt:

$$\forall v \in V : a(u_*, v) = \langle \ell, v \rangle - \langle \lambda_H, \beta(v) \rangle, \quad (\text{IV.6})$$

Diese Idee geht ursprünglich auf Braess [26] zurück, der einen ähnlichen Ansatz speziell für Hindernisprobleme anwendet. Das Hilfsproblem (IV.6) ist jedoch universell einsetzbar.

¹vgl. entsprechende Bemerkungen in [94]

²Die Einbeziehung einer im Allgemeinen unbekanntem generischen Konstante führt dazu, dass der Fehler quantitativ nicht bestimmt werden kann, was augenscheinlich widersprüchlich zur Bezeichnung Fehlerschätzer ist. Tatsächlich ist man aber häufig nur an einer Schätzung der Fehlerordnung oder der Fehlerverteilung interessiert. Eine wichtige Motivation für Fehlerschätzer mit generischen Konstanten erhält man aus der Anwendung von Gitteradaptierungsstrategien. Hierzu sei auf Abschnitt IV.7 verwiesen.

³Jedoch sind die Überlegungen nicht nur auf Minimierungsprobleme beschränkt, da als Ausgangspunkt eine Sattelpunktformulierung gewählt wird. Insbesondere kann auf die Symmetrie der Bilinearform verzichtet werden.

Lemma IV.1 *Es erfüllen $(u, \lambda) \in V \times \Lambda$ und $u_\star \in V$ die Variationsgleichung (IV.1) bzw. (IV.6). Dann existiert ein $\tilde{C} > 0$, so dass gilt:*

$$\|\lambda - \lambda_H\| \leq \tilde{C} \|u - u_\star\|.$$

Beweis: Nach Satz A.10 existiert ein $c > 0$, so dass

$$\forall w \in U : \exists v \in V : \beta(v) = w, \|v\| \leq c \|w\|$$

ist. Für $\tilde{V} := \{v \in V \mid \|v\| \leq c \|\beta(v)\|\}$ gilt demnach $\beta(\tilde{V}) = U$. Daraus folgt

$$\begin{aligned} \|\lambda - \lambda_H\| &= \sup_{w \in U \setminus \{0\}} \frac{\langle \lambda - \lambda_H, w \rangle}{\|w\|} = \sup_{v \in \tilde{V} \setminus \{0\}} \frac{a(u - u_\star, v)}{\|\beta(v)\|} \\ &\leq c \sup_{v \in \tilde{V} \setminus \{0\}} \frac{a(u - u_\star, v)}{\|v\|} \leq c \nu_0 \|u - u_\star\|. \end{aligned}$$

□

Lemma IV.2 *Es seien (u, λ) und (u_h, λ_H) Lösung von (IV.1)-(IV.2) bzw. (IV.3)-(IV.4) sowie u_\star Lösung von (IV.6). Dann existieren $C' > 0$ und $C'' > 0$, so dass gilt:*

$$\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \leq C' \|u_\star - u_h\|^2 + C'' \langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle.$$

Beweis: Aus der Stetigkeit und der Elliptizität von a sowie aus der Youngschen Ungleichung (III.11) folgt:

$$\begin{aligned} &\nu_1 \|u - u_h\|^2 \\ &\leq a(u - u_h, u - u_h) \\ &= a(u - u_\star, u - u_h) + a(u_\star - u_h, u - u_h) \\ &\leq \langle \lambda_H - \lambda, \beta(u - u_h) \rangle + \nu_0 \|u_\star - u_h\| \|u - u_h\| \\ &= \langle \lambda_H - \lambda, \beta(u) - \tilde{g} \rangle + \langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle + \nu_0 \|u_\star - u_h\| \|u - u_h\| \\ &\leq \langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle + \nu_0 \|u_\star - u_h\| \|u - u_h\| \\ &\leq \langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle + \frac{\nu_0^2}{2\tilde{\epsilon}} \|u_\star - u_h\|^2 + \frac{\tilde{\epsilon}}{2} \|u - u_h\|^2. \end{aligned}$$

Daraus gewinnt man:

$$\|u - u_h\|^2 \leq \frac{\nu_0^2}{\tilde{\epsilon}(2\nu_1 - \tilde{\epsilon})} \|u_\star - u_h\|^2 + \frac{2}{2\nu_1 - \tilde{\epsilon}} \langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle,$$

wobei $0 < \tilde{\epsilon} < 2\nu_1$ gewählt wird. Aus Lemma IV.1 folgt damit schließlich

$$\begin{aligned} & \|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \\ & \leq \|u - u_h\|^2 + \tilde{C}\|u - u_\star\|^2 \\ & \leq \|u - u_h\|^2 + 2\tilde{C}(\|u - u_h\|^2 + \|u_\star - u_h\|^2) \\ & \leq \left(\frac{\nu_0^2(1 + 2\tilde{C})}{\tilde{\epsilon}(2\nu_1 - \tilde{\epsilon})} + 2\tilde{C} \right) \|u_\star - u_h\|^2 + \frac{2(1 + 2\tilde{C})}{2\nu_1 - \tilde{\epsilon}} \langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle. \end{aligned}$$

□

Wesentlich ist, dass die Lösung u_h von (IV.3)-(IV.4) auch gleichzeitig eine diskrete Lösung des Hilfsproblems (IV.6) ist. Kennt man einen Fehlerschätzer η_\star für $\|u_\star - u_h\|$, also $\|u_\star - u_h\|^2 \leq C_\star \eta_\star^2$ mit $C_\star > 0$, dann folgt aus Lemma IV.2 unmittelbar

$$\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \leq C' C_\star \eta_\star^2 + C'' \langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle.$$

Da das Hilfsproblem (IV.6) eine Variationsgleichung ist, können aus der Literatur bekannte Fehlerschätzer für Variationsgleichungen verwendet werden, so dass nur noch der Ausdruck $\langle \lambda - \lambda_H, \beta(u_h) - \tilde{g} \rangle$ abzuschätzen ist.

IV.1.1 Elliptische Minimierungsprobleme erster Art

Betrachtet wird ein elliptische Minimierungsproblem erster Art mit K wie in (I.24) und einem surjektiven $\beta_0 \in L(V, U_0)$.

Das Paar $(u, \lambda_0) \in V \times \Lambda_0$ erfülle das System (I.26)-(I.27) und das Paar $(u_h, \lambda_{0,H}) \in V_h \times \Lambda_{0,H}$ das System (III.4)-(III.5). Nach Satz I.12 ist $u \in V$ dann die eindeutig bestimmte Lösung des elliptischen Minimierungsproblems erster Art und u_h eine Approximationslösung.

Für die Herleitung eines Fehlerschätzers erhält man aus Lemma IV.2

$$\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \leq C' C_0 \eta_0^2 + C'' \langle \lambda_0 - \lambda_{0,H}, \beta_0(u_h) - g \rangle. \quad (\text{IV.7})$$

Dabei sei $\eta_0 \in \mathbb{R}_{\geq 0}$ und $C_0 > 0$, so dass

$$\|u_0 - u_h\|^2 \leq C_0 \eta_0^2 \quad (\text{IV.8})$$

gilt, wobei $u_0 \in V$ Lösung des folgenden Hilfsproblems ist:

$$\forall v \in V : a(u_0, v) = \langle \ell, v \rangle - \langle \lambda_{0,H}, \beta_0(v) \rangle. \quad (\text{IV.9})$$

Da Einsetzen von 0 und $2\lambda_{0,H}$ in (III.5) $\langle \lambda_{0,H}, \beta_0(u_h) - g \rangle = 0$ liefert, folgt aus (IV.7)

$$\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \leq C' C_0 \eta_0^2 + C'' \langle \lambda_0, \beta_0(u_h) - g \rangle.$$

Im Allgemeinen ist $\langle \lambda_0, \beta_0(u_h) - g \rangle = -\langle \lambda_0, g - \beta_0(u_h) \rangle \not\leq 0$, da $g - \beta_0(u_h)$ nicht im Kegel G sein muss.

Der folgende Ansatz beruht auf dem Hinzufügen einer Korrektur d , so dass $g - \beta(u_h + d) \in G$ und damit $\langle \lambda_0, \beta_0(u_h + d) - g \rangle \leq 0$ ist. Die Menge aller Korrekturen mit dieser Eigenschaft sei

$$\tilde{K} := \{v \in V \mid g - \beta_0(u_h + v) \in G\}.$$

Für $d \in \tilde{K}$ folgt dann aus (I.26) und (III.4) sowie der Youngschen Ungleichung (III.11)

$$\begin{aligned} \langle \lambda_0, \beta_0(u_h) - g \rangle &= -\langle \lambda_0, g - \beta_0(u_h + d) \rangle - \langle \lambda_0, \beta_0(d) \rangle \\ &\leq a(u, d) - \langle \ell, d \rangle \\ &= a(u - u_h, d) + a(u_h, d) - \langle \ell, d \rangle \\ &\leq \nu_0 \|u - u_h\| \|d\| + a(u_h, d) - \langle \ell, d \rangle \\ &\leq \frac{\epsilon}{2} \|u - u_h\|^2 + \frac{\nu_0^2}{2\epsilon} \|d\|^2 + a(u_h, d) - \langle \ell, d \rangle. \end{aligned}$$

Der Ausdruck $a(u_h, d) - \langle \ell, d \rangle$ kann unter Ausnutzung von (IV.9) weiter betrachtet werden:

$$\begin{aligned} a(u_h, d) - \langle \ell, d \rangle &= a(u_h - u_0, d) - \langle \lambda_{0,H}, \beta_0(d) \rangle \\ &\leq \nu_0 \|u_0 - u_h\| \|d\| - \langle \lambda_{0,H}, \beta_0(d) \rangle \\ &\leq \frac{C_0}{2} \eta_0^2 + \frac{\nu_0^2}{2} \|d\|^2 + |\langle \lambda_{0,H}, \beta_0(d) \rangle|. \end{aligned}$$

Aus (IV.7) folgt damit

$$\begin{aligned} \left(1 - \frac{C''\epsilon}{2}\right) (\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2) &\leq \\ \max\left\{\left(C' + \frac{C''}{2}\right)C_0, \frac{C''(1+\epsilon)\nu_0^2}{2\epsilon}, C''\right\} (\eta_0^2 + \|d\|^2 + |\langle \lambda_{0,H}, \beta_0(d) \rangle|), & \end{aligned}$$

wobei $0 < \epsilon < 2/C''$ zu wählen ist. Insgesamt erhält man schließlich:

Satz IV.1 *Es seien $(u, \lambda_0) \in V \times \Lambda_0$ und $(u_h, \lambda_{0,H}) \in V_h \times \Lambda_{0,H}$ Lösungen von (I.26)-(I.27) bzw. (III.4)-(III.5) sowie u_0 Lösung von (IV.9). Außerdem gelte (IV.8) für $\eta_0 \in \mathbb{R}_{\geq 0}$. Dann existiert ein $C > 0$, so dass mit $d \in \tilde{K}$ gilt:*

$$\|u - u_h\|^2 + \|\lambda_0 - \lambda_{0,H}\|^2 \leq C(\eta_0^2 + \|d\|^2 + |\langle \lambda_{0,H}, \beta_0(d) \rangle|).$$

Nicht jede Wahl von $d \in \tilde{K}$ führt zu einem konsistenten Fehlerschätzer. Wenn $u_h = u = u_*$, $\lambda_H = \lambda$ und $d := w - u_h$ mit $\beta(w) = g$ gesetzt wird, ist zwar $\eta_0^2 = 0$ und $\langle \lambda_0, d \rangle = 0$,⁴ aber $\|d\|$ ist für $w \neq u$ von null verschieden. Eine geeignete Wahl von $d \in \tilde{K}$ für einen konsistenten Fehlerschätzer besteht offenbar in der Bestimmung von $\min_{\tilde{d} \in \tilde{K}} (\|\tilde{d}\|^2 + |\langle \lambda_{0,H}, \beta_0(\tilde{d}) \rangle|)$. In diesem Fall muss jedoch ein restringiertes elliptisches Minimierungsproblem zweiter Art gelöst werden, was im Allgemeinen nicht praktikabel ist.

Eine weitere Möglichkeit erhält man folgendermaßen: Da G als abgeschlossen und konvex vorausgesetzt wird, existiert ein $r \in G$ als Projektion von $g - \beta_0(u_h)$, d.h. $\|g - \beta_0(u_h) - r\| = \min_{\tilde{r} \in G} \|g - \beta_0(u_h) - \tilde{r}\|$. Aus der Surjektivität von β_0 folgt, dass ein $d \in V$ mit $\beta_0(d) = g - \beta_0(u_h) - r$ existiert. Weil $g - \beta_0(u_h) - \beta_0(d) = r \in G$ ist, gilt $d \in \tilde{K}$. Diese Wahl führt offensichtlich zu einem konsistenten Fehlerschätzer.

Bei Signorini-Problemen ist der Operator β_0 im Allgemeinen ein (nicht injektiver) Spuoperator, wodurch d als eine geeignete Fortsetzung von $g - \beta_0(u_h) - r$ auf V zu wählen ist. Da $\|d\|$ möglichst klein sein soll, erscheint etwa im Fall $V = H^1(\Omega, \Gamma_0)$ die harmonische Fortsetzung als geeignet.

Sowohl bei Hindernisproblemen als auch bei Signorini-Problemen kann darüber hinaus auch der positive (oder negative) Anteil $(g - \beta_0(u_h))_+$ (oder $(g - \beta_0(u_h))_-$) bzw. die harmonische Fortsetzung hiervon für d gewählt werden. Auch hier ist die Konsistenz des Fehlerschätzers direkt einsehbar. Hierzu sei insbesondere auf Abschnitt IV.3.1 verwiesen.

IV.1.2 Unrestringierte elliptische Minimierungsprobleme zweiter Art

Für die Herleitung von Fehlerschätzern für elliptische Minimierungsprobleme zweiter Art mit $K = V$ kann in ähnlicher Weise wie oben vorgegangen werden, sofern das hierzu verwendete Funktional $j : V \rightarrow \mathbb{R}$ die Darstellung wie in (I.59) mit einer konvexen und abgeschlossenen Menge Λ_1 besitzt.⁵

⁴vgl. Abschnitt I.4, (I.45) und (I.48)

⁵Diese Darstellung ist typisch für Reibungsprobleme oder für Bingham-Fluid-Probleme (vgl. Abschnitt II.3 bzw. Abschnitt II.5), jedoch ist dies nicht der allgemeine Fall.

Hierzu erfülle $(u, \lambda_1) \in V \times \Lambda_1$ das System (I.33)-(I.34) und das Paar $(u_h, \lambda_{1,H}) \in V_h \times \Lambda_{1,H}$ das System (III.6)-(III.7). Nach Satz I.14 ist $u \in V$ dann die eindeutig bestimmte Lösung des elliptischen Minimierungsproblems zweiter Art und u_h eine zugehörige Approximationslösung.

Lemma IV.2 liefert in diesem Fall

$$\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \leq C' \eta_1^2 + C'' \langle \lambda_1 - \lambda_{1,H}, \beta_1(u_h) \rangle. \quad (\text{IV.10})$$

Hierbei ist $\eta_1 \in \mathbb{R}_{\geq 0}$, so dass

$$\|u_1 - u_h\|^2 \leq C_1 \eta_1^2 \quad (\text{IV.11})$$

gilt, wobei $u_1 \in V$ Lösung des Hilfsproblems

$$\forall v \in V : a(u_1, v) = \langle \ell, v \rangle - \langle \lambda_{1,H}, \beta_1(v) \rangle \quad (\text{IV.12})$$

ist. Aus (I.59) und (IV.10) folgt damit unmittelbar

$$\begin{aligned} \|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 &\leq C' C_1 \eta_1^2 + C'' (j(u_h) - \langle \lambda_{1,H}, \beta_1(u_h) \rangle) \\ &\leq \max\{C' C_1, C''\} (\eta_1^2 + |j(u_h) - \langle \lambda_{1,H}, \beta_1(u_h) \rangle|) \end{aligned}$$

und damit:

Satz IV.2 *Es seien $(u, \lambda_1) \in V \times \Lambda_1$ und $(u_h, \lambda_{1,H}) \in V_h \times \Lambda_{1,H}$ Lösungen von (I.33)-(I.34) bzw. (III.6)-(III.7), sowie u_1 Lösung von (IV.12). Außerdem gelte (IV.11) für $\eta_1 \in \mathbb{R}_{\geq 0}$. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|^2 + \|\lambda_1 - \lambda_{1,H}\|^2 \leq C (\eta_1^2 + |j(u_h) - \langle \lambda_{1,H}, \beta_1(u_h) \rangle|).$$

Dass der in Satz IV.2 aufgeführte Fehlerschätzer zumindest in der Konfiguration von Abschnitt I.4 mit $u \in \text{dom}(A)$ konsistent ist, folgt unmittelbar aus (I.56) und (I.60).

IV.1.3 Restringierte elliptische Minimierungsprobleme zweiter Art

Zur Untersuchung eines restringierten elliptischen Minimierungsproblems zweiter Art müssen die bisherigen Überlegungen kombiniert werden. Hierzu betrachtet man die Tripel $(u, \lambda_0, \lambda_1) \in V \times \Lambda_0 \times \Lambda_1$ und $(u, \lambda_{0,H}, \lambda_{1,H}) \in V_h \times \Lambda_{0,H} \times \Lambda_{1,H}$, die die Systeme (I.35)-(I.36) bzw. (III.8)-(III.9) erfüllen sollen. Nach Satz I.15 ist u Lösung des elliptischen Minimierungsproblems

zweiter Art mit $K \subset V$ wie in (I.24) und j wie in (I.59) und u_h eine entsprechende approximative Lösung.

Aus Lemma IV.2 folgt unmittelbar

$$\|u - u_h\|^2 \leq C' C_{0,1} \eta_{0,1}^2 + C'' (\langle \lambda_0 - \lambda_{0,H}, \beta_0(u_h) - g \rangle + \langle \lambda_1 - \lambda_{1,H}, \beta_1(u_h) \rangle). \quad (\text{IV.13})$$

Hierbei ist $\eta_{0,1} \in \mathbb{R}_{\geq 0}$, so dass

$$\|u_{0,1} - u_h\|^2 \leq C_{0,1} \eta_{0,1}^2 \quad (\text{IV.14})$$

ist, wobei $u_{0,1} \in V$ das Hilfsproblem

$$\forall v \in V : a(u_{0,1}, v) = \langle \ell, v \rangle - \langle \lambda_{0,H}, \beta_0(v) \rangle - \langle \lambda_{1,H}, \beta_1(v) \rangle \quad (\text{IV.15})$$

löst. Durch Einsetzen von $(0, \lambda_{1,H})$ und $(2\lambda_{0,H}, \lambda_{1,H})$ in (III.9) erhält man $\langle \lambda_{0,H}, \beta_0(u_h) - g \rangle = 0$. Aus (I.59) und (IV.13) gewinnt man damit:

$$\|u - u_h\|^2 \leq C' C_{0,1} \eta_{0,1}^2 + C' (\langle \lambda_0, \beta_0(u_h) - g \rangle + j(u_h) - \langle \lambda_{1,H}, \beta_1(u_h) \rangle). \quad (\text{IV.16})$$

Aus (I.35) und (IV.15) sowie der Youngschen Ungleichung (III.11) folgt für $d \in \tilde{K}$

$$\begin{aligned} \langle \lambda_0, \beta_0(u_h) - g \rangle &= -\langle \lambda_0, g - \beta_0(u_h) - \beta(d) \rangle - \langle \lambda_0, \beta_0(d) \rangle \\ &\leq a(u, d) - \langle \ell, d \rangle + \langle \lambda_1, \beta_1(d) \rangle \\ &\leq a(u - u_h, d) + a(u_h, d) - \langle \ell, d \rangle + j(d) \\ &\leq \nu_0 \|u - u_h\| \|d\| + a(u_h, d) - \langle \ell, d \rangle + j(d) \\ &\leq \frac{\epsilon}{2} \|u - u_h\|^2 + \frac{\nu_0^2}{2\epsilon} \|d\|^2 + a(u_h, d) - \langle \ell, d \rangle + j(d). \end{aligned}$$

Mit Hilfe von (IV.15) kann $a(u_h, d) - \langle \ell, d \rangle + j(d)$ weiter untersucht werden:

$$\begin{aligned} &a(u_h, d) - \langle \ell, d \rangle + j(d) \\ &= a(u_h - u_{0,1}, d) - \langle \lambda_{0,H}, \beta_0(d) \rangle + j(d) - \langle \lambda_{1,H}, \beta_1(d) \rangle \\ &\leq \nu_0 \|u_{0,1} - u_h\| \|d\| - \langle \lambda_{0,H}, \beta_0(d) \rangle + j(d) - \langle \lambda_{1,H}, \beta_1(d) \rangle \\ &\leq \frac{1}{2} \|u_{0,1} - u_h\|^2 + \frac{\nu_0^2}{2} \|d\|^2 + |\langle \lambda_{0,H}, \beta_0(d) \rangle| + |j(d) - \langle \lambda_{1,H}, \beta_1(d) \rangle|. \end{aligned}$$

Insgesamt folgt aus (IV.13)

$$\begin{aligned} (1 - \frac{C''\epsilon}{2})(\|u - u_h\|^2 + \|\lambda - \lambda_H\|^2) \leq \\ \max\{(C' + \frac{C''}{2})C_0, \frac{C''(1+\epsilon)\nu_0^2}{2\epsilon}, C''\}(\eta_0^2 + \|d\|^2 + \\ |\langle \lambda_{0,H}, \beta_0(d) \rangle| + |j(d) - \langle \lambda_{1,H}, \beta_1(d) \rangle|), \end{aligned}$$

wobei $0 < \epsilon < 2/C''$ zu wählen ist. Damit gilt schließlich:

Satz IV.3 *Es seien $(u, \lambda_0, \lambda_1) \in V \times \Lambda_0 \times \Lambda_1$ und $(u_h, \lambda_{0,H}, \lambda_{1,H}) \in V_h \times \Lambda_{0,H} \times \Lambda_{1,H}$ Lösungen von (I.35)-(I.36) bzw. (III.8)-(III.9) sowie $u_{0,1}$ Lösung von (IV.15). Außerdem gelte (IV.14) für $\eta_{0,1} \in \mathbb{R}_{\geq 0}$. Dann existiert ein $C > 0$, so dass mit $d \in \tilde{K}$ gilt:*

$$\begin{aligned} \|u - u_h\|^2 + \|\lambda - \lambda_H\|^2 \leq C(\eta_{0,1}^2 + \|d\|^2 + |\langle \lambda_{0,H}, \beta_0(d) \rangle| \\ + |j(d) - \langle \lambda_{1,H}, \beta_1(d) \rangle| + |j(u_h) - \langle \lambda_{1,H}, \beta_1(u_h) \rangle|). \end{aligned}$$

Auch hier ist die Konsistenz des in Satz IV.3 dargestellten Fehlerschätzers für die in Abschnitt I.4 beschriebenen Voraussetzungen sofort einsichtig, wenn hierzu (I.63) und (I.64) sowie (I.68) und (I.69) berücksichtigt werden.

IV.2 Residuale Fehlerschätzer für Variationsgleichungen

Für die Fehlerschätzer aus den Sätzen IV.1, IV.2 und IV.3 sind Fehlerschätzer für Variationsgleichungen erforderlich, die durch die Hilfsprobleme (IV.9), (IV.12) und (IV.15) gegeben sind. Für Finite-Elemente-Diskretisierungen der in Abschnitt II.2 und II.3 vorgestellten Problemklassen existiert in der Literatur eine Fülle von Ansätzen, die prinzipiell alle eingesetzt werden können. Für einen Überblick bezüglich Fehlerschätzer sei etwa auf [2] und [105] verwiesen.

Insbesondere sind sogenannte residuale Fehlerschätzer sehr etabliert. Gegenüber anderen Fehlerschätzern wie etwa dem in der Ingenieurliteratur weitverbreiteten ZZ-Fehlerschätzer sind sie auch für Finite-Elemente-Ansätze höherer Ordnung zugänglich.

In diesem Abschnitt werden kurz die wesentlichen Herleitungsschritte zur

Gewinnung residueller Fehlerschätzer beschrieben. Es werden residuale Fehlerschätzer sowohl für Modellprobleme als auch für linear-elastische Probleme vorgestellt, mit denen Fehlerschätzer für die vorgestellten vereinfachten Kontaktprobleme und für Kontaktprobleme des linear-elastischen Falls hergeleitet werden können. Hierbei werden speziell Finite-Elemente-Ansätze höherer Ordnung berücksichtigt.

IV.2.1 Residuale Fehlerschätzer für Modellprobleme

Im Folgenden werden die Bezeichnungen aus Abschnitt II.2 und Abschnitt III.1.1 vorausgesetzt. Insbesondere sei $u \in H^1(\Omega, \Gamma_0)$ Lösung von (II.6) und $u_h \in S^p(\mathcal{T}_h)$ Lösung von (III.13). Als Galerkin-Orthogonalität (III.10) erhält man in diesem Fall

$$\forall v_h \in S^p(\mathcal{T}_h) : (\nabla(u - u_h), \nabla v_h)_0 = 0. \quad (\text{IV.17})$$

Für die Herleitung residueller Fehlerschätzer ist die Verwendung eines geeigneten Interpolationsoperators $I : H^1(\Omega, \Gamma_0) \rightarrow S^p(\mathcal{T}_h)$ von großer Bedeutung. Seine Aufgabe besteht darin, die L_2 -Norm des Diskretisierungsfehlers $e := u - u_h$ durch Potenzen der lokalen Maschenweite h_T und des lokalen Polynomgrads p_T zu erfassen.

Es gilt unter Ausnutzung der Galerkin-Orthogonalität:

$$\begin{aligned} (\nabla e, \nabla e)_0 &= (\nabla e, \nabla(e - I(e)))_0 = (\nabla u, \nabla(e - I(e)))_0 - (\nabla u_h, \nabla(e - I(e)))_0 \\ &= (f, e - I(e))_0 + (q, \gamma(e - I(e)))_{0, \Gamma_1} - (\nabla u_h, \nabla(e - I(e)))_0. \end{aligned} \quad (\text{IV.18})$$

Da für $T \in \mathcal{T}_h$ die Einschränkung $\nabla u_h|_T \in H^1(T)^2$ ist, folgt mit Hilfe der Greenschen Formel (II.2) und der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned} (\nabla e, \nabla e)_0 &= \sum_{T \in \mathcal{T}_h} (f, e - I(e))_{0, T} + (q, \gamma(e - I(e)))_{0, \Gamma_1} \\ &\quad - \sum_{T \in \mathcal{T}_h} (\nabla u_h, \nabla(e - I(e)))_{0, T} \\ &= \sum_{T \in \mathcal{T}_h} (f, e - I(e))_{0, T} + (q, \gamma(e - I(e)))_{0, \Gamma_1} \\ &\quad + \sum_{T \in \mathcal{T}_h} \left((\Delta u_h, e - I(e))_{0, T} - \sum_{E \in \mathcal{E}_T} (\partial_n u_h, \gamma(e - I(e)))_{0, E} \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{T \in \mathcal{T}_h} (f + \Delta u_h, e - I(e))_{0,T} + \sum_{E \in \mathcal{E}^\circ} ([\partial_n u_h], \gamma(e - I(e)))_{0,E} \\
&\quad + \sum_{E \in \mathcal{E}_1} (q - \partial_n u_h, \gamma(e - I(e)))_{0,E} \\
&\leq \sum_{T \in \mathcal{T}_h} \left(R_T \|e - I(e)\|_{0,T} + \sum_{E \in \mathcal{E}_T} R_E \|e - I(e)\|_{0,E} \right), \quad (\text{IV.19})
\end{aligned}$$

wobei \mathcal{E}° die Menge aller inneren Kanten von \mathcal{T}_h , \mathcal{E}_T die Menge aller Kanten von T und \mathcal{E}_1 die Menge aller Kanten auf Γ_1 sowie $[\cdot]$ den Sprung⁶ auf einer Kante bei vorher festgelegtem Normalenvektor bezeichnet. Die element- bzw. kantenbezogenen Residuen R_T bzw. R_E sind definiert als

$$R_T := \|f + \Delta u_h\|_{0,T}, \quad R_E := \begin{cases} \frac{1}{2} \|[\partial_n u_h]\|_{0,E}, & E \in \mathcal{E}^\circ \\ \|q - \partial_n u_h\|_{0,E}, & E \in \mathcal{E}_1. \end{cases} \quad (\text{IV.20})$$

Sind die Polynomgrade $p \equiv 1$ und \mathcal{T}_h quasi-uniform, so wird üblicherweise die sogenannte Clément-Interpolierende⁷ I als Interpolationsoperator herangezogen, um die Ausdrücke $\|v - I(v)\|_{0,T}$ und $\|\gamma(v - I(v))\|_{0,E}$ durch die lokale Maschenweite auszudrücken: Es existiert eine Konstanten $C_I > 0$, so dass $\|v - I(v)\|_{0,T} \leq C_I h_T \|v\|_{1,\omega_T^1}$ und $\|v - I(v)\|_{0,E} \leq C_I h_T^{1/2} \|v\|_{1,\omega_T^1}$ mit $T \in \mathcal{T}_h$ und $E \in \mathcal{E}_T$ gilt. Hierbei ist $\omega_V := \bigcup_{T \in \mathcal{T}_h, V \in \overline{T}} \overline{T}$ und $\omega_T^1 := \bigcup_{V \in \overline{T}} \omega_V$. Da ein Element T eines nicht-degenerierten Gitters \mathcal{T}_h nur in endlich vielen $\omega_{\tilde{T}}^1$ mit $\tilde{T} \in \mathcal{T}_h$ enthalten ist und deren Anzahl unabhängig von der Maschenweite beschränkt ist, folgt aus (IV.19), dass es ein $\tilde{C} > 0$ gibt, so dass gilt:

$$\begin{aligned}
(\nabla e, \nabla e)_0 &\leq C_I \sum_{T \in \mathcal{T}_h} \left(h_T R_T \|e\|_{1,\omega_T^1} + \sum_{E \in \mathcal{E}_T} h_T^{1/2} R_E \|e\|_{1,\omega_T^1} \right) \\
&\leq \tilde{C} \|e\|_1 \left(\sum_{T \in \mathcal{T}_h} \left(h_T^2 R_T^2 + \sum_{E \in \mathcal{E}_T} h_T R_E^2 \right) \right)^{1/2}
\end{aligned}$$

Durch Ausnutzen der Poincaré-Friederichsschen Ungleichung und Dividie-

⁶Für die Kante E und $x \in E$ ist der Sprung einer Funktion v definiert als $[v](x) := \lim_{t \rightarrow 0^+} v(x + tn_E) - \lim_{t \rightarrow 0^+} v(x - tn_E)$.
⁷vgl. [36]

ren erhält man schließlich mit einer Konstanten $C > 0$

$$\|u - u_h\|_1^2 \leq C \sum_{T \in \mathcal{T}_h} \left(h_T^2 R_T^2 + \sum_{E \in \mathcal{E}_T} h_T R_E^2 \right). \quad (\text{IV.21})$$

Sollen die Polynomgrade p_T beliebig, also insbesondere nicht konstant sein, ist eine Verallgemeinerung der Clément-Interpolierenden erforderlich, um auch die lokalen Polynomgrade p_T in die Fehlerschätzung einfließen zu lassen. Für eine Kante $E \in \mathcal{E}_h$ sei hierbei $\omega_E^1 := \bigcup_{V \in \overline{E}} \omega_V$ und $p_E := \min\{p_T \mid E \in \mathcal{E}_T, T \in \mathcal{T}_h\}$, sowie $\omega_T^{j+1} := \bigcup_{V \in \overline{\omega_T^j}} \omega_V$ und $\omega_E^{j+1} := \bigcup_{V \in \overline{\omega_E^j}} \omega_V$.

Definition IV.1 *Ein linearer Operator $I : H^1(\Omega, \Gamma_0) \rightarrow S^p(\mathcal{T}_h)$ heißt hp -Interpolierende vom Clément-Typ, wenn eine Konstante $C_I > 0$ und ein $j \geq 1$ existieren, so dass für alle $T \in \mathcal{T}_h$ und $E \in \mathcal{E}_V$ sowie $v \in H^1(\Omega, \Gamma_0)$ gilt:*

$$\begin{aligned} \|v - I(v)\|_{0,T} &\leq C_I (h_T/p_T) \|\nabla v\|_{0,\omega_T^j} \\ \|\gamma(v - I(v))\|_{0,E} &\leq C_I (h_E/p_E)^{1/2} \|\nabla v\|_{0,\omega_E^j}. \end{aligned}$$

Im Fall $\Omega \subset \mathbb{R}^2$ kann die Frage nach der Existenz einer hp -Interpolierenden vom Clément-Typ positiv beantwortet werden:

Satz IV.4 *Es sei $\Omega \subset \mathbb{R}^2$ und \mathcal{T}_h regulär. Dann existiert eine hp -Interpolierende vom Clément-Typ.*

Beweis: Th.3.1 in [80]. □

Für den Fall $\Omega \subset \mathbb{R}^3$ sind dem Autor keine Ergebnisse bekannt, numerische Untersuchungen lassen jedoch auf die Existenz einer hp -Interpolierenden vom Clément-Typ auch für $\Omega \subset \mathbb{R}^3$ schließen.

Der wesentliche Unterschied zur üblichen Clément-Interpolierenden besteht darin, dass nun, wie gewünscht, auch die Polynomgrade p_T in die lokale Fehlerschätzung eingehen. Mit anderen Worten: Die Konstante C ist unabhängig von der lokalen Maschenweite und dem lokalen Polynomgrad, ein Ansteigen des Polynomgrads zur Genauigkeitsverbesserung wird berücksichtigt. Außerdem erstreckt sich die lokale L_2 -Norm auf das u.U. größere Gebiet ω_T^j bzw. ω_E^j . Um dann in ähnlicher Weise wie oben Gebrauch von der Abschätzung (IV.19) machen zu können, ist es erforderlich, dass benachbarte Elemente von $T \in \mathcal{T}_h$ vergleichbare Durchmesser besitzen und dass ihnen

vergleichbare Polynomgrade zugeordnet sind. Es existieren also $\varrho_0 > 0$ und $\varrho_1 > 0$, so dass für alle $T, T' \in \mathcal{T}_h$ mit $\overline{T} \cap \overline{T'} \neq \emptyset$ gilt:

$$\varrho_0^{-1} h_T \leq h_{T'} \leq \varrho_0 h_T, \quad \varrho_1^{-1} p_T \leq p_{T'} \leq \varrho_1 p_T. \quad (\text{IV.22})$$

Aus (IV.19) folgt damit unmittelbar die im Vergleich zu (IV.21) allgemeinere Fehlerabschätzung⁸:

Satz IV.5 *Es seien $u \in H^1(\Omega, \Gamma_0)^3$ und $u_h \in S^p(\mathcal{T}_h)$ Lösung von (II.6) bzw. (III.13) sowie*

$$\eta^2 := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_T^2 + \sum_{E \in \mathcal{E}_T} (h_E/p_E) R_E^2 \right).$$

Ferner existiere eine hp -Clément-Interpolierende, und (IV.22) sei erfüllt mit $\varrho_0 > 0$ und $\varrho_1 > 0$. Dann existiert eine von \mathcal{T}_h und p unabhängige Konstante $C > 0$, so dass gilt:

$$\|u - u_h\|_1^2 \leq C \eta^2.$$

Beweis: Wie oben folgt aus (IV.19) mit $\tilde{C} > 0$:

$$\begin{aligned} & (\nabla e, \nabla e)_0 \\ & \leq C_I \sum_{T \in \mathcal{T}_h} \left((h_T/p_T) R_T \|\nabla e\|_{0, \omega_T^j} + \sum_{E \in \mathcal{E}_T} (h_E/p_E)^{1/2} R_E \|\nabla e\|_{0, \omega_E^j} \right) \\ & \leq \tilde{C} \|\nabla e\|_0 \left(\sum_{T \in \mathcal{T}_h} (h_T/p_T)^2 R_T^2 + \sum_{E \in \mathcal{E}_T} (h_E/p_E) R_E^2 \right)^{1/2}. \end{aligned}$$

Die Poincaré-Friederichsschen Ungleichung liefert dann die Behauptung. \square

Dass die Konstante C tatsächlich weder von der lokalen Maschenweite h noch von dem lokalen Polynomgrad abhängt, kann mit Hilfe des Effektivitätsindex veranschaulicht werden. Der Effektivitätsindex ist definiert als $\|u - u_h\|_1 / \eta$ (oder auch $\|\nabla e\|_0 / \eta$) und bildet damit eine untere Schranke für \sqrt{C} . Insbesondere kann bei einem annähernd konstanten Effektivitätsindex davon ausgegangen werden, dass die globalen wie auch die lokalen Fehlerordnungen richtig wiedergegeben werden.

⁸vgl. Prop.4.1 in [80]

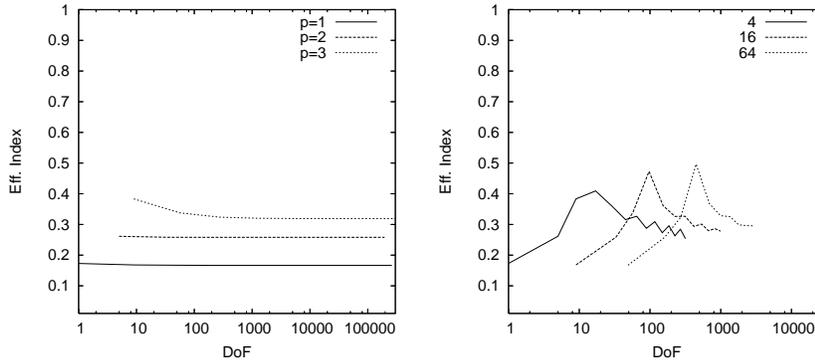


Abb. IV.1: Effektivitätsindizes für h - und p -Verfeinerungen des Einheitsquadrats.

In Abbildung IV.1 sind die Effektivitätsindizes des in Satz IV.5 aufgeführten Fehlerschätzers η für globale h - und p -Verfeinerungen des Einheitsquadrats dargestellt. Die Daten sind so gewählt, dass $u := \cos(0.5\pi x) \cos(0.5\pi y)$ die analytische Lösung ist. Wird die Maschenweite bei konstantem Polynomgrad verkleinert (h -Methode), sind die zugehörigen Effektivitätsindizes konstant. Wird der Polynomgrad bei fester Maschenweite erhöht (p -Methode), schwankt der Effektivitätsindex zwischen 0.15 und 0.45. Die in Abbildung IV.1 (rechts) zugrunde liegenden Anfangsgitter bestehen aus 4, 16 und 64 Zellen. Insgesamt erkennt man, dass die Effektivitätsindizes bei der h - und p -Methode nur relativ kleinen Schwankungen unterworfen sind. Damit wird bestätigt, dass die Konstante C im wesentlichen unabhängig von der lokalen Maschenweite und dem lokalen Polynomgrad ist.

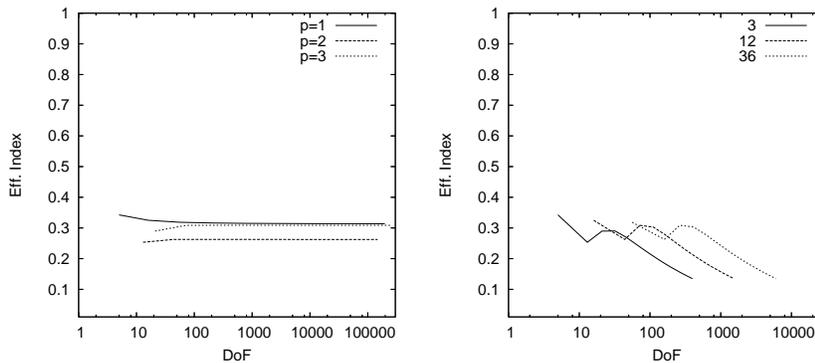


Abb. IV.2: Effektivitätsindizes für h - und p -Verfeinerungen des L-Gebiets.

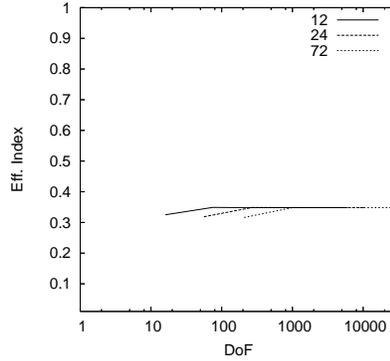


Abb. IV.3: Effektivitätsindizes für hp -Verfeinerungen des L-Gebiets.

In den Abbildungen IV.2 und IV.3 sind die Effektivitätsindizes für h -, p - und hp -Verfeinerungen des L-Gebiets dargestellt. Die analytische Lösung ist hier wie in (III.16) gewählt, also eine Funktion, die nicht H^2 -regulär ist. Für h - und hp -Verfeinerungen sind die Effektivitätsindizes konstant, bei p -Verfeinerungen monoton fallend. Der Abbildung IV.2 (rechts) ist nicht zu entnehmen, ob der Effektivitätsindex konvergiert.

IV.2.2 Residuale Fehlerschätzer für linear-elastische Probleme

In analoger Weise zum vorherigen Abschnitt lassen sich residuale Fehlerschätzer auch für Variationsgleichungen angeben, die linear-elastischen Problemen zugeordnet sind.

Hierzu sei $u \in H^1(\Omega, \Gamma_0)^3$ und $u_h \in S^p(\mathcal{T}_h)^3$ Lösung von (II.39) bzw. (III.30). Die zugehörige Galerkin-Orthogonalität (III.10) ist in diesem Fall

$$\forall v_h \in S^p(\mathcal{T}_h)^3 : (\sigma(u - u_h), \varepsilon(v_h))_0 = 0.$$

Wie in Abschnitt IV.2.1 gewinnt man daraus mit Hilfe eines geeigneten Interpolationsoperators $I : H^1(\Omega, \Gamma_0)^3 \rightarrow S^p(\mathcal{T}_h)^3$ für $e := u - u_h$

$$\begin{aligned} (\sigma(e), \varepsilon(e))_0 &= (\sigma(e), \varepsilon(e - I(e)))_0 \\ &= (\sigma(u), \varepsilon(e - I(e)))_0 - (\sigma(u_h), \varepsilon(e - I(e)))_0 \\ &= (f, e - I(e))_0 + (q, \gamma(e - I(e)))_{0, \Gamma_1} - (\sigma(u_h), \varepsilon(e - I(e)))_0. \end{aligned}$$

Da für $T \in \mathcal{T}_h$ die Einschränkung $\sigma(u_h|_T) \in H^1(T)^{3 \times 3}$ ist, erhält man mit Hilfe der Greenschen Formel (II.3)

$$\begin{aligned}
 & (\sigma(e), \varepsilon(e))_0 \\
 &= \sum_{T \in \mathcal{T}_h} (f, e - I(e))_{0,T} + (q, \gamma(e - I(e)))_{0,\Gamma_1} - \sum_{T \in \mathcal{T}_h} (\sigma(u_h), \varepsilon(e - I(e)))_{0,T} \\
 &= \sum_{T \in \mathcal{T}_h} (f, e - I(e))_{0,T} + (q, \gamma(e - I(e)))_{0,\Gamma_1} \\
 &\quad + \sum_{T \in \mathcal{T}_h} \left((\operatorname{div} \sigma(u_h), e - I(e))_{0,T} - \sum_{F \in \mathcal{F}_T} (\sigma_n(u_h), \gamma(e - I(e)))_{0,F} \right) \\
 &= \sum_{T \in \mathcal{T}_h} (f + \operatorname{div} \sigma(u_h), e - I(e))_{0,T} + \sum_{F \in \mathcal{F}^\circ} ([\sigma_n(u_h)], \gamma(e - I(e)))_{0,F} \\
 &\quad + \sum_{F \in \mathcal{F}_1} (q - \sigma_n(u_h), \gamma(e - I(e)))_{0,F} \\
 &\leq \sum_{T \in \mathcal{T}_h} \left(R_T \|e - I(e)\|_{0,T} + \sum_{F \in \mathcal{F}_T} R_F \|e - I(e)\|_{0,F} \right), \tag{IV.23}
 \end{aligned}$$

wobei die element- bzw. facettenbezogenen Residuen R_T bzw. R_F als

$$R_T := \|f + \operatorname{div} \sigma(u_h)\|_{0,T}, \quad R_F := \begin{cases} \frac{1}{2} \|[\sigma_n(u_h)]\|_{0,F}, & F \in \mathcal{F}^\circ \\ \|q - \sigma_n(u_h)\|_{0,F}, & F \in \mathcal{F}_1 \end{cases}$$

definiert sind. Die Menge \mathcal{F}° ist die Menge aller inneren Facetten von \mathcal{T}_h , \mathcal{F}_T die Menge aller Facetten von T und \mathcal{F}_1 die Menge aller Facetten auf Γ_1 .

Ist I die übliche Clément-Interpolierende, erhält man aus (IV.23) den Ausdruck (IV.21) mit den zu diesem Problem angepassten Normen. Setzt man dagegen die Existenz einer hp -Clément-Interpolierende voraus, erhält man eine analoge Aussage zu Satz IV.5:

Satz IV.6 *Es seien $u \in H^1(\Omega, \Gamma_0)^3$ und $u_h \in S^p(\mathcal{T}_h)^3$ Lösung von (II.39) bzw. (III.30) sowie*

$$\eta^2 := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_T^2 + \sum_{F \in \mathcal{F}_T} (h_F/p_F) R_F^2 \right).$$

Ferner existiere eine hp -Clément-Interpolierende, und (IV.22) sei erfüllt mit $\varrho_0 > 0$ und $\varrho_1 > 0$. Dann existiert eine von \mathcal{T}_h und p unabhängige Konstante $C > 0$, so dass gilt:

$$\|u - u_h\|_1^2 \leq C\eta^2.$$

In Abbildung IV.4 sind die Effektivitätsindizes für den in Satz IV.5 aufgeführten Fehlerschätzer η bezüglich h - und p -Verfeinerungen des Einheitswürfels zu sehen. Die analytische Lösung ist hier $u_i := \cos(0.5\pi x) \cos(0.5\pi y) \cos(\pi/2z)$, $i = 1, 2, 3$. Für h -Verfeinerungen sind die zugehörigen Effektivitätsindizes konstant. Bei p -Verfeinerungen, schwanken die Effektivitätsindizes im Bereich von 0.1 bis 0.85.

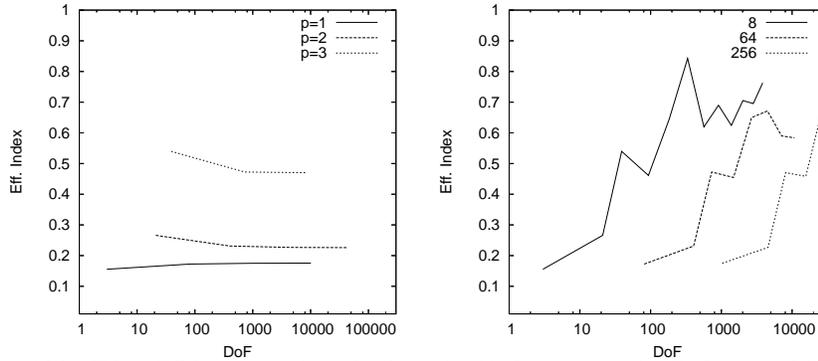


Abb. IV.4: Effektivitätsindizes für h - und p -Verfeinerungen des Einheitswürfels.

IV.3 Fehlerschätzer für vereinfachte Kontaktprobleme

Das Ziel dieses Abschnitts besteht in der Konkretisierung der in Abschnitt IV.1 hergeleiteten allgemeinen Fehlerschätzer für die in Abschnitt II.2 beschriebenen Modellprobleme. Hierzu werden die in Abschnitt IV.2.1 für modellhafte Variationsgleichungen bereitgestellten residualen Fehlerschätzer verwendet.

IV.3.1 Ein residualer Fehlerschätzer für vereinfachte Signorini-Probleme

Es sei $(u, \lambda_0) \in H^1(\Omega, \Gamma_0) \times \Lambda_0$ Lösung von (II.19)-(II.20) und $(u_h, \lambda_{0,H}) \in S^p(\mathcal{T}_h) \times \Lambda_0 \cap M^{p_1}(\mathcal{T}_{1,H})$ Lösung von (III.19)-(III.20). Das unter (IV.9) aufgeführte Hilfsproblem lautet:

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla u_0, \nabla v)_0 = (f, v)_0 + (q, \gamma(v))_{0, \Gamma_1} - (\lambda_{0,H}, \gamma(v))_{0, \Gamma_1}. \quad (\text{IV.24})$$

Aus Satz IV.5 folgt, dass η_0 mit

$$\eta_0^2 := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_{0,T}^2 + \sum_{E \in \mathcal{E}_T} (h_E/p_E) R_{0,E}^2 \right), \quad (\text{IV.25})$$

$$R_{0,T} := \|f + \Delta u_h\|_{0,T}, \quad R_{0,E} := \begin{cases} \frac{1}{2} \|\partial_n u_h\|_{0,E}, & E \in \mathcal{E}^\circ \\ \|q - \partial_n u_h - \lambda_{0,H}\|_{0,E}, & E \in \mathcal{E}_1 \end{cases}$$

ein Fehlerschätzer für das Hilfsproblem (IV.24) ist.

Aus Satz IV.1 bekommt man damit

$$\|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0,H}\|_{-1/2,\Gamma_1}^2 \leq C(\eta_0^2 + \|d\|_1^2 + |(\lambda_{0,H}, \gamma(d))_{0,\Gamma_1}|) \quad (\text{IV.26})$$

für alle $d \in \tilde{K} := \{v \in H^1(\Omega, \Gamma_0) \mid g - \gamma(u_h) - \gamma(d) \leq 0\}$.

Wie bereits in Abschnitt IV.1.1 erwähnt, kann d als harmonische Fortsetzung der Projektion von $g - \gamma(u_h)$ in G oder des positiven Anteils $(g - \gamma(u_h))_+$ gewählt werden. Im Folgenden wird die letzte Möglichkeit bevorzugt, da sie praktikabler erscheint. Jedoch lassen sich alle Schritte auch für die Projektion von $g - \gamma(u_h)$ in G in analoger Weise durchführen.

Gesucht wird ein $d^* \in W := \{v \in H^1(\Omega, \Gamma_0) \mid \gamma(v) = (g - \gamma(u_h))_+\}$ als Lösung des Minimierungsproblems

$$\|d^*\|_1^2 = \inf_{v \in W} \|v\|_1^2. \quad (\text{IV.27})$$

Die Menge W ist als affiner Unterraum von $H^1(\Omega, \Gamma_0)$ abgeschlossen und konvex. Zudem ist $\|\cdot\|_1^2$ konvex, stetig und koerziv. Nach Satz I.3 besitzt (IV.27) eine Lösung $d^* \in W$.⁹ Für diese ist offenbar $g - \gamma(u_h) - \gamma(d^*) \leq 0$, also gilt tatsächlich $d^* \in \tilde{K}$. Wegen $\|d^*\|_1 = \|(g - \gamma(u_h))_+\|_{1/2}$ folgt aus (IV.26)

$$\begin{aligned} & \|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0,H}\|_{-1/2,\Gamma_1}^2 \\ & \leq C(\eta_0^2 + \|d^*\|_1^2 + |(\lambda_{0,H}, \gamma(d^*))_{0,\Gamma_1}|) \\ & = C(\eta_0^2 + \|(g - \gamma(u_h))_+\|_{1/2}^2 + |(\lambda_{0,H}, (g - \gamma(u_h))_+)_{0,\Gamma_1}|). \end{aligned}$$

⁹Nach Satz I.1 ist die Bedingung $\forall v \in H^1(\Omega, \Gamma_0) : (\nabla d^*, \nabla v)_0 + (d^*, v)_0 = 0$ notwendig und hinreichend.

Satz IV.7 *Es sei $(u, \lambda_0) \in H^1(\Omega, \Gamma_0) \times \Lambda_0$ Lösung von (II.19)-(II.20) und $(u_h, \lambda_{0,H}) \in S^p(\mathcal{T}_h) \times \Lambda_0 \cap MP^1(\mathcal{T}_H)$ Lösung von (III.19)-(III.20). Der Fehlerschätzer η_0 sei wie in (IV.25) definiert. Dann existiert ein $C > 0$, so dass gilt:*

$$\begin{aligned} & \|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0,H}\|_{-1/2, \Gamma_1}^2 \\ & \leq C(\eta_0^2 + \|(g - \gamma(u_h))_+\|_{1/2}^2 + |(\lambda_{0,H}, (g - \gamma(u_h))_+)_{0, \Gamma_1}|). \end{aligned} \quad (\text{IV.28})$$

Für die praktische Umsetzung von (IV.28) ist die Berechnung von $\|(g - \gamma(u_h))_+\|_{1/2}$ erforderlich, was allerdings mit gewissen Schwierigkeiten verbunden ist. Wird jedoch $g \in H^1(\Gamma)$ vorausgesetzt, dann ist $(g - \gamma(u_h))_+ \in H^1(\Gamma)$, da $\beta(u_h)$ stückweise polynomial ist.¹⁰ Mit Hilfe der folgenden Aussage über Interpolationsräume¹¹ kann der Ausdruck $\|(g - \gamma(u_h))_+\|_{1/2}$ weiter abgeschätzt werden:

Lemma IV.3 *Es ist*

$$[L_2(\Gamma), H^1(\Gamma)]_{1/2, 2} = H^{1/2}(\Gamma).$$

Beweis: Th.7.7. in [76]. □

Damit folgt aus Satz A.9 mit $\tilde{C} := C(1/2, 2)$

$$\|(g - \gamma(u_h))_+\|_{1/2}^2 \leq \tilde{C} \|(g - \gamma(u_h))_+\|_{0, \Gamma} \|(g - \gamma(u_h))_+\|_{1, \Gamma} \quad (\text{IV.29})$$

und schließlich aus (IV.28)

$$\begin{aligned} & \|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0,H}\|_{-1/2, \Gamma_1}^2 \\ & \leq \max\{C, \tilde{C}\}(\eta_0^2 + \|(g - \gamma(u_h))_+\|_{0, \Gamma} \|(g - \gamma(u_h))_+\|_{1, \Gamma} \\ & \quad + |(\lambda_{0,H}, (g - \gamma(u_h))_+)_{0, \Gamma_1}|). \end{aligned}$$

Korollar IV.1 *Es gelten die Voraussetzungen von Satz IV.7. Ferner sei $g \in H^{1/2}(\Gamma_1) \cap H^1(\Gamma)$. Dann existiert ein $\hat{C} > 0$, so dass gilt:*

$$\begin{aligned} & \|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0,H}\|_{-1/2, \Gamma_1}^2 \\ & \leq \hat{C}(\eta_0^2 + \|(g - \gamma(u_h))_+\|_{0, \Gamma} \|(g - \gamma(u_h))_+\|_{1, \Gamma} + |(\lambda_{0,H}, (g - \gamma(u_h))_+)_{0, \Gamma_1}|). \end{aligned}$$

¹⁰vgl. z.B. Ch. I, Cor.2.1 in [56]

¹¹vgl. Satz A.9

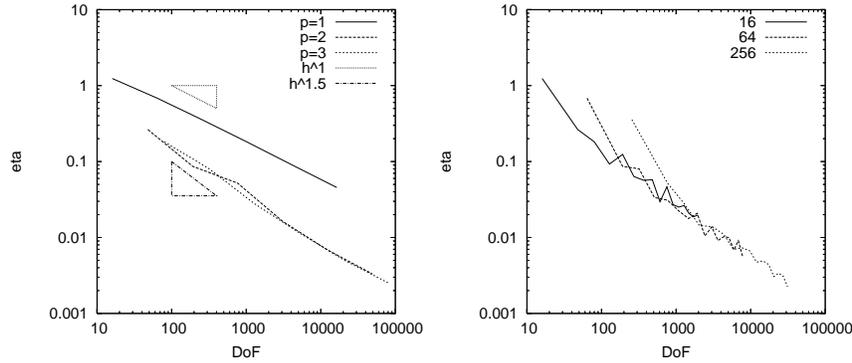


Abb. IV.5: Geschätztes Konvergenzverhalten bei h - und p -Verfeinerungen.

In Abbildung IV.5 ist das durch den Fehlerschätzer aus Korollar IV.1 geschätzte Konvergenzverhalten für h - und p -Verfeinerungen dargestellt. Die den Diagrammen zugrunde liegende Beispielrechnung entspricht der Beispielkonfiguration aus Abschnitt II.2.1.

Das geschätzte Konvergenzverhalten bei h -Verfeinerungen mit $p = 1$ entspricht exakt den a-priori Vorhersagen der Größenordnung $\mathcal{O}(h)$ aus den Sätzen III.6 und III.8. Das geschätzte Konvergenzverhalten für $p = 2$ oder $p = 3$ ist etwa $\mathcal{O}(h^{3/2})$. Demnach führt die Verwendung von Finite-Elemente-Ansätzen höherer Ordnung nicht zu dem gewünschten Konvergenzgewinn, sofern globale h -Verfeinerungen vorgenommen werden. Der Grund hierfür ist die fehlende Regularität der Lösung u , die im Allgemeinen nicht mehr als H^2 -regulär ist.¹²

Das geschätzte Konvergenzverhalten bei reinen p -Verfeinerungen kommt gemäß Abbildung IV.5 (rechts) ebenfalls nicht über die Ordnung $\mathcal{O}(h^{3/2})$ hinaus.

In den Abbildungen IV.6 und IV.7 sind die Fehleranteile

$$s_0 := \|(g - \gamma(u_h))_+\|_{0,\Gamma}^{1/2} \|(g - \gamma(u_h))_+\|_{1,\Gamma}^{1/2}$$

$$s_1 := |(\lambda_{0,H}, (g - \gamma(u_h))_+)_{0,\Gamma_1}|^{1/2}$$

für h - und p -Verfeinerungen abgebildet. Offenbar sind die Fehleranteile s_0 von höherer Ordnung und können deshalb u.U. vernachlässigt werden. Die Fehleranteile s_1 verringern sich in der gleichen Größenordnung wie η , fallen aber bei den Beispielrechnungen um den Faktor 10 kleiner aus.

¹²vgl. [28], [57], S. 416

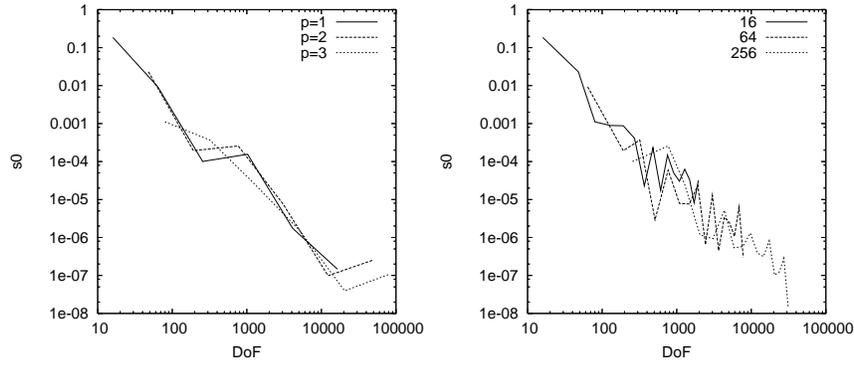


Abb. IV.6: Fehleranteile s_0 bei h - und p -Verfeinerungen.

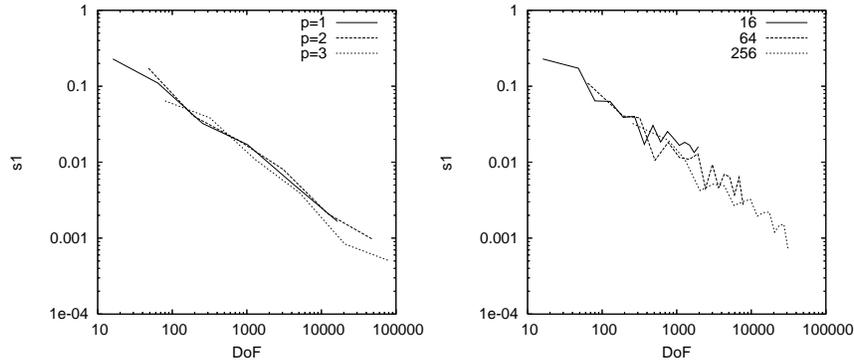


Abb. IV.7: Fehleranteile s_1 bei h - und p -Verfeinerungen.

IV.3.2 Ein residualer Fehlerschätzer für idealisierte Reibungsprobleme

Es sei $(u, \lambda_1) \in H^1(\Omega, \Gamma_0) \times \Lambda_1$ Lösung von (II.33)-(II.34), sowie $(u_h, \lambda_{1,H}) \in S^p(\mathcal{T}_h) \times \Lambda_1 \cap M^{p_1}(\mathcal{T}_{1,H})$ Lösung von (III.28)-(III.29). Das unter (IV.12) aufgeführte Hilfsproblem lautet:

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla u_1, \nabla v)_0 = (f, v)_0 + (q, \gamma(v))_{0, \Gamma_1} - (\lambda_{1,H}, \gamma(v))_{0, \Gamma_1}. \quad (\text{IV.30})$$

Einen Fehlerschätzer η_1 für das Hilfsproblem (IV.30) erhält man gemäß Satz IV.5 mit

$$\eta_1^2 := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_{1,T}^2 + \sum_{E \in \mathcal{E}_T} (h_E/p_E) R_{1,E}^2 \right), \quad (\text{IV.31})$$

$$R_{1,T} := \|f + \Delta u_h\|_{0,T}, \quad R_{1,E} := \begin{cases} \frac{1}{2} \|[\partial_n u_h]\|_{0,E}, & E \in \mathcal{E}^\circ \\ \|q - \partial_n u_h - \lambda_{1,H}\|_{0,E}, & E \in \mathcal{E}_1. \end{cases}$$

Damit folgt aus Satz IV.2:

Satz IV.8 *Es sei $(u, \lambda_1) \in H^1(\Omega, \Gamma_0) \times \Lambda_1$ Lösung von (II.33)-(II.34) und $(u_h, \lambda_{1,H}) \in S^p(\mathcal{T}_h) \times \Lambda_1 \cap M^{p_1}(\mathcal{T}_H)$ Lösung von (III.28)-(III.29). Der Fehlerschätzer η_1 sei wie in (IV.31) definiert. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1^2 + \|\lambda_1 - \lambda_{1,H}\|_{-1/2, \Gamma_1}^2 \leq C(\eta_1^2 + |j(u_h) - (\lambda_{1,H}, \gamma(u_h))_{0, \Gamma_1}|).$$

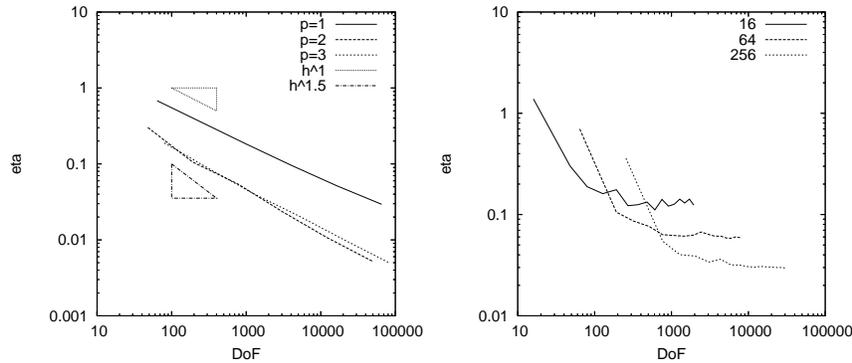


Abb. IV.8: Geschätztes Konvergenzverhalten bei h - und p -Verfeinerungen.

Für das hier behandelte Modellproblem erhält man zu Abschnitt IV.3.1 analoge Resultate bei Rechnungen mit h - und p -Verfeinerungen des in Abschnitt II.2.2 dargestellten Beispiels (vgl. Abbildung IV.8).

Das geschätzte Konvergenzverhalten bei h -Verfeinerungen mit $p = 1$ ist von der Ordnung $\mathcal{O}(h)$ und korrespondiert demnach zu den entsprechenden a-priori Aussagen in Satz III.9. Bei Finite-Elemente-Ansätzen höherer Ordnung ist kein wesentlicher Konvergenzzugewinn festzustellen. Die Konvergenzordnung liegt hier zwischen $\mathcal{O}(h)$ und $\mathcal{O}(h^{3/2})$. Die fehlende Regularität

der Lösung u verhindert auch hier eine höhere Konvergenzordnung. Insbesondere ist der Einsatz reiner p -Verfeinerungen bei dem hier gewählten Beispiel nicht sinnvoll.

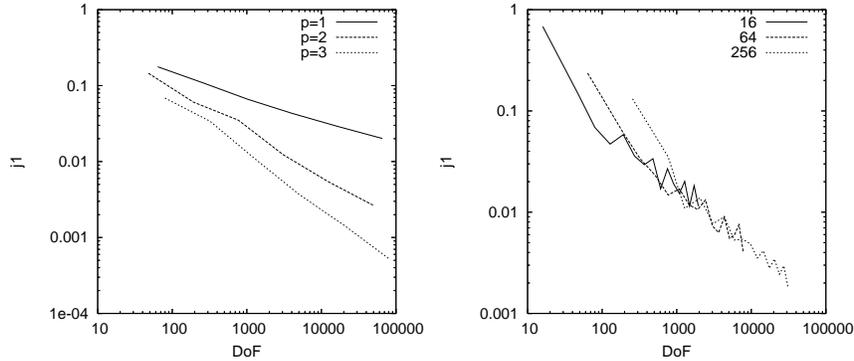


Abb. IV.9: Fehleranteile j_1 bei h - und p -Verfeinerungen.

In Abbildung IV.9 sind die Anteile $j_1 := |j(u_h) - (\lambda_{1,H}, \gamma(u_h))_{0,\Gamma_1}|^{1/2}$ des in Satz IV.8 dargestellten Fehlerschätzers abgebildet. Sowohl für h - als auch für p -Verfeinerungen sind die Fehleranteile j_1 in etwa von der gleichen Größen- bzw. Konvergenzordnung wie η .

IV.4 Fehlerschätzer für Kontaktprobleme

Wie zuvor besteht auch in diesem Abschnitt das Ziel darin, die in Abschnitt IV.1 hergeleiteten allgemeinen Fehlerschätzer mit Hilfe von Fehlerschätzern für Variationsgleichungen zu konkretisieren. Hierbei werden die in Abschnitt IV.2.2 aufgeführten Fehlerschätzer verwendet.

IV.4.1 Ein residualer Fehlerschätzer für reibungsfreie Kontaktprobleme

Es sei $(u, \lambda_n) \in H^1(\Omega, \Gamma_0)^3 \times \Lambda_n$ Lösung von (II.56)-(II.57) und $(u_h, \lambda_{n,H}) \in S^p(\mathcal{T}_h) \times \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H})$ Lösung von (III.32)-(III.33). Das unter (IV.9) aufgeführte Hilfsproblem ist hier gegeben durch:

$$\forall v \in H^1(\Omega, \Gamma_0)^3 : (\sigma(u_n), \varepsilon(v))_0 = (f, v)_0 + (q, \delta(v))_{0,\Gamma_1} - (\lambda_{n,H}, \delta_n(v))_{0,\Gamma_1}. \quad (\text{IV.32})$$

Aus Satz IV.6 folgt, dass η_n mit

$$\begin{aligned} \eta_n^2 &:= \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_{n,T}^2 + \sum_{F \in \mathcal{F}_T} (h_F/p_F) R_{n,F}^2 \right), & (IV.33) \\ R_{n,T} &:= \|f + \operatorname{div} \sigma(u_h)\|_{0,T}, \\ R_{n,F} &:= \begin{cases} \frac{1}{2} \|[\sigma_n(u_h)]\|_{0,F}, & F \in \mathcal{F}^\circ \\ (\|q_n - \sigma_{nn}(u_h) - \lambda_{n,H}\|_{0,F}^2 + \|q_t - \sigma_{nt}(u_h)\|_{0,F}^2)^{1/2}, & F \in \mathcal{F}_1 \end{cases} \end{aligned}$$

ein Fehlerschätzer für das Hilfsproblem (IV.32) ist.

Aus Satz IV.1 erhält man damit

$$\|u - u_h\|_1^2 + \|\lambda_n - \lambda_{n,H}\|_{-1/2,n}^2 \leq C(\eta_n^2 + \|d\|_1^2 + |(\lambda_{n,H}, \delta_n(d))_{0,\Gamma_1}|) \quad (IV.34)$$

für alle $d \in \tilde{K}_n := \{v \in H^1(\Omega, \Gamma_0)^3 \mid g - \delta_n(u_h) - \delta_n(v) \geq 0\}$.

Eine geeignete Wahl von $d \in \tilde{K}$ ist, wie in Abschnitt IV.3.1 bereits beschrieben, die harmonische Fortsetzung von $(g - \delta_n(u_h))_-$: Gesucht ist ein $d^* \in W_n := \{v \in H^1(\Omega, \Gamma_0)^3 \mid (\delta_n(v), \delta_t(v)) = ((g - \delta_n(u_h))_-, 0)\}$, das Lösung des Minimierungsproblems

$$\|d^*\|_1^2 = \inf_{v \in W_n} \|v\|_1^2$$

ist. Wie in Abschnitt IV.3.1 sichert auch hier Satz I.3 die Existenz einer Lösung $d^* \in W_n$.¹³ Zudem ist $g - \delta_n(u_h) - \delta_n(d^*) \geq 0$ und damit $d^* \in \tilde{K}$.

Damit folgt aus (IV.34):

Satz IV.9 *Es sei $(u, \lambda_0) \in H^1(\Omega, \Gamma_0)^3 \times \Lambda_0$ Lösung von (II.56)-(II.57) und $(u_h, \lambda_{0,H}) \in S^p(\mathcal{T}_h) \times \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H})$ Lösung von (III.32)-(III.33). Der Fehlerschätzer η_n sei wie in (IV.33) definiert. Dann existiert ein $C > 0$, so dass gilt:*

$$\begin{aligned} \|u - u_h\|_1^2 + \|\lambda_n - \lambda_{n,H}\|_{-1/2,n}^2 \\ \leq C(\eta_n^2 + \|(g - \delta_n(u_h))_-\|_{1/2,n} + |(\lambda_{n,H}, (g - \delta_n(u_h))_-)_{0,\Gamma_1}|). \end{aligned}$$

Wie in Korollar IV.1 zeigt man:

¹³Nach Satz I.1 ist $\forall v \in H^1(\Omega, \Gamma_0)^3 : (\varepsilon(d^*), \varepsilon(v))_0 + (d^*, v)_0 = 0$ eine notwendige und hinreichende Bedingung für die Minimalität von d^* .

Korollar IV.2 *Es gelten die Voraussetzungen von Satz IV.9. Ferner sei $\tilde{\delta}^{-1}((g - \delta_n(u_h), 0)_-) \in H^{1/2}(\Gamma_1)^3 \cap H^1(\Gamma)^3$. Dann existiert ein $\hat{C} > 0$, so dass gilt:*

$$\begin{aligned} & \|u - u_h\|_1^2 + \|\lambda_n - \lambda_{n,H}\|_{-1/2,n}^2 \\ & \leq \hat{C}(\eta_n^2 + \|\tilde{\delta}^{-1}((g - \delta_n(u_h))_-, 0)\|_{0,\Gamma} \|\tilde{\delta}^{-1}((g - \delta_n(u_h))_-, 0)\|_{1,\Gamma} \\ & \quad + |(\lambda_{n,H}, (g - \delta_n(u_h))_-, 0)_{0,\Gamma_1}|). \end{aligned}$$

IV.4.2 Ein residualer Fehlerschätzer für Reibungsprobleme mit vorgegebener Normalspannung

Es sei $(u, \lambda_t) \in H^1(\Omega, \Gamma_0)^3 \times \Lambda_t$ Lösung von (II.74)-(II.75), sowie $(u_h, \lambda_{t,H}) \in S^p(\mathcal{T}_h) \times \Lambda_t \cap M^{p_1}(\mathcal{T}_{1,H})^2$ Lösung von (III.35)-(III.36). Das Hilfsproblem (IV.12) hat in diesem Fall die Gestalt:

$$\forall v \in H^1(\Omega, \Gamma_0)^3 : (\sigma(u_t), \varepsilon(v))_0 = (f, v)_0 + (q, \delta(v))_{0,\Gamma_1} - (\lambda_{t,H}, \delta_t(v))_{0,\Gamma_1}. \quad (\text{IV.35})$$

Ein Fehlerschätzer η_t für das Hilfsproblem (IV.35) ist laut Satz IV.6 mit

$$\eta_t^2 := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_{t,T}^2 + \sum_{F \in \mathcal{F}_T} (h_F/p_F) R_{t,F}^2 \right), \quad (\text{IV.36})$$

$$R_{t,T} := \|f + \operatorname{div} \sigma(u_h)\|_{0,T},$$

$$R_{t,F} := \begin{cases} \frac{1}{2} \|\sigma_n(u_h)\|_{0,F}, & F \in \mathcal{F}^\circ \\ (\|q - \sigma_{nn}(u_h)\|_{0,F}^2 + \|q - \sigma_{nt}(u_h) - \lambda_{t,H}\|_{0,F}^2)^{1/2}, & F \in \mathcal{F}_1 \end{cases}$$

gegeben. Nach Satz IV.2 erhält man schließlich:

Satz IV.10 *Es sei $(u, \lambda_t) \in H^1(\Omega, \Gamma_0)^3 \times \Lambda_t$ Lösung von (II.74)-(II.75) und $(u_h, \lambda_{t,H}) \in S^p(\mathcal{T}_h) \times \Lambda_t \cap M^{p_1}(\mathcal{T}_{1,H})^2$ Lösung von (III.35)-(III.36). Der Fehlerschätzer η_t sei wie in (IV.36) definiert. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1^2 + \|\lambda_t - \lambda_{t,H}\|_{-1/2,t}^2 \leq C(\eta_t^2 + |j(u_h) - (\lambda_{t,H}, \delta_t(u_h))_{0,\Gamma_1}|).$$

IV.4.3 Ein residualer Fehlerschätzer für Kontaktprobleme mit Reibung

Es sei $(u, \lambda_n, \lambda_t) \in H^1(\Omega, \Gamma_0)^3 \times \Lambda_n \times \Lambda_t$ Lösung von (II.88)-(II.89) und $(u_h, \lambda_{n,H}, \lambda_{t,H}) \in S^p(\mathcal{T}_h) \times \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H}) \times \Lambda_t \cap M^{p_1}(\mathcal{T}_{1,H})^2$ Lösung von

(III.38)-(III.39). Das Hilfsproblem (IV.15) ist gegeben durch:

$$\begin{aligned} \forall v \in H^1(\Omega, \Gamma_0)^3 : (\sigma(u_{nt}), \varepsilon(v))_0 \\ = (f, v)_0 + (q, \delta(v))_{0, \Gamma_1} - (\lambda_{n,H}, \delta_n(v))_{0, \Gamma_1} - (\lambda_{t,H}, \delta_t(v))_{0, \Gamma_1}. \end{aligned} \quad (\text{IV.37})$$

Aus Satz IV.6 folgt, dass η_{nt} mit

$$\begin{aligned} \eta_{nt}^2 &:= \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_{nt,T}^2 + \sum_{F \in \mathcal{F}_T} (h_F/p_F) R_{nt,F}^2 \right), \quad (\text{IV.38}) \\ R_{nt,T} &:= \|f + \operatorname{div} \sigma(u_h)\|_{0,T}, \\ R_{nt,F} &:= \begin{cases} \frac{1}{2} \|\sigma_n(u_h)\|_{0,F}, & F \in \mathcal{F}^\circ \\ (\|q_n - \sigma_{nn}(u_h) - \lambda_{n,H}\|_{0,F}^2 \\ + \|q_t - \sigma_{nt}(u_h) - \lambda_{t,H}\|_{0,F}^2)^{1/2}, & F \in \mathcal{F}_1 \end{cases} \end{aligned}$$

ein Fehlerschätzer für das Hilfsproblem (IV.37) ist.

Mit den Bezeichnungen $\langle \lambda, (\cdot, \cdot) \rangle := \langle \lambda_n, \cdot \rangle + \langle \lambda_t, \cdot \rangle$ und $\langle \lambda_H, (\cdot, \cdot) \rangle := \langle \lambda_{n,H}, \cdot \rangle + \langle \lambda_{t,H}, \cdot \rangle$ liefert Satz IV.3 für alle $d \in \tilde{K}_n$:

$$\begin{aligned} \|u - u_h\|_1^2 + \|\lambda - \lambda_H\|_{-1/2, nt}^2 \leq C(\eta_n^2 + \|d\|_1^2 + |(\lambda_{n,H}, \delta_n(d))_{0, \Gamma_1}| \\ + |j(d) - (\lambda_{t,H}, \delta_t(d))_{0, \Gamma_1}| + |j(u_h) - (\lambda_{t,H}, \delta_t(u_h))_{0, \Gamma_1}|). \end{aligned} \quad (\text{IV.39})$$

Eine geeignete Wahl für $d \in \tilde{K}$ ist wie in Abschnitt IV.4.1 die harmonische Fortsetzung d^* von $((g - \delta_n(u_h))_-, 0)$.

Damit folgt aus (IV.39) schließlich:

Satz IV.11 *Es sei $(u, \lambda_n, \lambda_t) \in H^1(\Omega, \Gamma_0)^3 \times \Lambda_n \times \Lambda_t$ Lösung von (II.88)-(II.89), zudem sei $(u_h, \lambda_{n,H}, \lambda_{t,H}) \in S^p(\mathcal{T}_h) \times \Lambda_n \cap M^{p_1}(\mathcal{T}_{1,H}) \times \Lambda_t \cap M^{p_1}(\mathcal{T}_{1,H})^2$ Lösung von (III.38)-(III.39). Der Fehlerschätzer η_{nt} sei wie in (IV.38) definiert. Dann existiert ein $C > 0$, so dass gilt:*

$$\begin{aligned} \|u - u_h\|_1^2 + \|\lambda - \lambda_H\|_{-1/2, nt}^2 \leq C(\eta_{nt}^2 + \|(g - \delta_n(u_h))_-\|_{1/2, n} \\ + |(\lambda_{n,H}, (g - \delta_n(u_h))_-)_{0, \Gamma_1}| + |j(u_h) - (\lambda_{t,H}, \delta_t(u_h))_{0, \Gamma_1}|). \end{aligned}$$

Wie in Korollar IV.1 zeigt man:

Korollar IV.3 *Es gelten die Voraussetzungen von Satz IV.11. Ferner sei $\tilde{\delta}^{-1}((g - \delta_n(u_h), 0)_-) \in H^{1/2}(\Gamma_1)^3 \cap H^1(\Gamma)^3$. Dann existiert ein $\hat{C} > 0$, so dass gilt:*

$$\begin{aligned} & \|u - u_h\|_1^2 + \|\lambda - \lambda_H\|_{-1/2, nt}^2 \\ & \leq \hat{C}(\eta_{nt}^2 + \|\tilde{\delta}^{-1}((g - \delta_n(u_h))_-, 0)\|_{0, \Gamma} \|\tilde{\delta}^{-1}((g - \delta_n(u_h))_-, 0)\|_{1, \Gamma} \\ & \quad + |(\lambda_{n, H}, (g - \delta_n(u_h))_-, 0)_{\Gamma_1}| + |j(u_h) - (\lambda_{t, H}, \delta_t(u_h))_{0, \Gamma_1}|). \end{aligned}$$

IV.5 Residuale Fehlerschätzer für weitere Modellprobleme

Auch auf die in den Abschnitten II.5.1 und II.5.3 dargestellten Hindernisprobleme und Bingham-Fluid-Probleme können die in Abschnitt IV.1 beschriebenen Techniken zur Fehlerkontrolle angewandt werden.

Das in Abschnitt II.5.2 formulierte modellhafte Torsionsproblem entzieht sich dem allgemeinen Kontext aus Abschnitt IV.1. Für Fehlerschätzer in Bezug auf Torsionsprobleme bei Beschränkung auf stückweise lineare bzw. biquadratische FE-Ansätze sei auf [95] und [97] verwiesen.

IV.5.1 Hindernisprobleme

Es sei $(u, \lambda_0) \in H^1(\Omega, \Gamma_0) \times \Lambda_0$ Lösung von (II.92)-(II.93) und $(u_h, \lambda_{0, H}) \in S^p(\mathcal{T}_h) \times \Lambda_0 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)$ Lösung von (III.44)-(III.45). Das unter (IV.9) aufgeführte Hilfsproblem lautet:

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla u_0, \nabla v)_0 = (f, v)_0 + (q, \gamma(v))_{0, \Gamma_1} - (\lambda_{0, H}, v)_0. \quad (\text{IV.40})$$

Aus Satz IV.5 folgt, dass η_0 mit

$$\begin{aligned} \eta_0^2 & := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_{0, T}^2 + \sum_{E \in \mathcal{E}_T} (h_E/p_E) R_{0, E}^2 \right), \quad (\text{IV.41}) \\ R_{0, T} & := \|f + \Delta u_h - \lambda_{0, H}\|_{0, T}, \quad R_{0, E} := \begin{cases} \frac{1}{2} \|[\partial_n u_h]\|_{0, E}, & E \in \mathcal{E}^\circ \\ \|q - \partial_n u_h\|_{0, E}, & E \in \mathcal{E}_1 \end{cases} \end{aligned}$$

ein Fehlerschätzer für das Hilfsproblem (IV.40) ist. Aus Satz IV.1 folgt damit

$$\|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0, H}\|^2 \leq C(\eta_0^2 + \|d\|_1^2 + |(\lambda_{0, H}, d)_0|) \quad (\text{IV.42})$$

für alle $d \in \tilde{K} := \{v \in H^1(\Omega, \Gamma_0) \mid g - u_h - v \leq 0\}$.

Einen konsistenten Fehlerschätzer erhält man, wenn d als Projektion von $g - u_h$ auf $G := \{v \in H^1(\Omega, \Gamma_0) \mid v \leq 0\}$ oder als positiver Anteil von $g - u_h$ gewählt wird. Bei Verwendung des positiven Anteils¹⁴ $d := (g - u_h)_+$ folgt aus (IV.42)

$$\|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0,H}\|^2 \leq C(\eta_0^2 + \|(g - u_h)_+\|_1^2 + |(\lambda_{0,H}, (g - u_h)_+)_0|).$$

Damit gilt der Satz:

Satz IV.12 *Es sei $(u, \lambda_0) \in H^1(\Omega, \Gamma_0) \times \Lambda_0$ Lösung von (II.92)-(II.93) und $(u_h, \lambda_{0,H}) \in S^p(\mathcal{T}_h) \times \Lambda_0 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)$ Lösung von (III.44)-(III.45). Der Fehlerschätzer η_0 sei wie in (IV.41) definiert. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1^2 + \|\lambda_0 - \lambda_{0,H}\|^2 \leq C(\eta_0^2 + \|(g - u_h)_+\|_1 + |(\lambda_{0,H}, (g - u_h)_+)_0|).$$

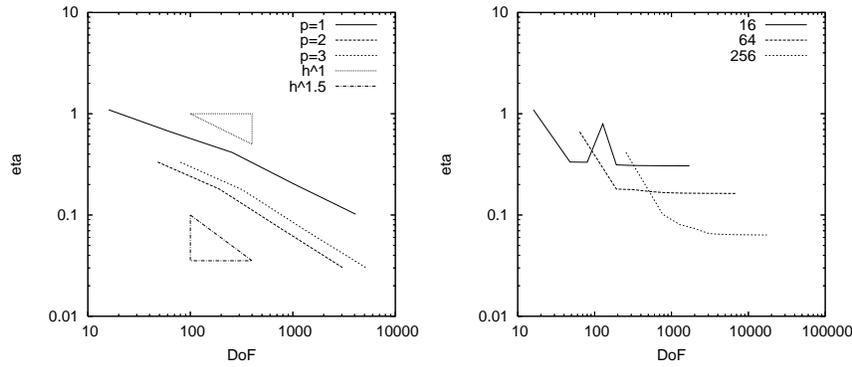


Abb. IV.10: Geschätztes Konvergenzverhalten bei h - und p -Verfeinerungen.

In Abbildung IV.10 ist das geschätzte Konvergenzverhalten für h - und p -Verfeinerungen zu sehen, das mit Hilfe des in Satz IV.12 dargestellten Fehlerschätzers ermittelt wird. Der Rechnung liegt das Beispiel aus Abschnitt II.5.1 zugrunde.

Das geschätzte Konvergenzverhalten bei h -Verfeinerungen mit $p = 1$ entspricht der in Satz III.11 vorhergesagten Konvergenzordnung $\mathcal{O}(h)$. Das geschätzte Konvergenzverhalten für $p = 2$ oder $p = 3$ liegt unterhalb von

¹⁴Man beachte hierbei, dass aus $g \in H^1(\Omega)$ auch $(g - u_h)_+ \in H^1(\Omega)$ folgt.

$\mathcal{O}(h^{3/2})$, was ebenfalls den theoretischen Vorhersagen entspricht. Demnach führt auch hier der Gebrauch von Finite-Elemente-Ansätzen höherer Ordnung bei globalen h -Verfeinerungen nicht zu einem signifikanten Konvergenzgewinn.

Die Anwendung reiner p -Verfeinerungen ist offenbar unzureichend. Gemäß Abbildung IV.10 (rechts) kommt es fast zu einer Stagnation des geschätzten Fehlers.

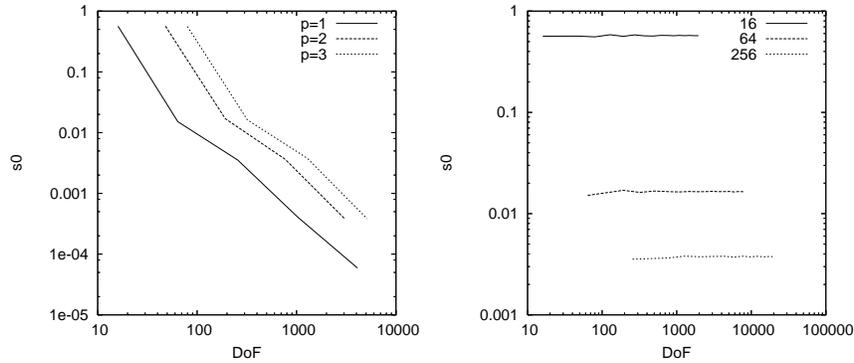


Abb. IV.11: Fehleranteile s_0 bei h - und p -Verfeinerungen.

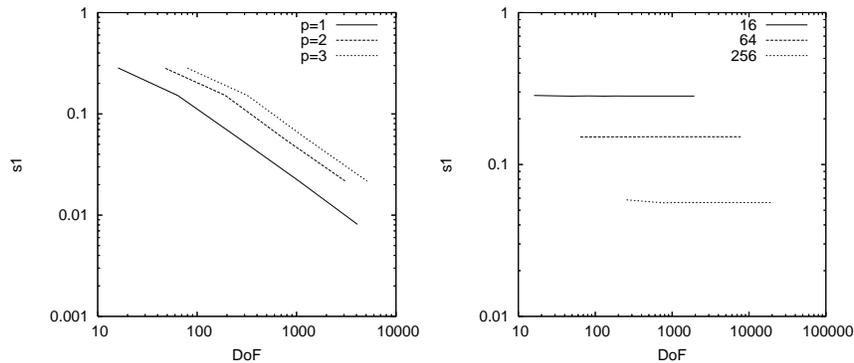


Abb. IV.12: Fehleranteile s_1 bei h - und p -Verfeinerungen.

In den Abbildungen IV.11 und IV.12 sind die Fehleranteile $s_0 := \|(g - u_h)_+\|_1$ und $s_1 := |(\lambda_{0,H}, (g - u_h)_+)_0|^{1/2}$ für h - und p -Verfeinerungen abgebildet. Wie bei Signorini-Problemen sind die Fehleranteile s_0 bei Verwen-

derung von h -Verfeinerungen offenbar von höherer Ordnung und können deshalb im Allgemeinen vernachlässigt werden. Berücksichtigt werden müssen dagegen die Fehleranteile s_1 , da sie der Größenordnung des geschätzten Gesamtfehlers entsprechen.

Die Abbildungen IV.11 und IV.12 (rechts) zeigen, dass bei Verwendung von p -Verfeinerungen die Fehleranteile s_0 und s_1 den Gesamtfehler dominieren, so dass es insgesamt zu einem Konvergenzstillstand kommt. Der Grund ist sicherlich die geringe Anpassung des verwendeten Gitters an das Hindernis.

IV.5.2 Bingham-Fluid-Probleme

Es sei $(u, \lambda_1) \in H_0^1(\Omega) \times \Lambda_1$ Lösung von (II.99)-(II.100) und $(u_h, \lambda_{1,H}) \in S^p(\mathcal{T}_h) \times \Lambda_1 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)^k$ Lösung von (III.46)-(III.47). Das Hilfsproblem (IV.9) lautet in diesem Fall:

$$\forall v \in H_0^1(\Omega) : (\nabla u_1, \nabla v)_0 = (f, v)_0 - (\lambda_{1,H}, \nabla v)_0. \quad (\text{IV.43})$$

Analog zu (IV.18) gilt mit $e := u_1 - u_h$

$$(\nabla e, \nabla e)_0 = (f, e - I(e))_0 + (\lambda_{1,H}, \nabla(e - I(e)))_0 - (\nabla u_h, \nabla(e - I(e)))_0.$$

Demnach kann der in Abschnitt (IV.2.1) vorgestellte residuale Fehler-schätzer nicht direkt auf das Hilfsproblem (IV.43) angewandt werden, da der Term $(\lambda_{1,H}, \nabla(e - I(e)))_0$ nicht die erforderlichen hp -Potenzen liefert. Ist $\mathcal{T}_h = \tilde{\mathcal{T}}_H$, erhält man durch Anwenden der Greenschen Formel (II.2):

$$\begin{aligned} & -(\lambda_{1,H}, \nabla(e - I(e)))_0 \\ &= \sum_{T \in \mathcal{T}_h} -(\lambda_{1,H}, \nabla(e - I(e)))_{0,T} \\ &= \sum_{T \in \mathcal{T}_h} (\operatorname{div} \lambda_{1,H}, e - I(e))_{0,T} - \sum_{E \in \mathcal{E}_T} (\gamma(\lambda_{0,H})n, \gamma(e - I(e)))_{0,E} \\ &= \sum_{T \in \mathcal{T}_h} (\operatorname{div} \lambda_{1,H}, e - I(e))_{0,T} - \frac{1}{2} \sum_{E \in \mathcal{E}^\circ} ([\lambda_{0,H}n], \gamma(e - I(e)))_{0,E}. \end{aligned}$$

Damit ist η_1 mit

$$\eta_1^2 := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_{1,T}^2 + \sum_{E \in \mathcal{E}_T} (h_E/p_E) R_{1,E}^2 \right), \quad (\text{IV.44})$$

$$R_{1,T} := \|f + \Delta u_h + \operatorname{div} \lambda_{0,H}\|_{0,T}, \quad R_{1,E} := \frac{1}{2} \|[\partial_n u_h] + [\lambda_{0,H}n]\|_{0,E}$$

ein Fehlerschätzer für das Hilfsproblem. Da $\beta \in L(H_0^1(\Omega), L^2(\Omega)^k)$ mit $\beta(v) := \beta_1(v) := \nabla v$ nicht surjektiv ist, ist Lemma IV.1 nicht anwendbar. Man erhält aber analog zu Lemma IV.2 und Satz IV.2:

Satz IV.13 *Es sei $(u, \lambda_1) \in H_0^1(\Omega) \times \Lambda_1$ Lösung von (II.99)-(II.100) und $(u_h, \lambda_{1,H}) \in S^p(\mathcal{T}_h) \times \Lambda_1 \cap \tilde{M}^{\tilde{p}}(\tilde{\mathcal{T}}_H)^k$ Lösung von (III.46)-(III.47) mit $\mathcal{T}_h = \tilde{\mathcal{T}}_H$. Der Fehlerschätzer η_1 sei wie in (IV.44) definiert. Dann existiert ein $C > 0$, so dass gilt:*

$$\|u - u_h\|_1^2 \leq C(\eta_1^2 + |(s, |\nabla u_h|)_0 - (\lambda_{1,H}, \gamma(u_h))_{0,\Gamma_1}|).$$

IV.6 Fehlerschätzer vom Goal-Oriented-Typ

Die bisher angegebenen Fehlerschätzer bilden obere Schranken für den Fehler $u - u_h$ gemessen in der H^1 -Norm $\|\cdot\|_1$, die letztlich ein gemitteltetes Fehlermaß darstellt. Wünschenswert ist jedoch häufig eine Fehlerkontrolle in Bezug auf problemorientierte Fehlermaße. Insbesondere sind Fehlermaße mit lokalem Charakter von großem Interesse. Als Beispiele seien der Fehler in einem Punkt oder der Fehler gemessen als Integralmittel über eine beliebige Teilmenge von Ω genannt.

Auch für die in Abschnitt II.4 beschriebenen fertigungstechnischen Anwendungen sind Fehlerkontrolltechniken mit lokal definierten Fehlermaßen von besonderer Bedeutung, da man oft an einer möglichst genauen Berechnung lokaler Größen wie Normalkraft oder Tangentialkraft interessiert ist.

Das Ziel des nun folgenden Abschnitts besteht darin, den grundlegenden Ansatz für Fehlerschätzer vom Goal-Oriented-Typ, der die Fehlerkontrolle für beliebige Fehlermaße aus V' ermöglicht, in dem hier gewählten allgemeinen Zusammenhang zu skizzieren.

Die Ausführungen für Variationsgleichungen basieren im Wesentlichen auf den Arbeiten [2] und [98]. Die Verallgemeinerungen für elliptische Minimierungsprobleme orientieren sich an den Arbeiten [94], [96] und [98].

IV.6.1 Fehlerschätzer vom Goal-Oriented-Typ für Variationsgleichungen

Es sei V ein Hilbertraum und $u \in V$ Lösung von

$$\forall v \in V : a(u, v) = \langle \ell, v \rangle \tag{IV.45}$$

mit einer stetigen und elliptischen Bilinearform $a : V \times V \rightarrow \mathbb{R}$, sowie $\ell \in V'$. Im Sinne von Diskretisierungsansätzen wie in Kapitel III sei $V_h \subset V$ ein weiterer Unterraum, und für ein $u_h \in V_h$ gelte:

$$\forall v_h \in V_h : a(u_h, v_h) = \langle \ell, v_h \rangle. \quad (\text{IV.46})$$

Es sei $J \in V'$ ein beliebiges Fehlermaß. Die Zielsetzung besteht in der Suche nach einem Fehlerschätzer für den Ausdruck

$$|\langle J, u \rangle - \langle J, u_h \rangle|. \quad (\text{IV.47})$$

Nach dem Lax-Milgram-Lemma¹⁵ existiert ein $z \in V$, so dass gilt:

$$\forall v \in V : a(v, z) = \langle J, v \rangle. \quad (\text{IV.48})$$

In diesem Zusammenhang heißt u die primale Lösung, während z duale Lösung genannt wird. Ist $V_{\bar{h}} \subset V$ ein Unterraum, sowie $z_{\bar{h}} \in V_{\bar{h}}$, so dass

$$\forall v_{\bar{h}} \in V_{\bar{h}} : a(v_{\bar{h}}, z_{\bar{h}}) = \langle J, v_{\bar{h}} \rangle \quad (\text{IV.49})$$

gilt, erhält man

$$\begin{aligned} \langle J, u - u_h \rangle &= a(u - u_h, z) = a(u - u_h, z - z_{\bar{h}}) + a(u - u_h, z_{\bar{h}}) \\ &= a(u - u_h, z - z_{\bar{h}}) + \langle \ell, z_{\bar{h}} \rangle - a(u_h, z_{\bar{h}}). \end{aligned} \quad (\text{IV.50})$$

Aus der Stetigkeit von a und aus (IV.50) folgt damit sofort:

Satz IV.14 *Es seien $u \in V$ und $u_h \in V_h$ Lösungen von (IV.45) bzw. (IV.46) sowie $z \in V$ und $z_{\bar{h}} \in V_{\bar{h}}$ Lösungen von (IV.48) bzw. (IV.49). Es seien $\eta_p, \eta_d \in \mathbb{R}_{\geq 0}$ Fehlerschätzer des primalen bzw. dualen Problems. Dann existieren $C, \bar{C} > 0$, so dass mit $\theta(u_h, z_{\bar{h}}) := \langle \ell, z_{\bar{h}} \rangle - a(u_h, z_{\bar{h}})$ gilt:*

$$|\langle J, u \rangle - \langle J, u_h \rangle| \leq C(\eta_p \eta_d + |\theta(u_h, z_{\bar{h}})|) \quad (\text{IV.51})$$

$$|\langle J, u \rangle - (\langle J, u_h \rangle + \theta(u_h, z_{\bar{h}}))| \leq \bar{C} \eta_p \eta_d. \quad (\text{IV.52})$$

Der Fehlerschätzer (IV.51) entspricht der gestellten Zielsetzung zur Kontrolle des Ausdrucks (IV.47). Dagegen wird durch den Fehlerschätzer (IV.52) der Fehler von $\langle J, u \rangle$ und einem durch $\theta(u_h, z_{\bar{h}})$ korrigierten Ausdruck von $\langle J, u_h \rangle$ kontrolliert. Dieses Vorgehen ist insofern vorteilhaft, da in der Regel weniger u sondern vielmehr $\langle J, u \rangle$ approximiert werden soll, und sich der

¹⁵vgl. z.B. Lemma 3.6 in [57]

Ausdruck $\langle J, u_h \rangle + \theta(u_h)$ diesbezüglich im Fall $h \approx \tilde{h}$ durch eine doppelte Approximationsgenauigkeit auszeichnet.

Der zusätzliche Fehler- bzw. Korrekturterm $\theta(u_h, z_{\tilde{h}})$ ist auf die allgemeine Voraussetzung zurückzuführen, dass $V_h \neq V_{\tilde{h}}$ ist. Bei Gleichheit erhält man wegen $z_{\tilde{h}} \in V_h$ unmittelbar:¹⁶

Korollar IV.4 *Es seien $u \in V$ und $u_h \in V_h$ Lösungen von (IV.45) bzw. (IV.46) sowie $z \in V$ und $z_{\tilde{h}} \in V_{\tilde{h}}$ Lösungen von (IV.48) bzw. (IV.49) mit $V_h = V_{\tilde{h}}$. Es seien $\eta_p, \eta_d \in \mathbb{R}_{\geq 0}$ Fehlerschätzer des primalen bzw. dualen Problems. Dann existiert ein $C > 0$, so dass gilt:*

$$|\langle J, u \rangle - \langle J, u_h \rangle| \leq C\eta_p\eta_d.$$

Die Wahl $V_h = V_{\tilde{h}}$ ist ohne weitere Maßnahmen in der Regel unangemessen. Eine geeignete (adaptive) Wahl von $V_{\tilde{h}}$ zur Approximation der dualen Lösung muss nicht notwendigerweise entsprechend geeignet für die Approximation des primalen Problems sein. Umgekehrt muss ein für die primale Lösung geeigneter Raum V_h nicht notwendigerweise auch für die duale Lösung passend sein. Zur Konstruktion eines Raumes V_h , der sowohl den Regularitätseigenschaften der primalen als auch der dualen Lösung Rechnung trägt, sei auf Abschnitt IV.7 verwiesen.

IV.6.2 Fehlerschätzer vom Goal-Oriented-Typ für Modellprobleme

Es seien $u \in H^1(\Omega, \Gamma_0)$ und $u_h \in S^p(\mathcal{T}_h)$ Lösung von (II.6) bzw. (III.13). Das Fehlermaß $J \in H^1(\Omega, \Gamma_0)'$ sei definiert als $\langle J, v \rangle := (\chi, v)_0 + (\psi, \gamma(v))_{0, \Gamma_1}$ mit $\chi \in L^2(\Omega)$ und $\psi \in L^2(\Gamma_1)$. Es seien $z \in H^1(\Omega, \Gamma_0)$ und $z_{\tilde{h}} \in S^{\tilde{p}}(\tilde{\mathcal{T}}_{\tilde{h}})$, so dass gilt:

$$\forall v \in H^1(\Omega, \Gamma_0) : (\nabla v, \nabla z)_0 = (\chi, v)_0 + (\psi, \gamma(v))_{0, \Gamma_1}, \quad (\text{IV.53})$$

$$\forall v_{\tilde{h}} \in S^{\tilde{p}}(\tilde{\mathcal{T}}_{\tilde{h}}, \Gamma_0) : (\nabla v_{\tilde{h}}, \nabla z_{\tilde{h}})_0 = (\chi, v_{\tilde{h}})_0 + (\psi, \gamma(v_{\tilde{h}}))_{0, \Gamma_1}. \quad (\text{IV.54})$$

Aus Satz IV.5 folgt, dass η_p mit

$$\eta_p^2 := \sum_{T \in \mathcal{T}_h} \left((h_T/p_T)^2 R_T^2 + \sum_{E \in \mathcal{E}_T} (h_E/p_E) R_E^2 \right)$$

¹⁶vgl. Ch.8 in [2]

und den Residuen R_T und R_E aus (IV.20) ein Fehlerschätzer für das primale Problem ist. Einen Fehlerschätzer für das duale Problem erhält man mit

$$\eta_d^2 := \sum_{T \in \tilde{\mathcal{T}}_h} \left((\tilde{h}_T / \tilde{p}_T)^2 \tilde{R}_T^2 + \sum_{E \in \mathcal{E}_T} (\tilde{h}_E / \tilde{p}_E) \tilde{R}_E^2 \right)$$

$$\tilde{R}_T := \|\chi + \Delta z_{\tilde{h}}\|_{0,T}, \quad \tilde{R}_E := \begin{cases} \frac{1}{2} \|\llbracket \partial_n z_{\tilde{h}} \rrbracket\|_{0,E}, & E \in \tilde{\mathcal{E}}^\circ \\ \|\psi - \partial_n z_{\tilde{h}}\|_{0,E}, & E \in \tilde{\mathcal{E}}_1, \end{cases}$$

wobei $\tilde{\mathcal{E}}^\circ$ die Menge der inneren Kanten von $\tilde{\mathcal{T}}_h$ bezeichnet, und $\tilde{\mathcal{E}}_1$ die Menge der Kanten auf Γ_1 . Mit $\theta(u_h, z_{\tilde{h}}) := (f, z_{\tilde{h}})_0 + (q, \gamma(z_{\tilde{h}}))_{0,\Gamma_1} - (\nabla u_h, \nabla z_{\tilde{h}})_0$ folgen dann die in Satz IV.14 angegebenen Fehlerschätzer (IV.51) und (IV.52).

Eine weitere Möglichkeit zur Fehlerkontrolle besteht darin, den Ausdruck $|\theta(u_h, z_{\tilde{h}})|$ in (IV.51) weiter abzuschätzen: Es folgt aus der Galerkin-Orthogonalität (IV.17) und partieller Integration mit einem Interpolationsoperator $I : H^1(\Omega, \Gamma_0) \rightarrow S^p(\mathcal{T}_h)$

$$\begin{aligned} & (f, z_{\tilde{h}})_0 + (q, \gamma(z_{\tilde{h}}))_{0,\Gamma_1} - (\nabla u_h, \nabla z_{\tilde{h}})_0 \\ &= (f, z_{\tilde{h}} - I(z_{\tilde{h}}))_0 + (q, \gamma(z_{\tilde{h}} - I(z_{\tilde{h}})))_{0,\Gamma_1} - (\nabla u_h, \nabla(z_{\tilde{h}} - I(z_{\tilde{h}})))_0 \\ &= \sum_{T \in \mathcal{T}_h} (f + \Delta u_h, z_{\tilde{h}} - I(z_{\tilde{h}}))_{0,T} + \sum_{E \in \mathcal{E}_h^\circ} ([\partial_n u_h], z_{\tilde{h}} - I(z_{\tilde{h}}))_{0,E} \\ & \quad + \sum_{E \in \mathcal{E}_{1,h}} (q - \partial_n u_h, z_{\tilde{h}} - I(z_{\tilde{h}}))_{0,E}. \end{aligned}$$

Die Fehlerkontrolle (IV.51) führt damit auf den Fehlerschätzer

$$\eta_{DWR,1} := \eta_p \eta_d + \sum_{T \in \mathcal{T}_h} \left((f + \Delta u_h, z_{\tilde{h}} - I(z_{\tilde{h}}))_{0,T} + \sum_{E \in \mathcal{E}_T} ([\partial_n u_h], z_{\tilde{h}} - I(z_{\tilde{h}}))_0 \right), \quad (\text{IV.55})$$

der einer Variante des dual-gewichteten residualen Fehlerschätzers (DWR-Fehlerschätzer) entspricht.¹⁷ Der Anteil $\eta_p \eta_d$ in (IV.55) wird im Allgemeinen vernachlässigt, da er von höherer Ordnung ist.

¹⁷vgl. z.B. [19], [21] und [22]

Alternativ hierzu erhält man nach Anwendung der Cauchy-Schwarzschen Ungleichung auf (IV.55) mit

$$\eta_{DWR,2} := \sum_{T \in \mathcal{T}_h} \left(\varrho_T (h_T/p_T)^2 R_T + \sum_{E \in \mathcal{E}_T} \varrho_E (h_E/p_E) R_E \right) \quad (\text{IV.56})$$

und den dualen Gewichten

$$\varrho_T := (h_T/p_T)^{-2} \|z_{\tilde{h}} - I(z_{\tilde{h}})\|_{0,T}, \quad \varrho_E := (h_E/p_E)^{-1} \|z_{\tilde{h}} - I(z_{\tilde{h}})\|_{0,E}$$

einen weiteren dual-gewichteten residualen Fehlerschätzer.

IV.6.3 Ein allgemeiner Ansatz für elliptische Minimierungsprobleme erster und zweiter Art

Für allgemeine elliptische Minimierungsprobleme erster und zweiter Art kann in ähnlicher Weise wie in Abschnitt IV.6.1 vorgegangen werden. Hierzu seien U ein normierter Raum, $\beta \in L(V, U)$ und $\lambda, \lambda_H \in U'$ sowie $u \in V$ bzw. $u_h \in V_h$ Lösungen von

$$\forall v \in V : a(u, v) = \langle \ell, v \rangle - \langle \lambda, \beta(v) \rangle, \quad (\text{IV.57})$$

$$\forall v_h \in V_h : a(u_h, v) = \langle \ell, v_h \rangle - \langle \lambda_H, \beta(v_h) \rangle. \quad (\text{IV.58})$$

Dann gilt:

$$\begin{aligned} & \langle J, u \rangle - \langle J, u_h \rangle \\ &= a(u - u_h, z) = a(u - u_h, z - z_{\tilde{h}}) + a(u - u_h, z_{\tilde{h}}) \\ &= a(u - u_h, z - z_{\tilde{h}}) + \langle \ell, z_{\tilde{h}} \rangle - \langle \lambda, \beta(z_{\tilde{h}}) \rangle - a(u_h, z_{\tilde{h}}) \\ &= a(u - u_h, z - z_{\tilde{h}}) + \langle \lambda_H - \lambda, \beta(z_{\tilde{h}}) \rangle + \langle \ell, z_{\tilde{h}} \rangle - \langle \lambda_H, \beta(z_{\tilde{h}}) \rangle - a(u_h, z_{\tilde{h}}). \end{aligned} \quad (\text{IV.59})$$

Analog zu Satz IV.14 erhält man aus der Stetigkeit von a und aus (IV.59):

Satz IV.15 *Es seien $u \in V$ und $u_h \in V_h$ Lösungen von (IV.57) bzw. (IV.58), sowie $z \in V$ und $z_{\tilde{h}}$ Lösungen von (IV.48) bzw. (IV.49). Es seien $\eta_p, \eta_d, \eta_l \in \mathbb{R}_{\geq 0}$ Fehlerschätzer des primalen bzw. dualen Problems sowie von $\|\lambda - \lambda_H\|$. Dann existieren $C, \bar{C} > 0$, so dass mit $\theta(u_h, z_{\tilde{h}}, \lambda_H) := \langle \ell, z_{\tilde{h}} \rangle - \langle \lambda_H, \beta(z_{\tilde{h}}) \rangle - a(u_h, z_{\tilde{h}})$ gilt:*

$$|\langle J, u \rangle - \langle J, u_h \rangle| \leq C(\eta_p \eta_d + \eta_l \|\beta(z_{\tilde{h}})\| + |\theta(u_h, z_{\tilde{h}}, \lambda_H)|), \quad (\text{IV.60})$$

$$|\langle J, u \rangle - (\langle J, u_h \rangle + \theta(u_h, z_{\tilde{h}}, \lambda_H))| \leq C(\eta_p \eta_d + \eta_l \|\beta(z_{\tilde{h}})\|). \quad (\text{IV.61})$$

Aus Satz IV.15 erhält man zusammen mit den Fehlerschätzern aus den Abschnitten IV.1.1, IV.1.2 und IV.1.3 unmittelbar Fehlerschätzer für elliptische Minimierungsprobleme erster und zweiter Art. Für elliptische Minimierungsprobleme erster Art setzt man $U := U_0$, $\beta := \beta_0$, $\lambda := \lambda_0$ und $\lambda_H := \lambda_{0,H}$. Ähnlich gewinnt man Fehlerschätzer für elliptische Minimierungsprobleme zweiter Art. Im unrestringierten Fall setzt man $U := U_1$, $\beta := \beta_1$, $\lambda := \lambda_1$ und $\lambda_H := \lambda_{1,H}$, sowie im restringierten Fall $U := U_0 \times U_1$, $\beta := (\beta_0, \beta_1)$, $\langle \lambda, \beta(v) \rangle := \langle \lambda_0, \beta_0(v) \rangle + \langle \lambda_1, \beta_1(v) \rangle$ und $\langle \lambda_H, \beta(v) \rangle := \langle \lambda_{0,H}, \beta_0(v) \rangle + \langle \lambda_{1,H}, \beta_1(v) \rangle$.

Die Fehlerschätzer aus Satz IV.15 entsprechen im Wesentlichen den Methoden zur Fehlerkontrolle in [96]. Insgesamt besteht das Problem dieses Vorgehens jedoch darin, einen Fehlerschätzer η_l für $\|\lambda - \lambda_H\|$ anzugeben. Nach den Sätzen IV.1, IV.2 und IV.3 ist η_p auch ein Fehlerschätzer von $\|\lambda - \lambda_H\|$. Der Fehlerschätzer (IV.61) erhält damit die Form:

$$|\langle J, u \rangle - (\langle J, u_h \rangle + \theta(u_h, z_{\tilde{h}}, \lambda_H))| \leq C\eta_p(\eta_d + \|\beta(z_{\tilde{h}})\|).$$

Im Vergleich zu (IV.52) geht hier die Verdoppelung der Genauigkeit des korrigierten Fehlermaßes $\langle J, u_h \rangle + \theta(u_h, z_{\tilde{h}}, \lambda_H)$ verloren. Dieses Vorgehen ist demnach offenbar nicht geeignet. In [96] wird dagegen die Verwendung von Average-Techniken vorgeschlagen, wie sie zum Beispiel in [32] und [33] beschrieben werden.

Erforderlich ist in jedem Fall eine Abschätzung von $\|\lambda - \lambda_H\|$, so dass $\eta_l\|\beta(z_{\tilde{h}})\|$ und $\eta_p\eta_d$ ungefähr von gleicher Ordnung sind.

IV.7 h - und hp -adaptive Verfeinerungen

Wie in den vorherigen Abschnitten dargestellt und etwa in den Abbildungen IV.5, IV.8 und IV.10 illustriert, führt die Verwendung höherer Ansätze bei elliptischen Minimierungsproblemen erster und zweiter Art im Allgemeinen nicht zu einem Gewinn höherer Konvergenzordnungen. So erhält man, sofern stückweise bilineare Ansatzfunktionen benutzt werden, für das in Abschnitt IV.3.1 untersuchte modellhafte Signorini-Problem unter Verwendung von h -Verfeinerungen eine Konvergenzordnung von $\mathcal{O}(h)$ gemessen in der H^1 -Norm. Dagegen ist die Konvergenzordnung bei Anwendung von stückweise quadratischen Ansatzfunktionen nicht mehr als $\mathcal{O}(h^{3/2})$. Wird von noch höheren Polynomgraden Gebrauch gemacht, wird kein weiterer Konvergenzgewinn erzielt. Auch bei dem modellhaften Reibungsproblem aus Abschnitt IV.3.2 und dem Hindernis-Problem aus Abschnitt IV.5.1 erhält man keine

höheren Konvergenzordnungen als $\mathcal{O}(h^{3/2})$.

Der Grund hierfür besteht darin, dass die Lösung nicht über die erforderliche Regularität verfügt. Diesbezüglich ist insbesondere auch die Verwendung reiner p -Verfeinerungen nicht zweckmäßig.

Ähnliche Phänomene lassen sich auch bereits bei unrestringierten elliptischen Minimierungsproblemen erster Art bzw. bei Variationsgleichungen beobachten. Weist das Grundgebiet Ω etwa eine einspringende Ecke auf, so ist die Lösung im Allgemeinen nicht H^2 -regulär. Schon bei der Verwendung von stückweise bilinearen Ansätzen ist die Konvergenzordnung $\mathcal{O}(h)$ nicht mehr zu erwarten. Dementsprechend ist der Gebrauch höherer Polynome bei h -Verfeinerungen nicht sinnvoll.

In allen drei Fällen ist die Ursache für die mangelnde Regularität der Lösung auf lokale Einflussgrößen innerhalb der Problemformulierung zurückzuführen. Diese lokalen Einflussgrößen sind bei Signorini-Problemen und bei Hindernisproblemen die Hindernisnebenbedingungen, bei Reibungsproblemen das in das Zielfunktional eingehende nicht-differenzierbare Funktional j und bei unrestringierten Problemen zum Beispiel einspringende Ecken oder ein unstetiger Wechsel in den Randdaten.

Die Frage ist nun, ob durch eine adäquate Anpassung des FE-Gitters bezüglich dieser lokalen Einflussgrößen eine dem Polynomgrad entsprechende Konvergenzordnung¹⁸ erreicht werden kann. Offenbar besteht eine sinnvolle Vorgehensweise darin, die jeweiligen Vorteile von h - und p -Verfeinerungen zu nutzen: In Bereichen hoher Regularität werden hohe lokale Polynomgrade gewählt, während in Bereichen mit geringer Regularität eher lokale Verfeinerungen bevorzugt werden. Gemäß Abschnitt III.1.1 können etwa bei Eckenproblemen FE-Gitter a-priori konstruiert werden, die bei Verwendung von h -Verfeinerungen und festem Polynomgrad p Konvergenzordnungen der Güte $\mathcal{O}(h^p)$ (Satz III.4) oder bei hp -Verfeinerungen exponentielle Konvergenz (Satz III.5) sichern.

Im Allgemeinen Fall ist jedoch im Voraus nicht bekannt, wo und in welcher Weise das FE-Gitter zu adaptieren ist, damit eine gewünschte Konvergenzordnung erreicht wird.

Strategien, die für dieses Adaptierungsproblem etabliert sind, bestehen üblicherweise darin, mit Hilfe eines Fehlerschätzers Bereiche großer Fehleranteile zu lokalisieren und in diesen Bereichen entweder die Maschenweite zu verringern oder die lokalen Polynomgrade anzupassen. Für die Lokalisierung

¹⁸Der Begriff Konvergenzordnung ist auch im Folgenden in Bezug auf die Anzahl der verwendeten Freiheitsgrade zu verstehen.

kommen nur Fehlerschätzer infrage, die Rückschlüsse auf die lokale Fehlerverteilung zulassen; etwa Fehlerschätzer η , die sich aus lokalen Fehleranteilen in der Form $\eta^2 = \sum_{T \in \mathcal{T}_h} \eta_T^2$ zusammensetzen. Die in den Abschnitten zuvor beschriebenen residualen Fehlerschätzer besitzen diese Eigenschaft und können deshalb eingesetzt werden. Auf Extrapolationsansätzen basierende Fehlerschätzer, wie sie zum Beispiel in [2], [99] oder [100] beschrieben werden, kommen für Adaptierungsstrategien dieser Art nicht in Frage.

Entscheidend für den Einsatz von Fehlerschätzern in Bezug auf Adaptierungsstrategien ist, dass die Fehlerverteilung richtig wiedergegeben wird. Die zahlenmäßige Erfassung des Fehlers spielt hierbei keine Rolle. Dementsprechend ist die Einbeziehung von generischen und deshalb unbekanntem Konstanten, wie sie in den Fehlerschätzern der vorangegangenen Abschnitte vorkommen, kein Nachteil, solange die eingehenden Konstanten nicht von der lokalen Maschenweite bzw. dem lokalen Polynomgrad abhängen.

Adaptierungsstrategien können im Wesentlichen in zwei Gruppen aufgeteilt werden: h -adaptive Strategien beschränken sich allein auf die Anpassung der Maschenweite, der Polynomgrad bleibt fest. Dagegen werden bei hp -adaptiven Strategien zusätzlich die lokalen Polynomgrade angepasst.

Sowohl h - als auch hp -Strategien werden in der Literatur in unterschiedlichen Varianten beschrieben. Für einen Überblick sei etwa auf [7], [84] oder [85] verwiesen. Gemeinsam ist nahezu allen Strategien, dass zunächst die Fehlerverteilung einer FE-Lösung mit Hilfe eines Fehlerschätzers bestimmt wird, dann das FE-Gitter bzw. die lokalen Polynomgrade angepasst werden und schließlich eine neue FE-Berechnung durchgeführt wird. Dieser Vorgang wird wiederholt, bis eine gewisse Fehlertoleranz (unter Berücksichtigung oder ggf. Vernachlässigung der Konstante C) erreicht wird oder die zur Verfügung stehenden Rechnerkapazitäten ausgeschöpft sind. Das Resultat ist in der Regel ein FE-Gitter mit einer kleinen Maschenweite in der Nähe der störenden lokalen Einflussgrößen und mit hohen Polynomgraden außerhalb hiervon. Die Maschenweite kann entweder durch Zerlegung der Gitterelemente, Verschieben der Gitterkoordinaten (r -Methode) oder durch Neuvernetzung (aufgrund einer vorgegebenen Knotendichte-Funktion) adaptiert werden.

Im Folgenden wird sowohl eine h -adaptive als auch hp -adaptive Strategie vorgestellt und auf die bisher beschriebenen Problemklassen elliptischer Minimierungsprobleme erster und zweiter Art angewandt. Hierzu werden die in den Abschnitten IV.3 und IV.4 dargestellten Fehlerschätzer verwendet. Wesentlich für die Fehlerkontrolle elliptischer Minimierungsprobleme ist die Nutzung von Fehlerschätzern für Variationsgleichungen, wie sie in Abschnitt

IV.2 vorgestellt werden. Aus diesem Grund wird der allgemeinen Betrachtung eine Untersuchung der Adaptierungsstrategien für Variationsgleichungen an dem Beispiel (III.16) vorangestellt, da hiermit ihre Anwendbarkeit gut demonstriert werden kann. Weitere Beispiele sind im Abschnitt A.1 enthalten.

Darüber hinaus werden Anwendungsmöglichkeiten bei Fehlerbetrachtungen vom Goal-oriented-Typ diskutiert.

Das wesentliche Resultat bei den folgenden Betrachtungen ist, dass die Verwendung der hier beschriebenen Adaptierungsstrategien in der Regel zu den gewünschten Konvergenzsteigerungen führt.

IV.7.1 h -adaptive Verfeinerungsstrategien

h -adaptive Strategien beruhen im Allgemeinen darauf, dass mit Hilfe eines Fehlerschätzers η mit der Darstellung $\eta^2 = \sum_{T \in \mathcal{T}_h} \eta_T^2$ diejenigen Gitterelemente bestimmt werden, die den größten Beitrag zum Gesamtfehler liefern. Zur Verringerung der Maschenweite werden diese Elemente dann zerlegt, wobei das hieraus resultierende verfeinerte Gitter im nächsten Iterationsschritt den zugrunde liegenden FE-Raum V_h definiert. Alternativ kann aber auch eine Neuvernetzung mit einer höheren Elementdichte im Bereich dieser Gitterelemente vorgenommen werden. Für die Auswahl \mathcal{K} der mit den größten Fehleranteilen behafteten Gitterelemente können unterschiedliche Kriterien angewandt werden.¹⁹

Ein einfaches Kriterium, das hier exemplarisch aufgeführt wird, ist das sogenannte Fixed-Fraction-Kriterium: In jedem Iterationsschritt wird ein vorher festgelegter, fester prozentualer Anteil $\kappa \in [0, 1]$ der Gitterelemente zerlegt. Die sich daraus ergebene h -adaptive Strategie kann mit $k^* := \lfloor (1 - \kappa)|\mathcal{T}_h| \rfloor + 1$ wie folgt beschrieben werden:

- (i) Ermittlung der η_T für alle $T \in \mathcal{T}_h$.
- (ii) Aufsteigende Sortierung zu $\eta_{T_1} \leq \eta_{T_2} \leq \dots$
- (iii) $\mathcal{K} := \{T \in \mathcal{T}_h \mid \eta_T \geq \eta_{T_{k^*}}\}$.
- (iv) Zerlegung aller $T \in \mathcal{K}$.

Zur Vermeidung k -irregulärer Knoten, Kanten bzw. Facetten kann nach der Verfeinerung eine Regularisierung des Gitters vorgenommen werden, etwa indem solange Gitterelemente, die k -irreguläre Kanten oder Facetten enthalten, zerlegt werden, bis diese innerhalb des Gitters nicht mehr vorkommen.

¹⁹vgl. [73], S. 178

IV.7.2 hp -adaptive Verfeinerungsstrategien

Bei hp -adaptiven Strategien muss über die Entscheidung hinaus, welche Elemente zerlegt werden sollen, zudem noch entschieden werden, welche Elemente mit einem höheren Polynomgrad versehen werden sollen. Wie bei h -adaptiven Strategien sind auch hier unterschiedliche Vorgehensweisen möglich. Hierzu sei etwa auf [3], [42], [81] und [85] verwiesen. Die wesentliche Idee zur Wahl des lokalen Polynomgrads besteht in den meisten Fällen in dem Versuch, die lokale Regularität der Lösung mit Hilfe eines Fehlerschätzers abzuschätzen. Wird die lokale Regularität für ein Gitterelement als hoch eingeschätzt, lohnt sich die Erhöhung des Polynomgrads, bei geringer lokaler Regularität ist die Zerlegung des Elements sinnvoll.

Im Folgenden wird eine Strategie vorgestellt, die auf einer Abschätzung der lokalen Regularität aufgrund von Fehlerschätzungen zweier aufeinander folgender FE-Berechnungen basiert und die sich im Wesentlichen an die hp -adaptive Strategie aus [93] orientiert, jedoch den FE-Raum zur Ermittlung der lokalen Regularität in einer verfeinerten Variante wählt.

Von zentraler Bedeutung bei dieser Strategie ist die Annahme, dass sich die lokalen Fehleranteile η_T für Gitterelemente T ungefähr wie

$$\eta_T \approx C_T p_T^{-\varrho_T + 1}$$

verhalten. Sind zwei Fehleranteile η_T^0 und η_T^1 für p_T bzw. $p_T + 1$ bekannt, kann ϱ_T als

$$\varrho_T = \frac{\log(\eta_T^1/\eta_T^0)}{\log(p_T/(p_T + 1))} + 1$$

berechnet werden. Der Parameter ϱ_T ist hier als ein Maß für die lokale Regularität interpretierbar. Ist die Lösung bezüglich des Gitterelements T hinreichend regulär, muss ϱ_T größer als $p_T + 1$ sein.

Der erste Schritt der im Folgenden beschriebenen hp -adaptiven Strategie besteht wie bei der oben vorgestellten h -adaptiven Strategie darin, die Gitterelemente mit den größten lokalen Fehleranteilen η_T^0 in einer Auswahl \mathcal{K} zu erfassen. Auch hier kann diesbezüglich wie oben ebenfalls das Fixed-Fraction-Kriterium für ein $\kappa \in [0, 1]$ angewandt werden. Der lokale Polynomgrad der in \mathcal{K} erfassten Gitterelemente wird um eins erhöht. Damit der volle lokale Polynomraum auf den Gitterelementen mit erhöhtem lokalen Polynomgrad zur Verfügung steht und nicht durch eventuell erforderliche Stetigkeitsanpassungen reduziert wird, wird ebenfalls der lokale Polynomgrad von denjenigen Elementen um eins erhöht, die zu Elementen aus \mathcal{K} über eine

Kante oder eine Facette benachbart sind und die nicht zu \mathcal{K} gehören. Diese Elemente werden in der Menge \mathcal{W} zusammengefasst. Anschließend wird der Fehler aufgrund einer weiteren FE-Berechnung durch die Fehleranteile η_T^1 erneut geschätzt. Im nächsten Schritt wird die ursprüngliche Verteilung der Polynomgrade wiederhergestellt, der Polynomgrad der Elemente $T \in \mathcal{K} \cup \mathcal{W}$ wird um eins verringert.

Mit den nun zur Verfügung stehenden lokalen Fehleranteilen η_T^0 und η_T^1 kann entschieden werden, ob eine Zerlegung eines Elements $T \in \mathcal{K}$ oder eine Erhöhung des Polynomgrads p_T vorgenommen werden soll: Ein Element $T \in \mathcal{K}$ wird zerlegt, falls die geschätzte lokale Regularität ϱ_T echt kleiner als $p_T + 1$ ist. Die neuen Elemente erhalten den Polynomgrad p_T . Dagegen wird der lokale Polynomgrad von T wieder erhöht, wenn $\varrho_T \geq p_T + 1$ ist.

Diese Strategie kann insgesamt wie folgt zusammengefasst werden:

- (i) Ermittlung der η_T^0 für alle $T \in \mathcal{T}_h$.
- (ii) Aufsteigende Sortierung zu $\eta_{T_1}^0 \leq \eta_{T_2}^0 \leq \dots$
- (iii) $\mathcal{K} := \{T \in \mathcal{T}_h \mid \eta_T^0 \geq \eta_{T_{k^*}}^0\}$.
- (iv) $\mathcal{W} := \{T \in \mathcal{T}_h \setminus \mathcal{K} \mid \exists T_0 \in \mathcal{K} : \emptyset \neq \overline{T} \cap \overline{T_0} \notin \mathcal{V}\}$.
- (v) $p_T := p_T + 1$, falls $T \in \mathcal{K} \cup \mathcal{W}$.
- (vi) Ermittlung der η_T^1 sowie ϱ_T für alle $T \in \mathcal{K}$.
- (vii) $p_T := p_T - 1$ für alle $T \in \mathcal{K} \cup \mathcal{W}$.
- (viii) Zerlegung von $T \in \mathcal{K}$, falls $\varrho_T < p_T + 1$.
- (ix) $p_T := p_T + 1$ für alle $T \in \mathcal{K}$, falls $\varrho_T \geq p_T + 1$

In den Abbildungen IV.13 und IV.14 werden die einzelnen Schritte der oben aufgeführten hp -Strategie an einem Beispiel mit $\kappa = 0.2$ demonstriert. Dabei sind die Schritte (i)-(v) in Abbildung IV.13 zu sehen, und die Schritte (vi)-(ix) in Abbildung IV.14.

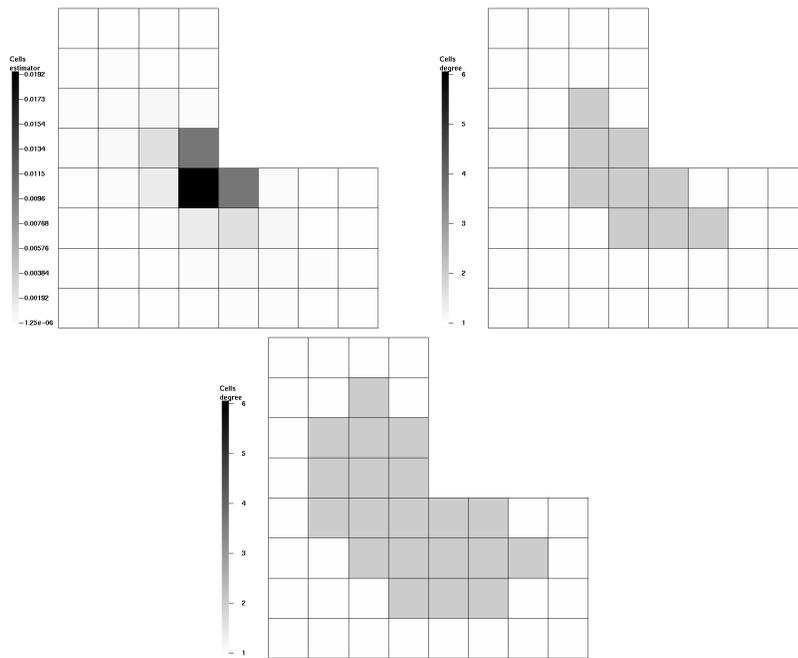


Abb. IV.13: Berechnung der η_T^0 , Ermittlung der Menge \mathcal{K} , Erhöhung des lokalen Polynomgrads für Elemente $T \in \mathcal{W}$.

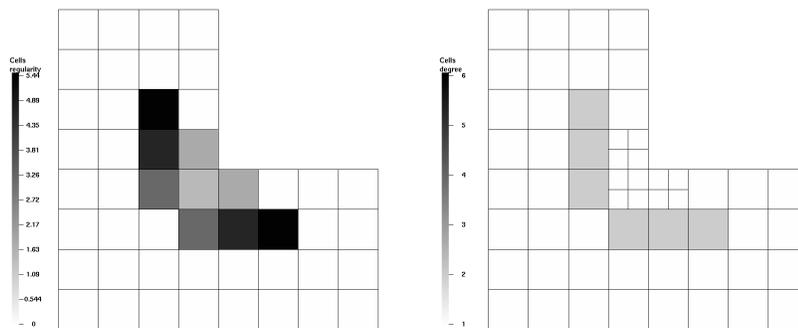


Abb. IV.14: Geschätzte lokale Regularität für Elemente $T \in \mathcal{K}$, *hp*-Adaptierung des FE-Gitters.

Sollen k -irreguläre Kanten bzw. Facetten vermieden werden, kann auch bei hp -adaptiven Strategien wie oben beschrieben eine Regularisierung des FE-Gitters vorgenommen werden. Außerdem ist es in der Regel günstiger, dass die Differenz lokaler Polynomgrade von über eine Kante oder Facette benachbarten Elemente nicht allzu groß ist, was gegebenenfalls weitere Anpassungen der Verteilung der lokalen Polynomgrade erforderlich macht.

Im Unterschied zu den meisten hp -adaptiven Strategien der oben referenzierten Arbeiten zeichnet sich die vorgestellte Strategie vor allem durch einen relativ vorsichtigen Umgang in Bezug auf die Ermittlung der beiden Fehleranteile η_T^0 und η_T^1 aus. Im Gegensatz hierzu werden diese Anteile z.B. in [80] ohne einen speziellen Zwischenschritt berechnet, in dem nur p -Verfeinerungen vorgenommen werden. Ob die dabei resultierenden Fehleranteile überhaupt in Beziehung miteinander stehen, ist fraglich, da die in jedem Schritt vorgenommenen Adaptierungsschritte sicherlich nicht ohne Einfluss auf die Fehleranteile bleiben. Tatsächlich sind die in [80] vorgestellten hp -adaptiven Gitter deutlich nicht optimal.

Das Hauptproblem der hier vorgestellten h - und hp -adaptiven Strategien sowie beinahe aller in der Literatur bekannten Strategien zur Verfeinerung besteht darin, dass diese letztlich auf heuristischen Annahmen beruhen. Problematisch ist vor allem, dass die lokalen Fehleranteile η_T des Fehlerschätzers η im strengen Sinne nicht notwendigerweise die lokalen Fehler wiedergeben müssen. Die Anwendung der Strategien führt zwar in den meisten Fällen zu fast optimalen Gittern bzw. zu Gittern, die eine geforderte Konvergenz sichern, jedoch ist die Optimalität der FE-Gitter im Allgemeinen unbewiesen. Ein rigoroser Nachweis der Optimalität der erzeugten Gitter ist in der Regel nur unter sehr eingeschränkten Bedingungen möglich, wie zum Beispiel bei Beschränkung auf eindimensionale Probleme. Hierzu sei etwa auf [15], [60] und [84] verwiesen.

IV.7.3 h - und hp -adaptive Verfeinerungen für Variationsgleichungen

Wie eingangs erwähnt, ist die Verwendung von Fehlerschätzern für Variationsgleichungen von zentraler Bedeutung für die in Abschnitt IV.1 bzw. IV.3, IV.4 und IV.5 entwickelten Fehlerschätzer für elliptische Minimierungsprobleme erster und zweiter Art. In Abschnitt IV.2 wird diesbezüglich ein residualer Fehlerschätzer, der auch für Finite-Elemente-Anwendungen mit höheren Polynomansätzen geeignet ist, vorgestellt und hinsichtlich seiner

Zuverlässigkeit über die Betrachtung des Effektivitätsindex an Beispielen studiert.

Im vorliegenden Abschnitt wird die Eignung des residualen Fehlerschätzers in Bezug auf die Anwendbarkeit von h - und hp -adaptiven Verfeinerungsstrategien untersucht. Im Vordergrund der Betrachtungen stehen numerische Experimente bzgl. des Modellproblems (II.6) mit der analytischen Lösung u wie in (III.16). Da u nicht H^2 -regulär ist, kann weder bei globalen h -Verfeinerungen eine Konvergenzordnung der Form $\mathcal{O}(h^p)$ noch bei reinen p -Verfeinerungen exponentielle Konvergenz erwartet werden. Anhand numerischer Beispiele wird untersucht, ob mit Hilfe von h - bzw. hp -adaptiven Strategien Konvergenzsteigerungen in angemessenen Größenordnungen erreicht werden können. Außerdem wird die Zuverlässigkeit des Fehlerschätzers mit Hilfe des Effektivitätsindex geprüft.

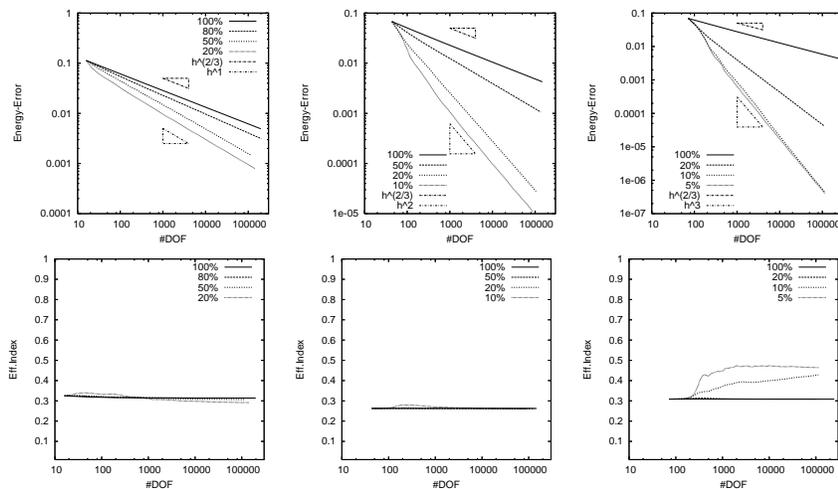


Abb. IV.15: Konvergenzverhalten bei h -adaptiver Verfeinerung (oben) und Effektivitätsindizes (unten) für $p = 1$ und $\kappa = 1.0, 0.8, 0.5, 0.2$ (links), $p = 2$ und $\kappa = 1.0, 0.5, 0.2, 0.1$ (mitte) und $p = 3$ und $\kappa = 1.0, 0.2, 0.1, 0.05$ (rechts).

In Abbildung IV.15 ist das Konvergenzverhalten bei h -adaptiven Verfeinerungen für die Polynomgrade $p = 1, 2, 3$ und verschiedenen Fixed-Fraction-Parametern κ zu sehen. Dass auch der Fehlerschätzer das richtige Konvergenzverhalten wiedergibt, ist anhand des rechts dargestellten Effektivitätsindex festzustellen. Dieser ist für alle Verfeinerungen im Wesentlichen kon-

stant. Zu beobachten ist, dass die Schwankungen des Effektivitätsindex abnehmen, je näher κ an 1.0 gewählt wird.

Schließlich ist zu bemerken, dass offenbar die geforderten Konvergenzordnungen $\mathcal{O}(h^p)$ für die hier dargestellten Polynomgrade erreicht werden. Insbesondere erkennt man, dass je kleiner κ gewählt wird, desto besser wird die gewünschte Konvergenzordnung erreicht. Der optimale Wert für κ ist jedoch abhängig von dem gewählten Polynomgrad. Wie aus den Abbildungen zu erkennen ist, ist dieser für $p = 1$ etwas kleiner als 0.5, für $p = 2$ und $p = 3$ etwas kleiner als 0.2 bzw. 0.1. Ein möglichst großer optimaler Wert für κ ist wünschenswert, da hiermit die Anzahl der Adaptierungsschritte gering bleibt.

Zusammenfassend ist aus den Abbildungen abzulesen, dass zwar ein hoher Fixed-Fraction-Parameter im besonderen Maße die Zuverlässigkeit der Fehlerschätzung sichert, dieser verhindert aber gleichzeitig die Optimalität der Konvergenzordnung.

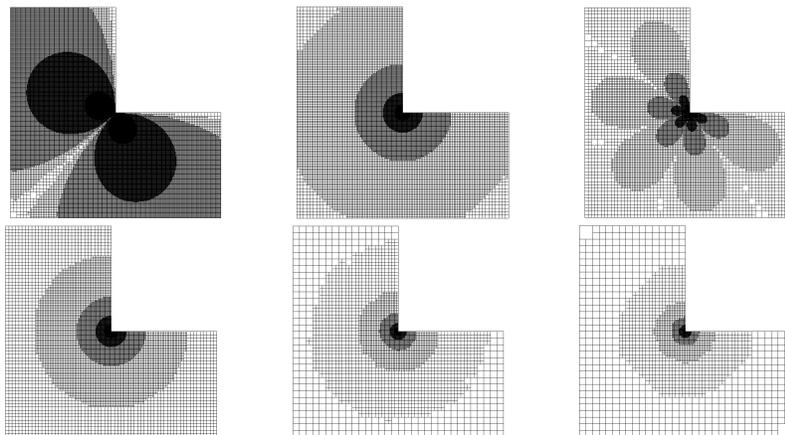


Abb. IV.16: h -adaptive Gitter mit optimaler Konvergenzordnung für $p = 1, \dots, 6$.

In der Abbildung IV.16 sind h -adaptive Gitter zu den Polynomgraden $p = 1, \dots, 6$ zu sehen, die gemäß Abbildung IV.15 sowie IV.17 eine jeweils optimale Konvergenzordnung $\mathcal{O}(h^p)$ gewährleisten. Auffällig sind die Gitterstrukturen zu $p = 1$ und $p = 3$. Die Gitter für die Polynomgrade $p = 2, 4, 5, 6$ haben kreisförmig angeordnete lokale Verfeinerungen mit der einspringenden Ecke als Zentrum und entsprechen damit den radialen Gittern aus Satz

III.4. Die Gitter für $p = 1$ und $p = 3$ weisen dagegen deutlich abweichende Strukturen auf.

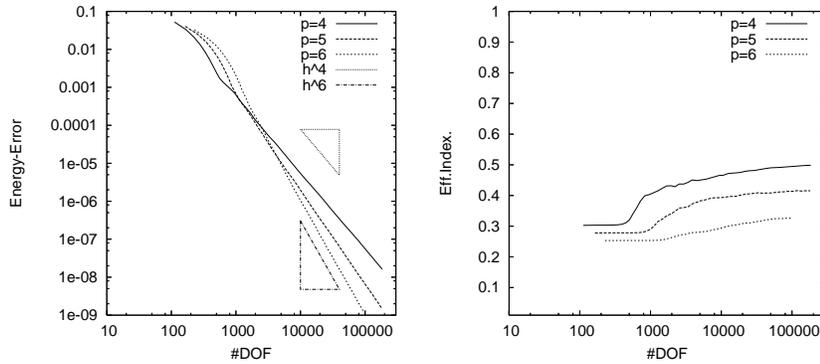


Abb. IV.17: Konvergenzverhalten bei h -adaptiver Verfeinerung, Effektivitätsindizes für $p = 4, 5, 6$.

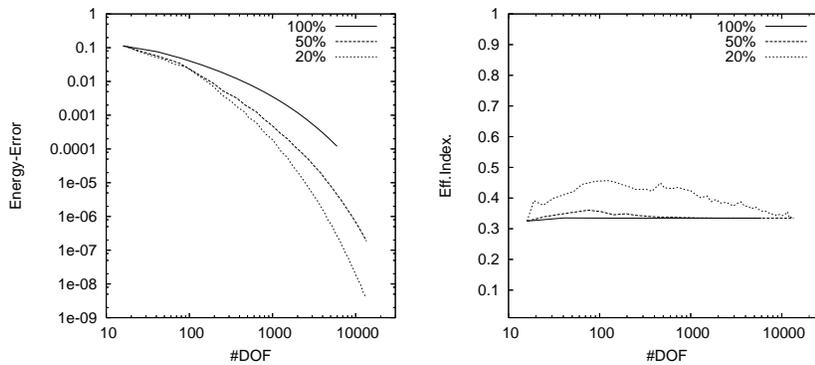


Abb. IV.18: Konvergenzverhalten bei hp -adaptiver Verfeinerung, Effektivitätsindizes für $\kappa = 1.0, 0.5, 0.2$.

In Abbildung IV.18 ist das Konvergenzverhalten bei der Verwendung hp -adaptiver Verfeinerungen dargestellt. Für die gewählten Fixed-Fraction-Parameter $\kappa \in \{0.2, 0.5, 1.0\}$ erhält man exponentielle Konvergenz. Wie bei der Anwendung von h -adaptiven Verfeinerungen ist in diesem Fall ebenfalls zu beobachten, dass kleine Fixed-Fraction-Parameter ein besseres

Konvergenzverhalten erzielen. Jedoch ist auch hier abzuwägen, dass kleine Parameter die Qualität der Fehlerschätzung mindern (vgl. Effektivitätsindex in Abbildung IV.18 (rechts)) und die Anzahl der erforderlichen Adaptierungsschritte erhöhen.

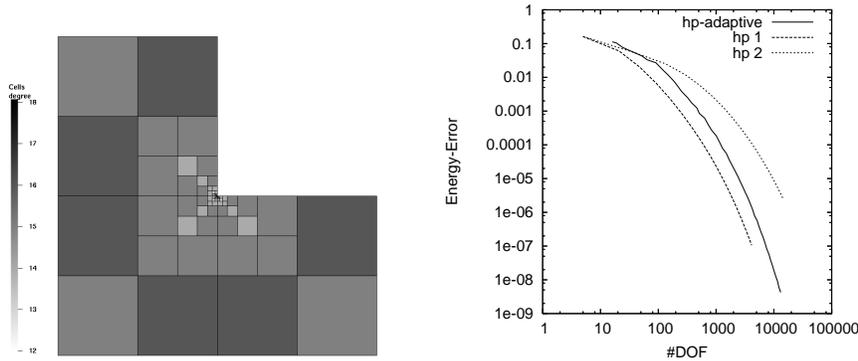


Abb. IV.19: hp -adaptives FE-Gitter, Vergleich von hp -adaptiven Verfeinerungen (vgl. Abb. III.3).

In Abbildung IV.19 (links) ist ein hp -adaptives FE-Gitter zu sehen, das mit einem Fixed-Fraction-Parameter von $\kappa = 0.2$ erzeugt wird. Die aus dem Verfeinerungsprozess resultierenden Gitter entsprechen exakt den Gittern der geometrischen Gitterfolgen aus Satz III.5. Dagegen weicht die Polynomgradverteilung deutlich von der a-priori vorgeschlagenen Verteilung ab. Bei letzterer steigt der lokale Polynomgrad ausgehend von der einspringenden Ecke schichtweise linear an (vgl. Abbildung III.4 (unten)), während sich der Anstieg bei der hp -adaptiven Verteilung über drei Schichten mit einem um eins verminderten Polynomgrad in den Ecken der ersten beiden Schichten vollzieht. Das FE-Gitter aus Satz III.5 erhält man als hp -adaptives FE-Gitter, wenn $\kappa = 1$ gewählt wird. Insgesamt wird deutlich, dass die FE-Gitter, die bei einem Fixed-Fraction-Parameter von $\kappa = 0.2$ erzeugt werden, ein wesentlich besseres Konvergenzverhalten als die üblicherweise in der Literatur vorgeschlagenen FE-Gitter bei gleichem Grading-Factor von $\sigma = 0.5$ liefern. Offenbar kompensiert eine günstigere Verteilung der lokalen Polynomgrade den eigentlich nicht optimal gewählten Grading-Factor, dessen Optimalwert bei etwa 0.15 liegt.

In Abbildung IV.19 (rechts) ist schließlich ein Vergleich der Konvergenzverhalten von FE-Gittern wie in Abbildung III.3, die mit dem laut Literatur op-

timalen Grading-Factor von $\sigma = 0.15$ erzeugt werden, und hp -adaptiven FE-Gittern zu sehen. Erkennbar ist, dass das Konvergenzverhalten der durch den hp -adaptiven Verfeinerungsprozess erzeugten FE-Gitter mindestens von der Güte der a-priori festgelegten FE-Gitter ist.

IV.7.4 h - und hp -adaptive Verfeinerungen mit Fehlerschätzern vom Goal-Oriented-Typ

In Abschnitt IV.6 werden Fehlerschätzer vom sogenannten Goal-Oriented-Typ vorgestellt, mit denen Fehlerkontrollen von lokalen, durch Funktionale $J \in V'$ beschriebenen Größen möglich sind. In diesem Abschnitt wird anhand numerischer Experimente untersucht, in welcher Weise die Anwendung der in Abschnitt IV.2 hergeleiteten Fehlerschätzer zusammen mit den h - und hp -adaptiven Strategien aus den Abschnitten IV.7.1 und IV.7.2 für Konvergenzverbesserungen in Bezug auf Variationsgleichungen genutzt werden können.

Neben Untersuchungen bezüglich Konvergenzverhalten und Gitteradaptierung wird auch die Entwicklung der Anzahl der Freiheitsgrade von primaler und dualer Diskretisierung diskutiert. Dieser Aspekt erscheint vor allem deshalb von Interesse, weil durch die Berechnung der dualen Lösung eine weitere FE-Berechnung durchgeführt werden muss, die vorzugsweise keinen höheren Aufwand als die FE-Berechnung für die primale Lösung verursachen sollte.

Wie im Abschnitt zuvor ist das primale Problem ebenfalls das durch (II.6) beschriebene Modellproblem, dem die Diskretisierung (III.13) zugeordnet ist. Als analytische Lösung u wird auch hier die Funktion aus (III.16) vorausgesetzt.

Das hierzu gewählte duale Problem entspricht der Formulierung (IV.53) mit der in (IV.54) angegebenen Diskretisierung. Als Fehlermaß wird hier exemplarisch das Funktional $\langle J, v \rangle := (\psi, v)_{0, \Gamma_1}$ mit

$$\psi(x_0, x_1) := \begin{cases} 1, & x_0 \in [-0.25, 0.25] \\ 0, & \text{sonst} \end{cases}$$

gewählt. Bei den folgenden Experimenten werden zunächst der Raum V_h für das primale Problem und der Raum V_h^c für das duale Problem getrennt voneinander mit Hilfe der oben beschriebenen Strategien adaptiert, um jeweils möglichst optimale Verfeinerungen zu gewinnen. Hierbei wird das Gitter bzw. die lokale Polynomgradverteilung sowohl für das primale Problem als

auch für das duale Problem in jedem Adaptierungsschritt angepasst.²⁰ Obwohl hierbei für das primale und das duale Problem unterschiedliche Gitter bzw. Polynomgradverteilungen erzeugt werden, zeigt sich, dass die Anzahl der verwendeten Freiheitsgrade im Wesentlichen gleich ist.

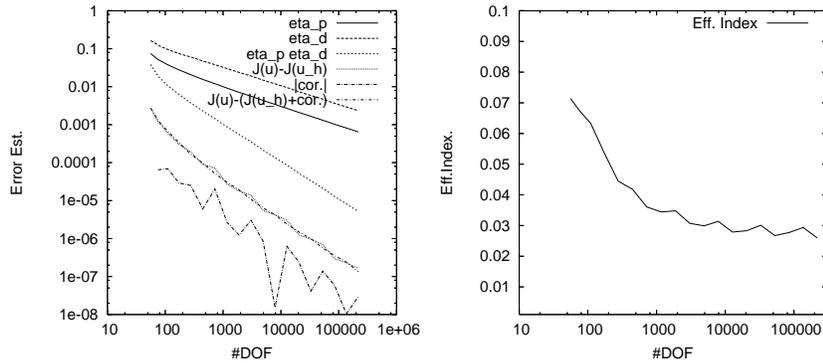


Abb. IV.20: Konvergenzverhalten für $p = 1$ und h -adaptive Verfeinerungen für die primale und duale Diskretisierung, Effektivitätsindizes.

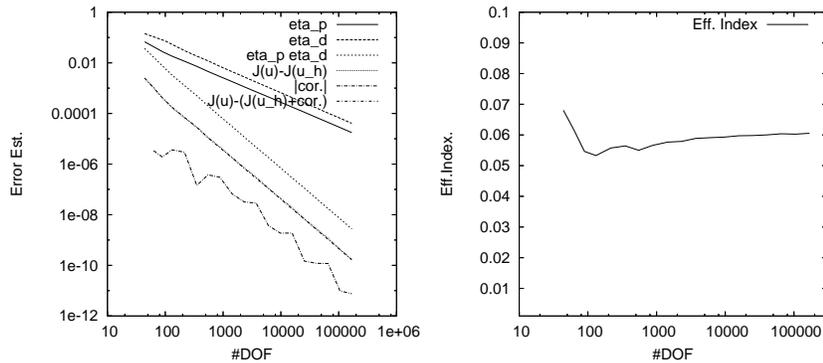


Abb. IV.21: Konvergenzverhalten für $p = 2$ und h -adaptive Verfeinerungen für die primale und duale Diskretisierung, Effektivitätsindizes.

²⁰Alternativ kann z.B. auch eine Adaptierung des dualen Gitters in jedem 2. oder 3. Schritt erfolgen, oder es wird zunächst das primale Problem adaptiert und anschließend mit einem gewissen Prozentsatz der für das primale Problem verwendeten Freiheitsgrade das duale Problem angepasst.

In den Abbildungen IV.20 und IV.21 (links) sind das wahre, das durch den Korrekturterm korrigierte und das durch den Ausdruck $\eta_p \eta_d$ geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen und Polynomgrad $p = 1, 2$ zu sehen. Deutlich erkennbar ist die zuvor beschriebene Verdopplung der Genauigkeit. Außerdem ist offenbar der Korrekturterm in dieser Konfiguration von höherer Ordnung und kann u.U. vernachlässigt werden. Mit Blick auf die in Abbildung IV.20 und IV.21 (rechts) dargestellten Effektivitätsindizes wird eine leichte Gitterabhängigkeit des für $p = 1$ gegebenen Fehlerschätzers sichtbar. Für den Polynomgrad $p = 2$ ist der Effektivitätsindex jedoch konstant.

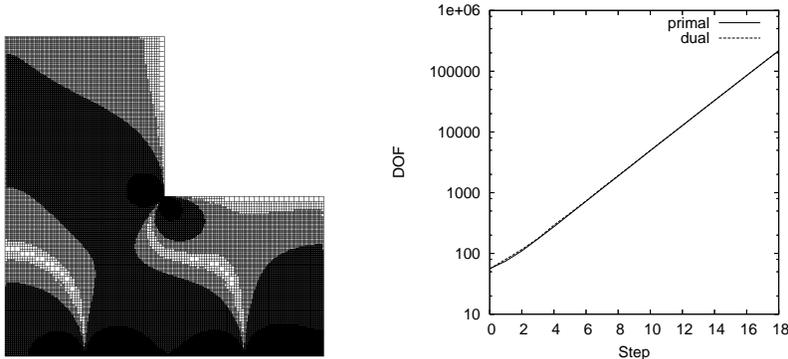


Abb. IV.22: h -adaptives Gitter für das duale Problem bei $p = 1$, Anzahl der Freiheitsgrade.

In den Abbildungen IV.22 und IV.23 (links) sind zum dualen Problem gehörige h -adaptive Gitter für $p = 1$ und $p = 2$ dargestellt. Die h -adaptiven Gitter für das primale Problem sind die bereits in Abbildung IV.16 gezeigten, den Polynomgraden $p = 1$ und $p = 2$ entsprechenden Gitter.

Bei den für das duale Problem erzeugten Gittern sind lokale Verfeinerungen sowohl in der Nähe der einspringenden Ecke als auch in der Umgebung von $(-0.25, -0.5)$ und $(0.25, -0.5)$ zu erkennen. Letzteres ist im Wesentlichen auf Singularitäten zurückzuführen, die sich durch Sprünge in den Neumannschen Randdaten ergeben, also insbesondere durch Eigenschaften von J . Dass die Anzahl der in jedem Adaptierungsschritt verwendeten Freiheitsgrade für das primale und für das duale Problem ungefähr gleich ist, ist in den Abbildungen IV.22 und IV.23 (rechts) zu erkennen.

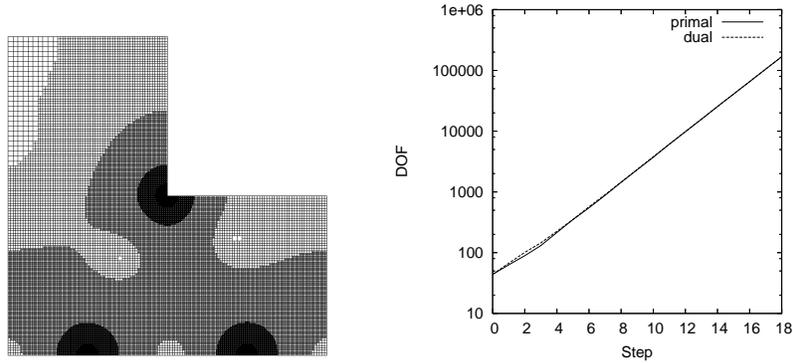


Abb. IV.23: h -adaptives Gitter für das duale Problem bei $p = 2$, Anzahl der Freiheitsgrade.

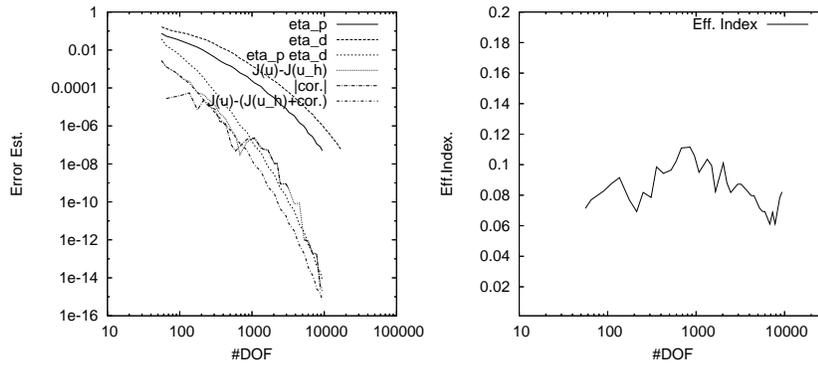


Abb. IV.24: Konvergenzverhalten für hp -adaptive Verfeinerungen für die primale und duale Diskretisierung, Effektivitätsindizes.

In Abbildung IV.24 (links) wird das exponentielle Konvergenzverhalten bei hp -adaptiven Verfeinerungen dargestellt. Auch hier ist die Verdoppelung der Genauigkeit bei Verwendung des korrigierten Fehlermaßes zu erkennen. Bei hp -adaptiven Verfeinerungen kann offenbar der Korrekturterm nicht vernachlässigt werden. Mit Blick auf die Kurven von $J(u) - J(u_h)$, $|cor.|$ und $J(u) - (J(u_h) + cor.)$ wird deutlich, dass erst durch die Addition des Korrekturterms die Schätzung zuverlässig wird, was durch den zugehörigen Effektivitätsindex schließlich bestätigt wird.

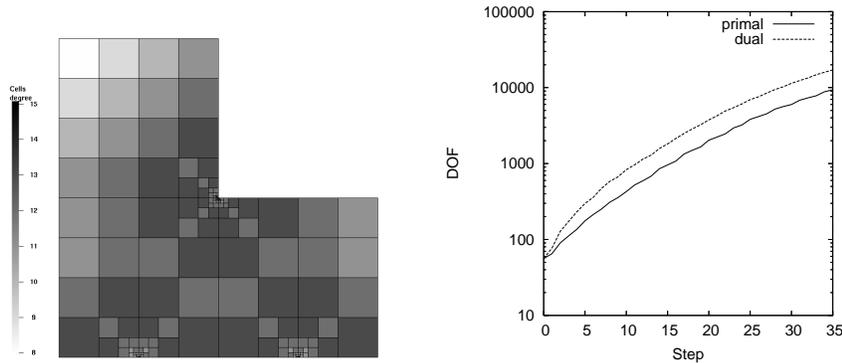


Abb. IV.25: hp -adaptives FE-Gitter für das duale Problem, Anzahl der Freiheitsgrade.

In Abbildung IV.25 (links) ist ein hp -adaptives FE-Gitter für das duale Problem mit den typischen geometrischen Verfeinerungen an Stellen mit Singularitäten und der oben beschriebenen Verteilung des lokalen Polynomgrads zu sehen.

Die Anzahl der Freiheitsgrade für das duale Problem ist, wie in Abbildung IV.25 (rechts) dargestellt, etwas höher als die entsprechende Anzahl für das primale Problem. Jedoch wächst die Anzahl der Freiheitsgrade beider Probleme im gleichen Maße.

Die Verwendung von zwei unterschiedlichen FE-Gittern, einem FE-Gitter für das primale Problem und einem für das duale Problem, ist aus implementierungstechnischen Gesichtspunkten nicht unproblematisch, da die Berechnung des Anteils $a(u_h, z_{\tilde{h}})$ innerhalb des Korrekturterms $\theta(u_h, z_{\tilde{h}})$ einen erheblichen Datentransfer u.U. über mehrere Verfeinerungsebenen hinweg erfordert.

Die Frage ist nun, ob FE-Gitter konstruiert werden können, in denen gleichermaßen die Adaptierungserfordernisse des primalen wie des dualen Problems Berücksichtigung finden, so dass letztlich die Verwendung nur eines FE-Gitters ausreichend ist.

Mit Blick auf die nun folgenden Experimente kann diese Frage durchaus bejaht werden, sofern eine moderate Überverfeinerung in Kauf genommen werden kann.

Die Vorteile, die die Verwendung nur eines FE-Gitters für beide Probleme mit sich bringt, sind offensichtlich: Zum einen ergeben sich erhebliche Vereinfachungen bei Implementierungsansätzen, da kein Datentransfer von

verschiedenen Gittern notwendig ist. Zum anderen entfällt wegen $z_{\tilde{h}} \in V_h$ der Korrekturterm $\theta(u_h, z_{\tilde{h}})$, womit sich die Fehlerabschätzungen (IV.51) und (IV.52) zu

$$|\langle J, u \rangle - \langle J, u_h \rangle| \leq C\eta_p\eta_d$$

vereinfachen.

Zur Gewinnung einer adaptiven Strategie ist der Ausdruck $\eta_p\eta_d$ nicht unmittelbar anwendbar, da hier keine lokalen Fehleranteile dargestellt werden. Mit $\epsilon := \eta_p(2\eta_d)^{-1}$ ist aber

$$\eta_p\eta_d = \frac{1}{4\epsilon}\eta_p^2 + \epsilon\eta_d^2 = \sum_{T \in \mathcal{T}} \frac{1}{4\epsilon}\eta_{p,T}^2 + \epsilon\eta_{d,T}^2.$$

Damit ist die Anwendung der in den Abschnitten IV.7.1 und IV.7.2 vorgestellten Strategien mit den lokalen Fehleranteilen

$$\eta_T := (4\epsilon)^{-1}\eta_{p,T}^2 + \epsilon\eta_{d,T}^2,$$

möglich. In den Abbildungen IV.26 und IV.27 ist links das Konvergenzverhalten bei h -adaptiven Verfeinerungen von nur einem Gitter sowohl für das primale als auch für das duale Problem dargestellt. Erkennbar ist, dass die Konvergenz für $p = 1$ und $p = 2$ von nahezu optimaler Ordnung ist. Ebenso erkennt man auch hier wieder die typische Verdoppelung der Genauigkeit.

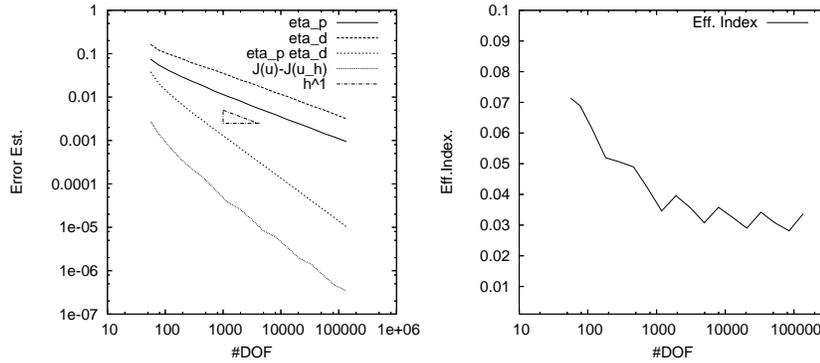


Abb. IV.26: Konvergenzverhalten für $p = 1$ und h -adaptive Verfeinerungen für die primal-duale Diskretisierung, Effektivitätsindizes.

Die Zuverlässigkeit der Fehlerschätzungen wird in den Abbildungen IV.26 und IV.27 (rechts) deutlich. Zwar treten auch hier gewisse Oszillationen bei den Effektivitätsindizes auf, diese fallen aber relativ moderat aus.

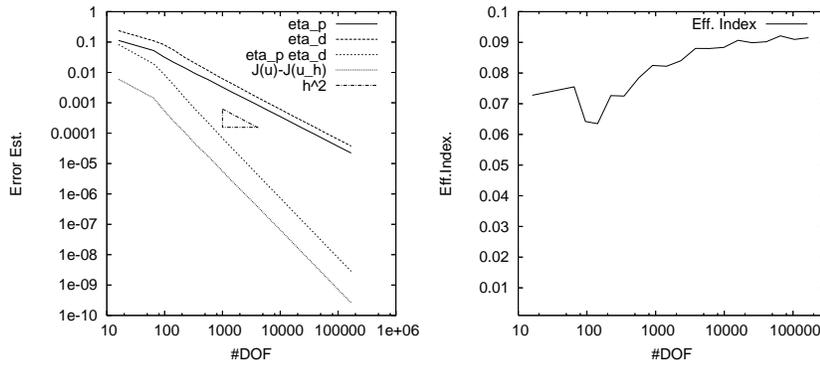


Abb. IV.27: Konvergenzverhalten für $p = 2$ und h -adaptive Verfeinerungen für die primal-duale Diskretisierung, Effektivitätsindizes.

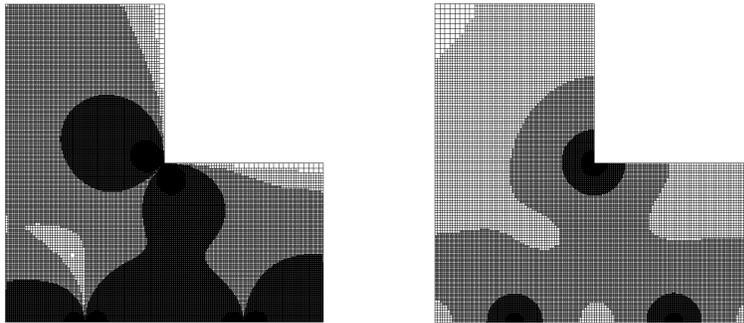


Abb. IV.28: h -adaptives Gitter für das primale und das duale Problem bei $p = 1$ und $p = 2$.

In Abbildung IV.28 sind h -adaptive Gitter für $p = 1$ und $p = 2$ dargestellt, die den oben erwähnten Berechnungen zugrunde liegen. Für beide der hier betrachteten Polynomgrade kommt es zumindest aus Sicht des primalen Problems zu den erwarteten Überverfeinerungen. Diese sind vor allem in den Bereichen zu lokalisieren, bei denen die Neumannschen Randdaten des dualen Problems Sprünge aufweisen. Letztlich ist jedoch die Anzahl der zusätzlichen Freiheitsgrade, die aus den Überverfeinerungen resultieren, gegenüber der Gesamtzahl der Freiheitsgrade vernachlässigbar.

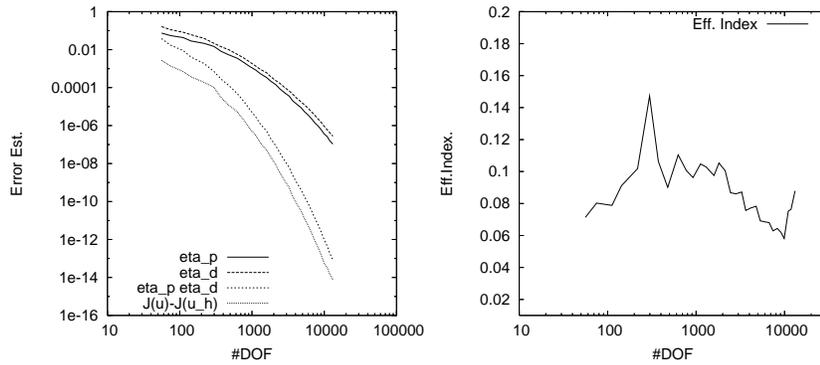


Abb. IV.29: Konvergenzverhalten für hp -adaptive Verfeinerungen für die primal-duale Diskretisierung, Effektivitätsindizes.

Abbildung IV.29 zeigt das exponentielle Konvergenzverhalten bei hp -adaptiven Verfeinerungen. Deutlich ist auch hier die Verdoppelung der Genauigkeit zu erkennen. Der zugehörige Effektivitätsindex bewegt sich in einem moderaten Bereich, womit die Schätzung als zuverlässig zu betrachten ist.

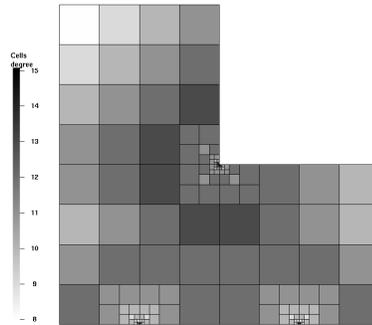


Abb. IV.30: hp -adaptives FE-Gitter für das primale und das duale Problem.

In Abbildung IV.30 ist ein hp -adaptives FE-Gitter, das dem primalen und dualen Problem zugrunde liegt, dargestellt. Wie bei den h -adaptiven Gittern zuvor ist auch hier erkennbar, dass alle Adaptierungsanforderungen sowohl des primalen als auch des dualen Problems berücksichtigt werden. So sind geometrische Verfeinerungen in der Nähe von Singularitäten des primalen wie des dualen Problems sichtbar. Auch die Polynomgradverteilung

entspricht den Verteilungen wie sie in Abschnitt IV.7.3 diskutiert werden. Die Überverfeinerung für das primale Problem im Bereich von $(-0.25, -0.5)$ und $(0.25, -0.5)$ wirkt sich nicht wesentlich auf das asymptotische Verhalten von η_p aus.

IV.7.5 h - und hp -adaptive Verfeinerungen für Modellprobleme

Die in Abschnitt IV.7.3 vorgestellten Experimente für H^1 -Fehlerabschätzungen bzgl. Variationsgleichungen zeigen, dass die Anwendung der in den Abschnitten IV.7.1 und IV.7.2 vorgestellten h - und hp -adaptiven Strategien in der Regel zu einem Konvergenzverhalten führt, das bezüglich der verwendeten Polynomgradverteilung angemessen ist.

In diesem Abschnitt wird anhand von experimentell gewonnenen Daten untersucht, ob Ergebnisse dieser Form auch für modellhafte elliptische Minimierungsprobleme erster und zweiter Art gewonnen werden können. Auch hierbei werden die h - und hp -adaptiven Strategien aus den Abschnitten IV.7.1 und IV.7.2 verwendet.

Entscheidend für die Herleitung der in Abschnitt IV.3 beschriebenen Methode zur Fehlerkontrolle sind Fehlerschätzer für Variationsgleichungen, wie sie etwa in Abschnitt IV.2 angegeben werden. Die experimentell ermittelten Daten aus den Abschnitten IV.2 und IV.7.3 lassen auf eine hohe Zuverlässigkeit dieser Schätzer schließen, wodurch sie im besonderen Maße für die Anwendung in diesem Zusammenhang prädestiniert sind.

Die Modellprobleme, die diesem Abschnitt zugrunde liegen, sind das vereinfachte Signorini-Problem (II.10) und das modellhafte Reibungsproblem (II.21) aus den Abschnitten II.2.1 bzw. II.2.2 mit den in diesen Abschnitten aufgeführten konkreten Funktionen f , q , g und s . In den folgenden Testrechnungen wird stets die Polynomgradverteilung $p_1 \leq p - 1$ für $M^{p_1}(\mathcal{T}_{1,H})$ vorausgesetzt. Das Gitter $\mathcal{T}_{1,H}$ ist so gewählt, dass $h/H \leq 0.5$ ist. Diese Wahl erweist sich bei den hier betrachteten Beispielen im Wesentlichen als stabil.²¹

Da weder für das vereinfachte Signorini-Problem noch für das modellhafte Reibungsproblem analytische Lösungen zur Verfügung stehen, können die Effektivitätsindizes der in den vorigen Abschnitten hergeleiteten Fehlerschätzer nicht bestimmt werden.²² Entscheidend bei den folgenden Un-

²¹vgl. Abschnitt III.1.1

²²Ist die exakte Lösung nicht bekannt, kann zur Ermittlung des Effektivitätsindex häufig eine Referenzlösung herangezogen werden, die mit hohem Aufwand z.B. auf ei-

tersuchungen ist jedoch, dass das Konvergenzverhalten qualitativ mit Hilfe der Fehlerschätzer geschätzt werden kann.²³

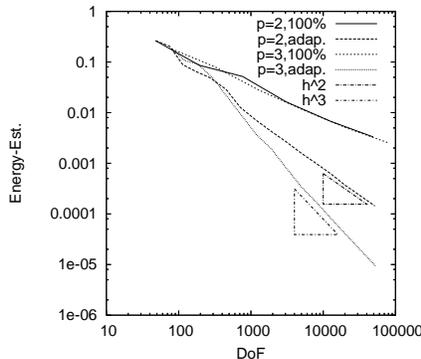


Abb. IV.31: Geschätztes Konvergenzverhalten für h -adaptive Verfeinerungen (Signorini-Problem).

In Abbildung IV.31 ist das geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen für die Polynomgrade $p = 2$ und $p = 3$ dargestellt. Den Berechnungen liegt das vereinfachte Signorini-Problem zugrunde. Da die Lösung für dieses Problem im Wesentlichen H^2 -regulär ist, liefern globale Verfeinerungen bei Polynomgrad $p = 1$ bereits die optimale Konvergenzordnung $\mathcal{O}(h)$. Eine Adaptierung des Gitters hat demnach keinen Sinn.²⁴ Diese ist erst für Polynomgrade $p \geq 2$ zweckmäßig, da in diesem Fall Konvergenzverbesserungen erwartet werden können und in Abbildung IV.31 auch tatsächlich zu erkennen sind: Für die Polynomgrade $p = 2$ und $p = 3$ wird die optimale Konvergenzordnung $\mathcal{O}(h^p)$ erreicht.

In Abbildung IV.32 sind h -adaptive Gitter für die Polynomgrade $p = 2$ und $p = 3$ dargestellt. Erkennbar sind lokale Verfeinerungen in den Ecken und vor allem an den Randbereichen der Kontaktzone.²⁵ Zudem gibt es lokale

nem stark verfeinerten, regelmäßigen Gitter berechnet wurde. Bei adaptiven Verfahren insbesondere mit hohen Polynomgraden ist dieses Vorgehen im Allgemeinen jedoch nicht geeignet, da in den meisten Fällen die Genauigkeit der adaptiven Lösung höher als die der Referenzlösung ist.

²³vgl. Abschnitt IV.3

²⁴Tatsächlich erhält man bei der Anwendung der oben vorgestellten h -adaptiven Strategie bis auf unwesentliche Verfeinerungen im Kontaktbereich nur globale Verfeinerungen.

²⁵vgl. Abbildung II.1 (rechts)

Verfeinerungen in der Kontaktzone selbst.

Auffällig sind auch hier die Gitterstrukturen für $p = 3$. Während für $p = 2$ radiale Verfeinerungen zu sehen sind, ist dies für $p = 3$ nicht der Fall.

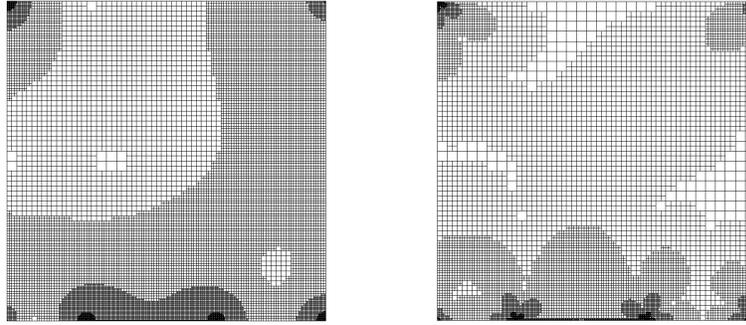


Abb. IV.32: h -adaptive Gitter für $p = 2$ und $p = 3$ (vereinfachtes Signorini-Problem).

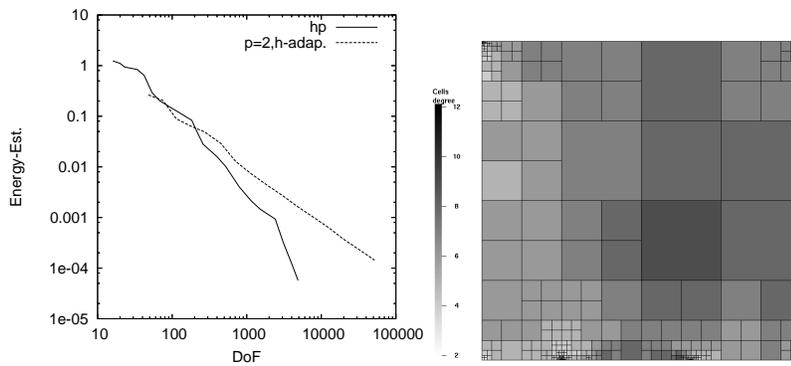


Abb. IV.33: Geschätztes Konvergenzverhalten für h - und hp -adaptive Verfeinerungen, hp -adaptives FE-Gitter für das vereinfachte Signorini-Problem.

In Abbildung IV.33 (links) ist das geschätzte Konvergenzverhalten bei hp -adaptiven Verfeinerungen zu sehen. Außerdem ist zum Vergleich das geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen mit $p = 2$ abgebildet. Erkennbar ist die exponentielle Konvergenz der hp -Variante. In Abbildung IV.33 (rechts) ist ein hp -adaptives FE-Gitter zu sehen, das während des Verfeinerungsprozesses erzeugt wurde. Im Wesentlichen sind

auch hier die typischen geometrischen Verfeinerungen bezüglich der Ecken und in den Randbereichen der Kontaktzone zu sehen. In der Kontaktzone selbst kommt es sowohl zu Verfeinerungen als auch zur Erhöhung des Polynomgrads.

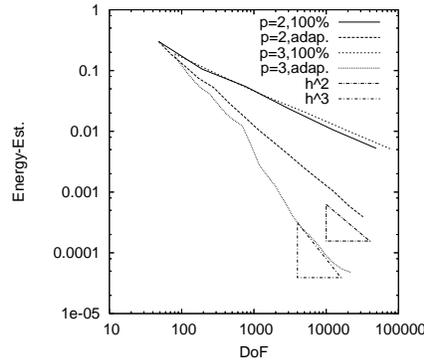


Abb. IV.34: Geschätztes Konvergenzverhalten für h -adaptive Verfeinerungen (Idealisiertes Reibungsproblem).

Abbildung IV.34 zeigt das geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen für das idealisierte Reibungsproblem. Der Polynomgrad ist hier $p = 2$ und $p = 3$. Wie bei dem vereinfachten Signorini-Problem ist die Anwendung von h -adaptiven Strategien bezüglich $p = 1$ nicht zweckmäßig, da auch hier die Lösung über die notwendige Regularität verfügt, so dass globale Verfeinerungen bereits die optimale Konvergenzordnung liefern. Für $p = 2$ und $p = 3$ erhält man dagegen erst bei Verwendung von Adaptierungsstrategien die angemessene Konvergenzordnung $\mathcal{O}(h^p)$.

In Abbildung IV.35 sind adaptive Gitter für die Polynomgrade $p = 2$ und $p = 3$ dargestellt. Man erkennt deutlich lokale Verfeinerungen in den Ecken und an den Übergangsstellen von $|q - \partial_n u| = s$ zu $|q - \partial_n u| < s$.²⁶ Für $p = 2$ bilden die Gitterstrukturen eher radiale Formen, für $p = 3$ erkennt man das für diesen Polynomgrad typische Muster wieder.

²⁶vgl. Abbildung II.2

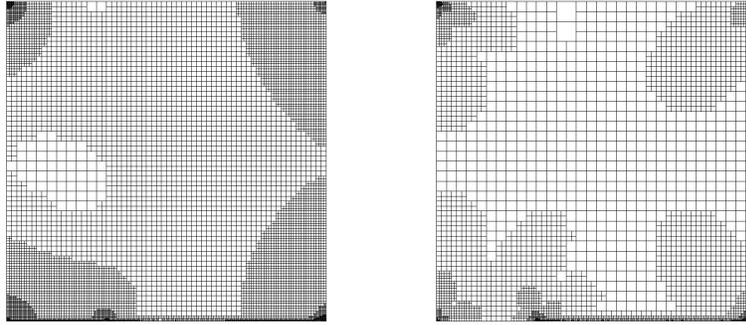


Abb. IV.35: h -adaptive Gitter für $p = 2$ und $p = 3$ (Idealisiertes Reibungsproblem).

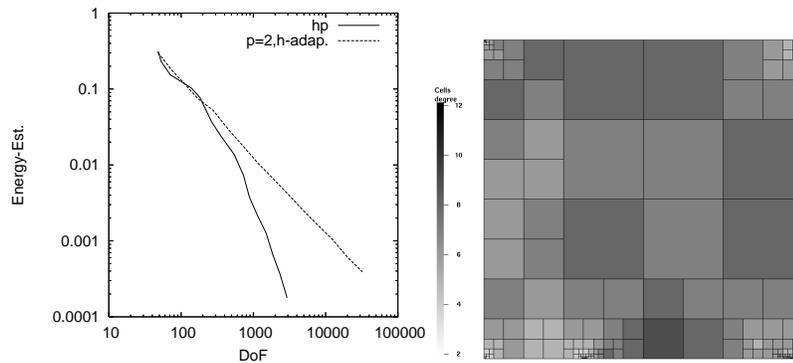


Abb. IV.36: Geschätztes Konvergenzverhalten für h - und hp -adaptive Verfeinerungen, hp -adaptives FE-Gitter für das idealisierte Reibungsproblem.

Abbildung IV.36 (links) zeigt das geschätzte Konvergenzverhalten bei hp -adaptiven Verfeinerungen. Zum Vergleich ist zusätzlich das geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen mit Polynomgrad $p = 2$ abgebildet.

Das in Abbildung IV.36 (rechts) dargestellte FE-Gitter, das während des hp -Verfeinerungsprozesses erzeugt wurde, weist ebenfalls die üblichen Charakteristika auf: geometrische Verfeinerungen in den Ecken und an den wie oben beschriebenen Übergangsstellen, zudem die für hp -Anwendungen typische Polynomgradverteilung.

IV.7.6 *h*- und *hp*-adaptive Verfeinerungen für linear-elastische Kontaktprobleme

Im Mittelpunkt dieses Abschnitts stehen numerische Experimente für die in Abschnitt II.4 gewonnenen Modelle für das Industrieroboter-gestützte Bandschleifen. Wie im Abschnitt zuvor wird untersucht, inwiefern *h*- und *hp*-adaptive Strategien zur Konvergenzverbesserung eingesetzt werden können. Das wesentliche Resultat ist, dass die im vorherigen Abschnitt für idealisierte Modellprobleme gewonnenen Erkenntnisse in der Regel auf elliptische Minimierungsprobleme erster und zweiter Art, denen ein linear-elastisches Modell zugrunde liegt, übertragen werden können: Durch lokale Verfeinerungen im Kontaktbereich und speziell bei *hp*-adaptiven Verfeinerungen zusätzlich durch kleine Polynomgrade im Kontaktbereich und hohe Polynomgrade außerhalb können Konvergenzgewinne erzielt werden, die dem gewählten Verfeinerungstyp (*h*- oder *hp*-Methode) entsprechen. Für die in diesem Abschnitt beschriebenen Testrechnungen wird ebenfalls $p_1 \leq p-1$ und $h/H \leq 0.5$ gewählt, da diese Wahl sich bei den Berechnungen als stabil erweist.

Das zur Verfügung stehende Datenmaterial ist aufgrund der hohen Komplexität dieser Problemklassen²⁷ an einigen Stellen nicht immer ausreichend, um eindeutige, den Erwartungen entsprechende Aussagen zu treffen. Dieses Manko wird hoffentlich in Zukunft durch den Einsatz effizienterer Rechnerumgebungen und verfeinerter Optimierungsverfahren überwunden.

Abbildung IV.37 zeigt das geschätzte Konvergenzverhalten bei *h*-adaptiven Verfeinerungen für $p = 1$ und $p = 2$. Das zugrunde liegende Modell entspricht dem Problem (II.46) bezüglich der in Abschnitt II.4.1 beschriebenen reibungsfreien Kontaktsituation beim Industrieroboter-gestützten Bandschleifen.

Erkennbar ist, dass sowohl für den Polynomgrad $p = 1$ als auch für den Polynomgrad $p = 2$ Konvergenzverbesserungen erzielt werden, die adäquat zum jeweils gewählten Polynomgrad sind. Insbesondere wird deutlich, dass im Gegensatz zu den Modellbeispielen wenige globale Verfeinerungen für den Polynomgrad $p = 1$ nicht ausreichen, um die theoretisch mögliche Konvergenzordnung von $\mathcal{O}(h)$ zu bestätigen.²⁸

²⁷Zu bedenken ist, dass die hohe Komplexität vor allem durch die dreidimensionale Modellierung bedingt ist. Möglichkeiten zur Dimensionsreduktion innerhalb der Modellierung (z.B. in Form eines ebenen Spannungs- oder Verzerrungszustands) sind aufgrund der speziellen Kontaktverhältnisse beim Industrieroboter-gestützten Bandschleifen nicht zu rechtfertigen.

²⁸vgl. Abschnitt III.1.2

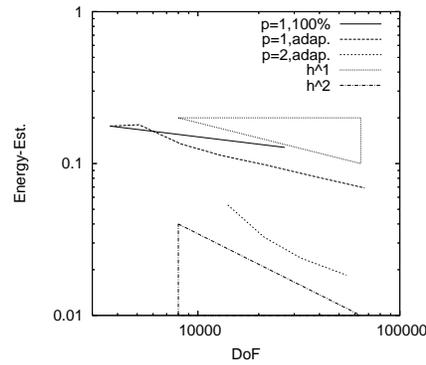


Abb. IV.37: Geschätztes Konvergenzverhalten für h -adaptive Verfeinerungen (Reibungsfreier Kontakt).

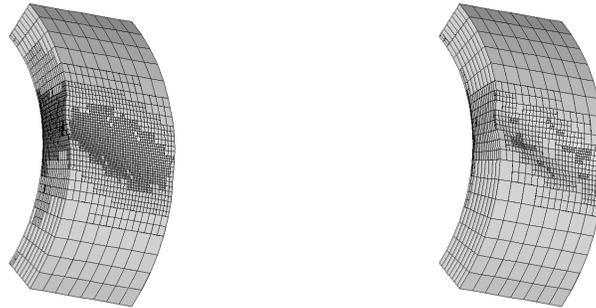


Abb. IV.38: h -adaptive Gitter für $p = 1$ und $p = 2$ (Reibungsfreier Kontakt).

In Abbildung IV.38 sind h -adaptive Gitter für $p = 1$ und $p = 2$ dargestellt. Vor allem im Kontaktbereich treten lokale Verfeinerungen auf. Insbesondere können aufgrund dieser lokalen Verfeinerungen die für das Industrieroboter-gestützte Bandschleifen interessanten Größen, wie zum Beispiel die Andruckkraft im Kontaktbereich, genauer bestimmt werden.²⁹ Vermutlich sind Gitter, die aufgrund von Goal-Oriented-Fehlerschätzern zielgerichtet für eine genauere Bestimmung der Andruckkräfte adaptiert werden, noch vorteilhafter. Das geschätzte Konvergenzverhalten bei hp -adaptiven Verfeine-

²⁹vgl. Abbildung II.10

rungen ist in Abbildung IV.39 (links) zu sehen. Ein erkennbarer exponentieller Verlauf kann nicht festgestellt werden. Der Grund hierfür ist vermutlich, dass die erwartete exponentielle Konvergenz erst relativ spät einsetzt und durch die geringe Anzahl der verfügbaren Verfeinerungsschritte noch nicht hervorgerufen wird. Dagegen besitzt das in Abbildung IV.39 (rechts) dargestellte hp -adaptive FE-Gitter die hp -typischen Merkmale mit lokalen Verfeinerungen im Kontaktbereich und einer entsprechenden Polynomgradverteilung.

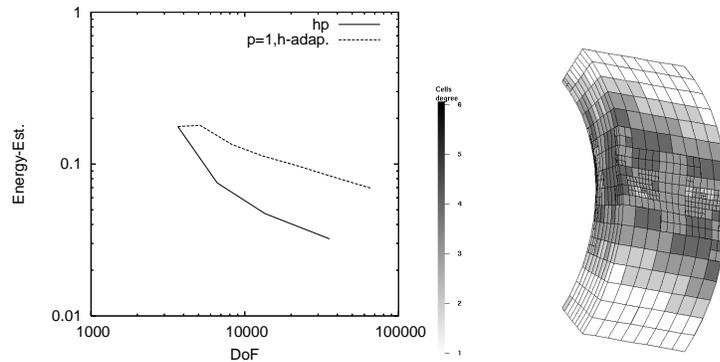


Abb. IV.39: Geschätztes Konvergenzverhalten für h - und hp -adaptive Verfeinerungen, hp -adaptives FE-Gitter für das reibungsfreie Kontaktproblem.

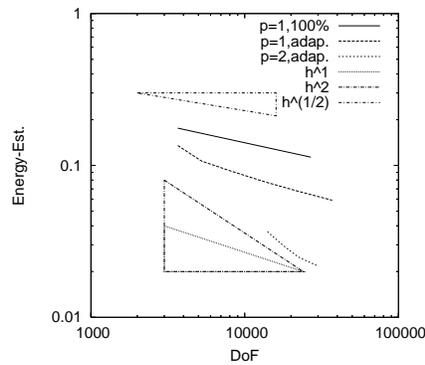


Abb. IV.40: Geschätztes Konvergenzverhalten für h -adaptive Verfeinerungen (Reibungsbehafteter Kontakt mit vorgegebener Normalspannung).

Für das Reibungsproblem mit vorgegebener Normalspannung (II.77) erhält man das in Abbildung IV.40 dargestellte geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen sowie $p = 1$ und $p = 2$. Diesem Problem liegt die Konfiguration aus Abschnitt II.4.2 zugrunde, die den reibungsbehafteten Kontakt mit vorgegebener Normalspannung beschreibt. Bei den hier vorgestellten Testrechnungen wird $q_t = 0$ gesetzt, um den Einfluss von Sprüngen in den Randdaten auf den Verfeinerungsprozess zu vermeiden. Erkennbar ist, dass der Konvergenzgewinn für den Polynomgrad $p = 1$ im Wesentlichen den Erwartungen entspricht. Für den Polynomgrad $p = 2$ ist die Güte des Konvergenzgewinns nicht klar erkennbar, da die zur Verfügung stehenden Daten für genauere Aussagen nicht ausreichen.

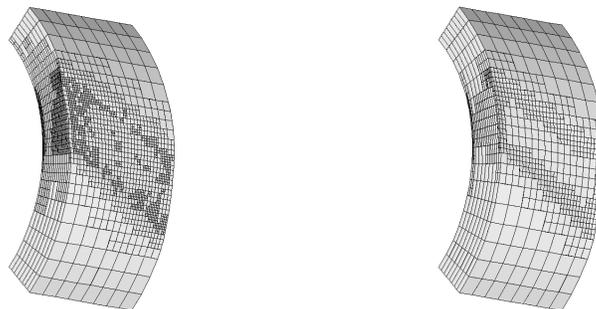


Abb. IV.41: h -adaptive Gitter für $p = 1$ und $p = 2$ (Reibungsbehafteter Kontakt mit vorgegebener Normalspannung).

In Abbildung IV.41 sind durch den Verfeinerungsprozess gewonnene h -adaptive Gitter für die Polynomgrade $p = 1$ und $p = 2$ dargestellt. An den FE-Gittern sind klar die erwarteten Strukturen mit einer deutlichen Tendenz zu lokalen Verfeinerungen im Kontaktbereich zu erkennen. Insbesondere sind lokale Verfeinerungen im Kontaktzonenrand auszumachen, was bereits bei den Modellrechnungen zu beobachten ist.

Abbildung IV.42 (links) zeigt das geschätzte Konvergenzverhalten bei hp -adaptiven Verfeinerungen. Ähnlich wie bei der entsprechenden Untersuchung des reibungsfreien Kontakts ist eine exponentielle Konvergenz nicht direkt erkennbar. Auch hier ist das zur Verfügung stehende Datenmaterial für genauere Aussagen unzureichend.

Dagegen besitzt das in Abbildung IV.42 (rechts) dargestellte hp -adaptive

FE-Gitter die erwarteten hp -typischen Merkmale in Bezug auf lokale Verfeinerungen und Polynomgradverteilung.

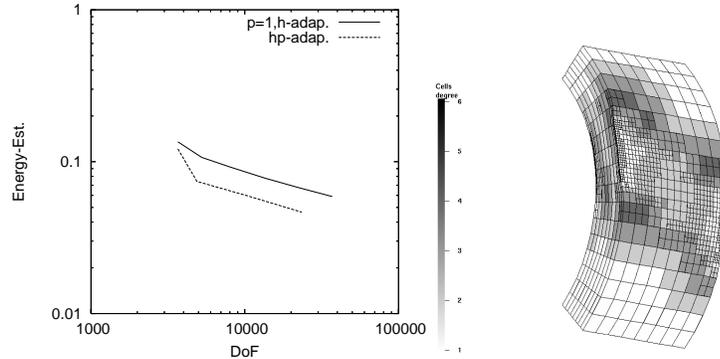


Abb. IV.42: Geschätztes Konvergenzverhalten für h - und hp -adaptive Verfeinerungen, hp -adaptives FE-Gitter für reibungsbehafteten Kontakt mit vorgegebener Normalspannung.

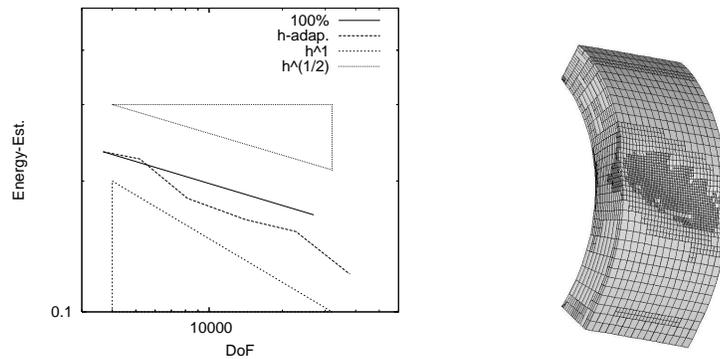


Abb. IV.43: Geschätztes Konvergenzverhalten für h -adaptive Verfeinerungen, h -adaptives Gitter für reibungsbehafteten Kontakt.

In Abbildung IV.43 (links) ist das geschätzte Konvergenzverhalten bezüglich des reibungsbehafteten Kontaktproblems für $p = 1$ dargestellt. Offenbar führt die Anwendung von h -adaptiven Strategien zu einer deutlichen Konvergenzsteigerung. Bei globalen Verfeinerungen erhält man im Wesentlichen

die in Satz III.10 vorhergesagte Konvergenzordnung von $\mathcal{O}(h^{1/2})$. Abbildung IV.43 (rechts) zeigt das zugehörige h -adaptive Gitter mit lokalen Verfeinerungen im Kontaktbereich.

Der Vergleich mit den Aussagen aus Satz III.10 ist jedoch kritisch zu beurteilen, da der Reibungswiderstand s durch die in Abschnitt II.3.4 vorgestellten Fixpunktiteration des diskreten Lagrangeschen Multiplikators $\lambda_{n,H}$ bestimmt wird, deren Konvergenzeigenschaften nicht bekannt sind.³⁰

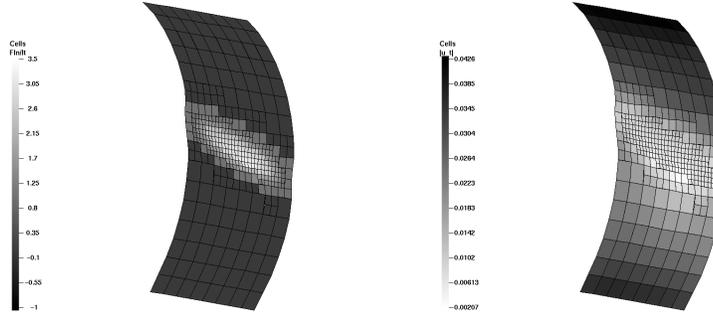


Abb. IV.44: $\mathcal{F}|\lambda_{n,H}|/|\lambda_{t,H}|$, Betrag der Tangentialverschiebungen $|\delta_t(u)|$.

Abbildung IV.44 zeigt den Quotienten $\mathcal{F}|\lambda_{n,H}|/|\lambda_{t,H}|$ und den Betrag der Tangentialverschiebungen $|u_t|$. Die Verteilung von Haft- und Gleitzonen korrespondiert auch hier zu den dargestellten Tangentialverschiebungen. Im Vergleich zu den Berechnungen, die Abbildung II.15 zugrunde liegen, sind diese Zonen durch den adaptiven Verfeinerungsprozess jedoch wesentlich besser aufgelöst.

IV.7.7 h - und hp -adaptive Verfeinerungen für Hindernisprobleme

Das Ziel dieses letzten Abschnitts besteht darin, anhand von numerischen Experimenten den Einsatz adaptiver Verfahren für das Hindernisproblem aus Abschnitt II.5.1 zu untersuchen. Die zugrunde liegende Beispielkonfiguration ist ebenfalls diesem Abschnitt zu entnehmen.

Im Gegensatz zu den vorherigen Abschnitten wird im Folgenden $\tilde{p} \leq p - 2$ und $h/H = 1$ für die Diskretisierung gewählt, da hierbei $\mathcal{T}_h = \tilde{\mathcal{T}}_H$ gilt. Diese

³⁰vgl. [46], [63]

Wahl erweist sich bei der Durchführung der Experimente als vorteilhaft hinsichtlich der Approximationsgüte. Allerdings treten bei den durchgeführten Testrechnungen Stabilitätsprobleme für sehr feine Gitter bzw. hohe Polynomgrade auf, so dass Experimente mit angemessenen Lagrangeschen Multiplikatoren nur für eine eher geringe Anzahl von Freiheitsgraden durchführbar sind.

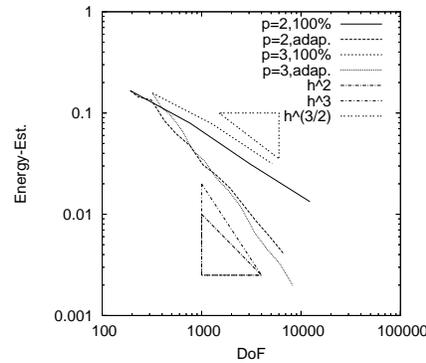


Abb. IV.45: Geschätztes Konvergenzverhalten für h -adaptive Verfeinerungen.

Abbildung IV.45 zeigt das geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen und $p = 2$ und $p = 3$. Darüberhinaus ist das geschätzte Konvergenzverhalten bei globalen Verfeinerungen abgebildet. Für den Polynomgrad $p = 1$ sind h -adaptive Verfeinerungen nicht notwendig, da bereits globale Verfeinerungen die Konvergenzordnung $\mathcal{O}(h)$ sichern. Dagegen wird für die Polynomgrade $p = 2$ und $p = 3$, wie in der Abbildung erkennbar, erst durch Anwendung h -adaptiver Verfeinerungen die optimale Konvergenzordnung $\mathcal{O}(h^p)$ erreicht.

Abbildung IV.46 zeigt die zugehörigen h -adaptiven Gitter. Deutlich erkennbar ist, dass alle kritischen Bereiche durch lokale Verfeinerungen erfasst werden. Hierzu gehören einerseits die Ecken als auch die Kontaktzone. Insbesondere wird der Grenzbereich zwischen Membran und Hindernis durch Verfeinerungen aufgelöst.

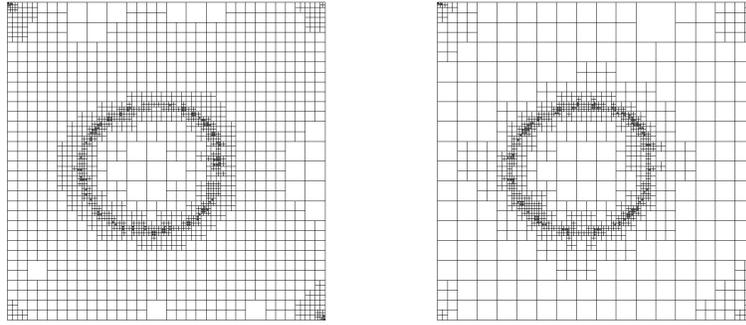


Abb. IV.46: h -adaptive Gitter für $p = 2$ und $p = 3$.

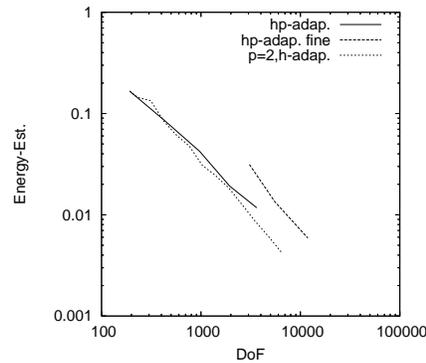


Abb. IV.47: Geschätztes Konvergenzverhalten für h - und hp -adaptive Verfeinerungen.

Abbildung IV.47 zeigt das geschätzte Konvergenzverhalten bei hp -adaptiven Verfeinerungen (64 und 1024 Elemente im Ausgangsgitter) sowie zum Vergleich das geschätzte Konvergenzverhalten bei h -adaptiven Verfeinerungen und $p = 2$. Die erwartete exponentielle Konvergenz ist nicht unmittelbar erkennbar, was auch hier auf eine ungenügende Anzahl von Verfeinerungsschritten zurückzuführen ist.

Dagegen erkennt man in Abbildung IV.48 die typischen hp -adaptiven Gitterstrukturen mit Verfeinerungen im Bereich der Kontaktzone und hohen Polynomgraden in kontaktfreien Bereichen.

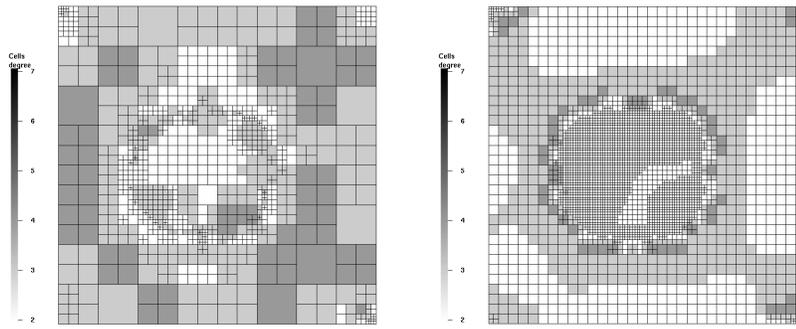


Abb. IV.48: hp -adaptives FE-Gitter mit 64 und 1024 Elementen im Ausgangsgitter.

Kapitel V

Ausblick

Diese Arbeit beinhaltet hauptsächlich drei Untersuchungsschwerpunkte: Modellierung von Kontaktproblemen, h - und hp -Finite-Elemente-Diskretisierungen sowie a-posteriori Fehlerkontrolle für Kontaktprobleme und Adaptivität. Exemplarisch werden Modellprobleme vom Signorini-Typ und idealisierte Reibungsprobleme sowie, als Anwendung in der Fertigungstechnik, Kontaktprobleme beim Industrieroboter-gestützten Bandschleifen betrachtet. Darüber hinaus werden Anwendungsmöglichkeiten auch für Modellprobleme anderen Typs diskutiert. Ein wichtiger Untersuchungsaspekt ist die Überprüfung und Bewertung der gewonnenen Resultate anhand numerischer Experimente. Jedoch bleibt in allen drei Untersuchungsschwerpunkten eine Reihe von Fragen ungeklärt:

Die Modellierung von Kontaktproblemen beschränkt sich auf Kontaktprobleme mit einseitigem Kontakt. Aus ingenieurwissenschaftlicher Sicht sind vor allem Mehrkörperkontaktsysteme interessant, bei denen auch das Hindernis als verformbar vorausgesetzt wird. Darüber hinaus ermöglicht zwar die Verwendung eines statischen, rein elastischen Modells einen überschaubaren mathematischen Zugang, jedoch entspricht in der Regel die Annahme von dynamischem und nichtlinearem Materialverhalten eher der physikalisch/mechanischen Realität. Insbesondere ist die Beschreibung des Materialverhaltens von Gummi-artigen Kontaktscheibenbelägen beim Industrieroboter-gestützten Bandschleifen mit Hilfe eines linear-elastischen Modells ein Grenzfall. Hier bleibt sicherlich Raum zur Diskussion. Die Anwendung von speziellen Gummi-spezifischen Stoffgesetzen, insbesondere die Berücksichtigung von Inkompressibilitäten, führt unter Umständen zu realitätsnä-

heren Modellierungen.¹ In jedem Fall ist bei der Ausweitung der Modellierung zu prüfen, ob und in welcher Weise eine Einbettung, ähnlich wie in dieser Arbeit für den einfacheren linear-elastischen Fall geschehen, in einen funktionalanalytischen Rahmen möglich ist. Überdies erfordern komplizierte Modellannahmen auch spezielle Finite-Elemente-Implementierungen und aufwändige Konvergenzanalysen. Resultate sind in der Regel nur für die h -Methode bekannt. Finite-Elemente-Zugänge mit höheren Polynomordnungen werden für nichtlineare Kontaktprobleme erst in jüngerer Zeit untersucht. Hier sei z.B. auf [74] und darin enthaltene Referenzen verwiesen. Fehlerkontrolltechniken für dynamische bzw. nichtlineare Kontaktprobleme werden in der Literatur häufig nur für Spezialfälle und auch hier in der Regel nur für Finite-Elemente-Ansätze niedriger Ordnung angegeben. Ob der in dieser Arbeit vorgeschlagene allgemeine Zugang mit Hilfe variationeller Sattelpunktformulierungen auch Fehlerkontrollen für Probleme dieses Typs ermöglicht, ist in Zukunft sicherlich noch zu untersuchen.

Die Stabilität der in Kapitel 3 vorgeschlagenen Diskretisierung von Sattelpunktformulierungen beruht im Wesentlichen auf der Verwendung gröberer Gitter bzw. niedrigerer Polynomgrade für die Diskretisierung der Lagrangeschen Multiplikatoren im Vergleich zum primalen Diskretisierungsansatz. Für den Nachweis der Stabilität ist das Erfülltsein eines schwer überprüfbar Dualitätsarguments hinreichend. Außerdem ist das zu wählende Verhältnis von Maschenweiten und Polynomgraden von einer im Allgemeinen unbekannt Konstanten abhängig. Letztlich ist in der Regel Gewissheit über die Stabilität dieser Diskretisierung nur in einem begrenzten Maße unter Umständen erst mit Hilfe numerischer Experimente zu erzielen. Rigorosere Stabilitätsnachweise sind für nichtkonforme Raviart-Thomas-Ansätze bekannt. Hierzu sei etwa auf [29] verwiesen. Die Frage ist jedoch, ob und in welcher Form stabile Diskretisierungen von diesem Typ auch im hp -Kontext und für a-posteriori Fehlerkontrolltechniken bei Kontaktproblemen verwendbar sind.

In Kapitel 3 werden hierarchische Basen auf der Grundlage von Gegenbaupolynomen für zwei- und dreidimensionale Diskretisierungen vorgestellt, und es werden Möglichkeiten zur Einarbeitung von Stetigkeitsanforderungen über irreguläre Kanten und Facetten hinweg aufgezeigt. Alternativ zu hierarchischen Basen können aber auch Basen mit Lagrange-artigen Basisfunktionen eingesetzt werden. Die in diesem Kapitel beschriebenen rekursiv definierten Datenstrukturen zur Behandlung irregulärer Knoten, Kanten und Facetten können in leicht modifizierter Form wiederverwendet werden.

¹vgl. z.B. [72]

Die Kopplungsgewichte sind hingegen neu zu bestimmen. Ob dies in ähnlich kompakter Form für alle Polynomgrade wie bei den rekursiv definierten hierarchischen Basen gelingt, ist allerdings fraglich.

Für die numerische Bestimmung von Diskretisierungslösungen sind im Allgemeinen effiziente Optimierungsverfahren erforderlich. Das SOR-Verfahren mit Projektion bzw. dessen beschleunigte SSOR-Variante ist sicherlich für moderat dimensionierte Probleme ein akzeptables Verfahren, zumal hochoptimierte Verfahren kommerzieller Optimierungspakete für spezielle Finite-Elemente-Anwendungen nicht notwendigerweise effizienter sein müssen. Jedoch ist der Einsatzbereich dieses einfachen Verfahrens relativ begrenzt, da insbesondere nur Finite-Elemente-Ansätze mit $p = 1$ behandelt werden können. Ähnlich ist die Problematik bei Sattelpunktproblemen. Die in dieser Arbeit beschriebenen Verfahren vom Uzawa- bzw. Arrow-Hurwicz-Typ sind zwar leicht implementierbar, jedoch sind diese im Zusammenhang von Finite-Elemente-Diskretisierungen höherer Ordnung nicht ohne weiteres einsetzbar. Hauptproblem ist, wie auch bei den vorgestellten projektiven SOR-Verfahren, die Angabe einer geeigneten Projektion auf die Restriktionsmenge. Durch Umformulieren in Minimierungsprobleme und unter zur Hilfenahme von kommerziellen Optimierungspaketen können auch Probleme mit höheren Polynomordnungen vom p - oder hp -Typ in einem gewissen Ausmaß numerisch gelöst werden. Jedoch sind auf diese Weise hochdimensionierte Probleme (> 25000 Variablen für die Lagrangeschen Multiplikatoren) mit den zur Verfügung stehenden Optimierungspaketen SQOPT und SNOPT nicht in den Griff zu bekommen. Die Entwicklung effizienter Optimierungsverfahren für hochdimensionierte Finite-Element-Diskretisierungen von Kontaktproblemen u.U. mit Hilfe erprobter Mehrgittertechniken ist nach wie vor eine große Herausforderung.

Die in Kapitel 4 vorgestellten allgemeinen Fehlerkontrolltechniken für Kontaktprobleme können prinzipiell mit allen Fehlerschätzern für Variationsgleichungen konkretisiert werden. In dieser Arbeit werden jedoch nur residuale Fehlerschätzer vorgestellt und numerisch getestet, da diese insbesondere für Finite-Elemente-Ansätze höherer Ordnung geeignet sind. Entscheidend für die Herleitung residualer Fehlerschätzer ist die Nutzung von Interpolationsabschätzungen vom Clément-Typ, die jedoch nur für den zweidimensionalen Fall nachgewiesen sind. Abschätzungen für dreidimensionale Clément-Interpolierende sind ein offenes Problem.

Ferner ist zu erwarten, dass der hier aufgezeigte allgemeine Zugang zur Fehlerkontrolle unter Umständen bereits in der Literatur bekannte Feh-

lerschätzer² für Kontaktprobleme liefert. Zur Bestätigung dieser Vermutung sind weitere Untersuchungen erforderlich.

In dieser Arbeit werden Fehlerschätzer für die H^1 -Norm bzw. Energie-Norm vorgestellt. Darauf basierend werden ebenfalls Fehlerschätzer vom Goal-Oriented-Typ für allgemeine lineare Fehlerfunktionale diskutiert und numerische Experimente im Fall von Variationsgleichungen beschrieben. Bei der Herleitung dieser Fehlerkontrolltechnik für Variationsungleichungen besteht jedoch die Schwierigkeit einer angemessenen Fehlerkontrolle des Lagrangeschen Multiplikators. Für stückweise konstante Ansätze führen u.U. Average-Techniken zu sinnvollen Resultaten. Jedoch ist zu diskutieren, inwieweit derartige Ansätze auch für höhere Polynomgrade geeignet sind.

Abschließend sei bemerkt, dass der Einsatz von Finite-Elemente-Diskretisierungen mit höheren Polynomansätzen sowie die Analyse und Anwendung von hp -adaptiven Verfahren in Bezug auf Kontaktprobleme große Forschungsspielräume bieten. Die thematische Relevanz ist insbesondere dadurch gegeben, dass Ingenieur-technische Kontaktprobleme durch die Verwendung moderner Finite-Elemente-Implementierungen mit Modulen für Fehlerkontrolle und Adaptivität sowie mit flexiblen h - und hp -Diskretisierungsmöglichkeiten effizient gelöst werden können.

²vgl. z.B. [104]

Anhang A

Beispiele, Kopplungsgewichte, Grundlagen

A.1 h - und hp -adaptive Gitter für Variationsgleichungen

Beispiel mit geschlitztem Gebiet

In Abbildung A.1 sind h - und hp -adaptive Gitter bzw. FE-Gitter für das Modellproblem (II.4) zu sehen. Außerdem ist das zugehörige geschätzte Konvergenzverhalten abgebildet. Im Vergleich zu dem in Abschnitt IV.7.3 verwendeten Beispiel hat das zugrunde liegende Gebiet $\Omega \subset [-1, 1] \times [-0.5, 0.5]$ drei einspringende Ecken. Es ist außerdem $\Gamma_0 := \{-0.5\} \times [0, 0.5] \cup [-0.5, 0.5] \times \{0.5\} \cup \{0.5\} \times [0, 0.5]$, $f := 0$, sowie $q(x_0, x_1) := 0.2\pi \sin(\pi x_0)$ für $x_1 = -0.5$ und 0 sonst.

Insgesamt können die gleichen Beobachtungen wie im Abschnitt IV.7.3 festgestellt werden.

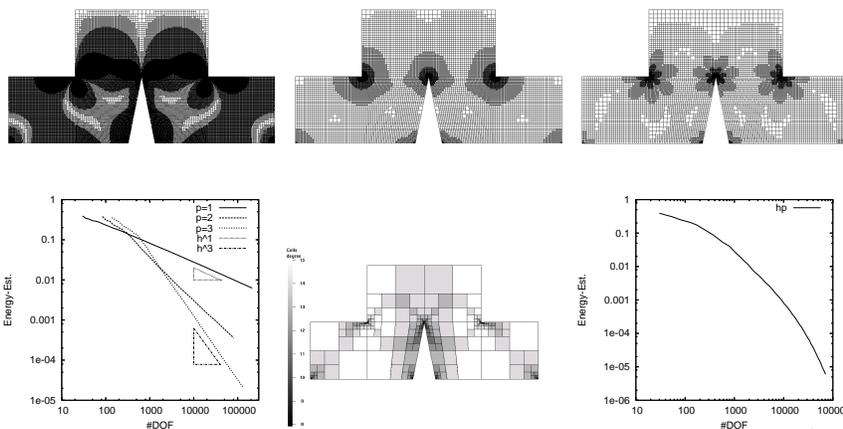


Abb. A.1: h -adaptive Gitter für $p = 1, 2, 3$ mit geschätztem Konvergenzverhalten (oben und links), hp -adaptive FE-Gitter mit geschätztem Konvergenzverhalten (unten rechts).

Interior-Layer-Beispiel

Betrachtet wird im Folgenden das Modellproblem (II.4) mit der analytischen Lösung $u := \arctan((x_1 - 2x_0)/\epsilon) - \pi/2$ auf $\Omega := (-1, 1)^2$. Dieses Beispiel wurde bereits in [4] hinsichtlich hp -adaptiver Verfahren untersucht. Für weitere Analysen bzgl. Interior-Layer-Probleme sei auf [90] verwiesen. Die Besonderheit dieses Beispiels besteht darin, dass sich die Lösung für einen relativ kleinen Wert $\epsilon \leq 0.05$ entlang einer dünnen Schicht im Innern des Gebiets (Interior-Layer) fast singulär verhält.

Zwar kann aus theoretischer Sicht eine optimale Konvergenz bereits bei globalen Verfeinerungen erwartet werden, solange aber das Interior-Layer durch das verwendete Gitter nicht hinreichend genau aufgelöst wird, ist das Konvergenzverhalten nicht optimal. Dieses Beispiel ist also besonders für den Einsatz adaptiver Verfahren prädestiniert.

Abbildung A.2 zeigt h -adaptive Gitter für die Polynomgrade $p = 1, 2, 3$ und hp -adaptive FE-Gitter sowie das daraus resultierende Konvergenzverhalten. Deutlich erkennbar ist, dass das Konvergenzverhalten zu Beginn des Verfeinerungsprozesses nicht optimal ist. Jedoch wird durch lokale Verfeinerungen bzw. zusätzlich durch eine hp -typische Polynomgradverteilung in der Umgebung des Interior-Layer die optimale Konvergenzordnung $\mathcal{O}(h^p)$ bzw. exponentielle Konvergenz erreicht.

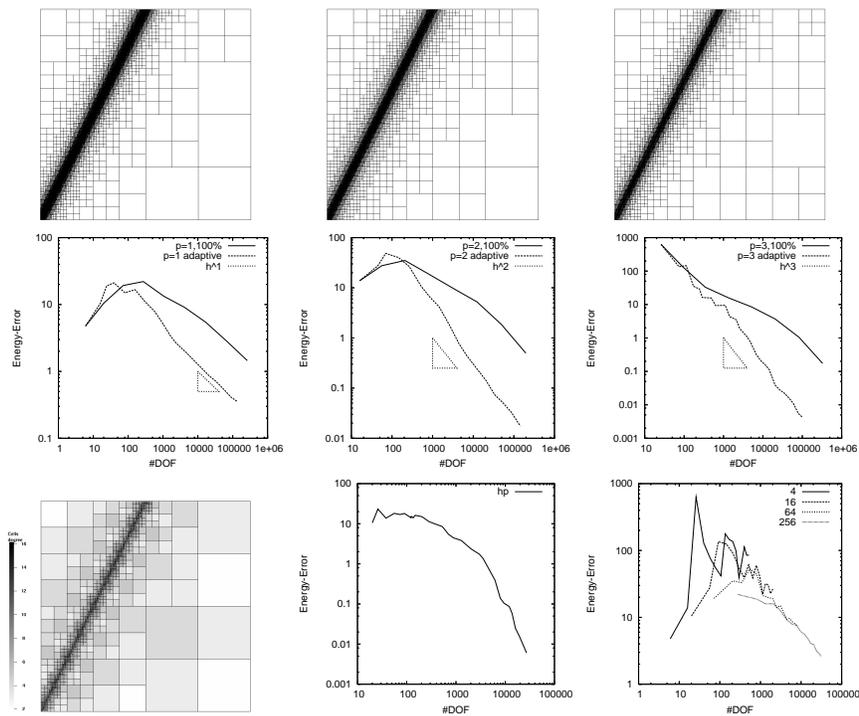


Abb. A.2: h -adaptive Gitter für $p = 1, 2, 3$ und Konvergenzverhalten, hp -adaptive FE-Gitter und Konvergenzverhalten (links unten), Konvergenzverhalten bei p -Verfeinerungen mit 4, 16, 64 und 256 Elementen (rechts unten).

In Abbildung A.2 (rechts unten) wird das Konvergenzverhalten bezüglich der reinen p -Methode gezeigt. Da u analytisch ist, kann theoretisch exponentielle Konvergenz erlangt werden. Bei ungenügender Auflösung (4 und 16 Gitterelemente) ist dies praktisch jedoch nicht der Fall. Erst bei 64 und 256 Gitterelementen kann exponentielle Konvergenz beobachtet werden.

Dreidimensionale Beispiele

In den Abbildungen A.3 und A.4 sind h - und hp -adaptive Gitter und das zugehörige geschätzte Konvergenzverhalten für das Modellproblem (II.4) und das linear-elastische Problem (II.38) mit einem jeweils dreidimensionalen Gebiet $\Omega \subset [-1, 1]^3$ bzw. $\Omega \subset [-3, 3] \times [-1, 1] \times [0, 4]$ abgebildet. In beiden Fällen hat das Gebiet Ω einspringende Ecken bzw. Kanten. Für das Mo-

dellproblem ist $f := 1$ und $\Gamma_0 := \partial\Omega$, für das linear-elastische Problem ist $f := (0.005, 0, -0.005)$ und $\Gamma_0 := [-3, -1]^2 \times \{0\} \cup [1, 3]^2 \times \{0\}$ sowie $E := 2$ und $\nu := 0.42$ gewählt.

Man erkennt, dass die erzeugten adaptiven Gitter alle typischen Merkmale von h - und hp -adaptiven Verfeinerungen aufweisen. Im Vergleich zu den dreidimensionalen Kontaktproblemen aus den Abschnitten IV.7.6 sind diese jedoch deutlich ausgeprägter.

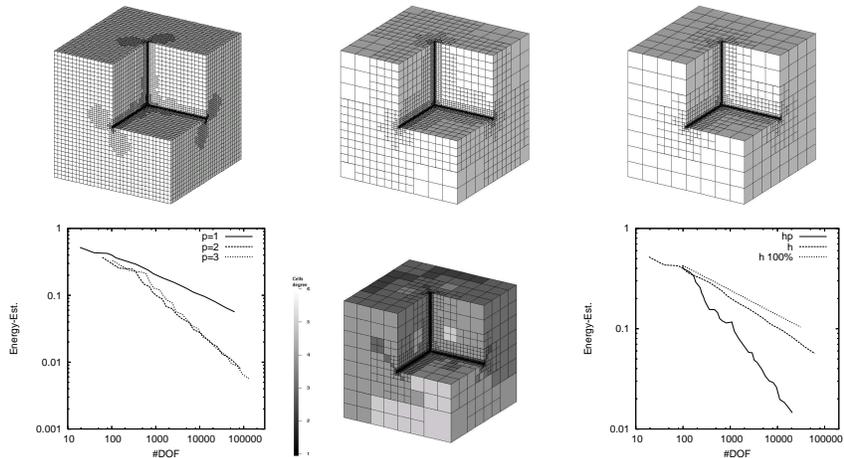


Abb. A.3: h -adaptive Gitter für $p = 1, 2, 3$ und geschätztes Konvergenzverhalten (oben und links), hp -adaptives FE-Gitter und geschätztes Konvergenzverhalten (unten und rechts).

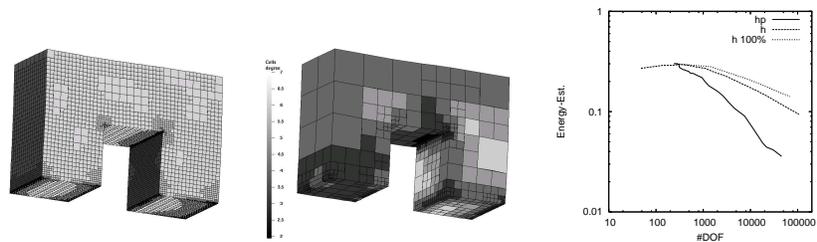


Abb. A.4: h -adaptive Gitter für $p = 1$ (links), hp -adaptive FE-Gitter (mitte) und geschätztes Konvergenzverhalten (rechts).

A.2 Kopplungsgewichte für hierarchische Basen

Die Kopplungsgewichte $\gamma_{d,c_0,\nu,c_1}^{h_0,r,h_1}$ aus Abschnitt III.2.2 gewinnt man, indem man Linearkombinationen der Basisfunktionen $\{\xi_{1,n}^q(-0.5 + 0.5x)\}$, die auf $[-1, 3]$ definiert sind, bezüglich der auf der ersten Teilungskante $[-1, 1]$ definierten Basisfunktionen $\{\xi_1^q(x)\}$ betrachtet. Hierbei ist

$$\{\xi_1^q(x)\} := \{0.5(1-x), 0.5(1+x)\} \cup \{G_{i+2}^{-1/2}(x)\}_{i=0,\dots,q-2}.$$

Satz A.1 *Es sei $\xi_{1,n}^n(-0.5 + 0.5x) = \sum_{i=0}^n \alpha_i^n \xi_{1,i}^n(x)$. Dann ist für $n \geq 2$:*

$$\begin{aligned} \alpha_0^0 &= 1, & \alpha_0^1 &= 0, & \alpha_0^2 &= 0, & \alpha_0^{n+1} &= 0, \\ \alpha_0^1 &= \frac{1}{2}, & \alpha_1^1 &= \frac{1}{2}, & \alpha_1^2 &= \frac{1}{2}, & \alpha_1^{n+1} &= -\frac{n-2}{n+1}\alpha_1^{n-1}, \\ & & & & \alpha_2^2 &= \frac{1}{4}, & \alpha_2^{n+2} &= \frac{2n-1}{2(n+1)} \left(\frac{1}{5}\alpha_3^n - \alpha_2^n + \alpha_0^n - \alpha_1^n \right) - \frac{n-2}{n+1}\alpha_2^{n-1}, \\ & & & & & & \alpha_i^{n+1} &= \frac{2n-1}{2(n+1)} \left(\frac{i}{2i-3}\alpha_{i-1}^n + \frac{i-1}{2i+1}\alpha_{i+1}^n - \alpha_i^n \right) \\ & & & & & & & - \frac{n-2}{n+1}\alpha_i^{n-1}, \quad i = 3, \dots, n-1, \\ & & & & & & \alpha_n^{n+1} &= \frac{2n-1}{2(n+1)} \left(\frac{n}{2n-3}\alpha_{n-1}^n - \alpha_n^n \right), \quad n > 2 \\ & & & & & & \alpha_{n+1}^{n+1} &= \frac{1}{2}\alpha_n^n. \end{aligned}$$

Beweis: Es sei $h_n(x) := \xi_{1,n}^n$. Aus Gleichung (III.59) erhält man für $n \geq 2$:

$$xh_n(x) = \frac{n+1}{2n-1}h_{n+1}(x) + \frac{n-2}{2n-1}h_{n-1}(x)$$

Daraus folgt:

$$\begin{aligned} & (n+1)h_{n+1}\left(-\frac{1}{2} + \frac{1}{2}x\right) \\ &= \frac{2n-1}{2}(-1+x)h_n\left(-\frac{1}{2} + \frac{1}{2}x\right) - (n-2)h_{n-1}\left(-\frac{1}{2} + \frac{1}{2}x\right) \\ &= -\frac{2n-1}{2}\left(\sum_{i=0}^n \alpha_i^n h_i(x)\right) + \frac{2n-1}{2}x\left(\sum_{i=0}^n \alpha_i^n h_i(x)\right) - (n-2)\sum_{i=0}^{n-1} \alpha_i^{n-1} h_i(x) \end{aligned}$$

$$\begin{aligned}
&= -\frac{2n-1}{2} \left(\sum_{i=0}^n \alpha_i^n h_i(x) \right) + \frac{2n-1}{2} \left(\sum_{i=2}^n \alpha_i^n \left(\frac{i+1}{2i-1} h_{i+1}(x) + \frac{i-2}{2i-1} h_{i-1}(x) \right) \right) \\
&\quad + \frac{2n-1}{2} \left(\alpha_0^n (-h_0(x) + h_2(x)) + \alpha_1^n (h_1(x) - h_2(x)) - (n-2) \sum_{i=0}^{n-1} h_i(x) \right) \\
&= \frac{n+1}{2} \alpha_n^n h_{n+1}(x) + \left(\frac{2n-1}{2} \cdot \frac{n}{2n-3} \alpha_{n-1}^n - \frac{2n-1}{2} \alpha_n^n \right) h_n(x) \\
&\quad + \frac{2n-1}{2} \sum_{i=3}^{n-1} \alpha_{i-1}^n \frac{i}{2i-3} h_i(x) + \frac{2n-1}{2} \sum_{i=2}^{n-1} \alpha_{i+1}^n \frac{i-1}{2i+1} h_i(x) \\
&\quad - \frac{2n-1}{2} \sum_{i=0}^{n-1} \alpha_i^n h_i(x) - (n-2) \sum_{i=0}^{n-1} \alpha_i^{n-1} h_i(x) + \frac{2n-1}{2} (\alpha_0^n - \alpha_1^n) h_2(x) \\
&\quad + \frac{2n-1}{2} \alpha_1^n h_1(x) - \frac{2n-1}{2} \alpha_0^n h_0(x) \\
&= \frac{n+1}{2} \alpha_n^n h_{n+1}(x) + \left(\frac{2n-1}{2} \cdot \frac{n}{2n-3} \alpha_{n-1}^n - \frac{2n-1}{2} \alpha_n^n \right) h_n(x) \\
&\quad + \sum_{i=3}^{n-1} \left(\frac{2n-1}{2} \cdot \frac{i}{2i-3} \alpha_{i-1}^n + \frac{2n-1}{2} \cdot \frac{i-1}{2i+1} \alpha_{i+1}^n - \frac{2n-1}{2} \alpha_i^n - (n-2) \alpha_i^{n-1} \right) h_i(x) \\
&\quad + \left(\frac{2n-1}{2} \cdot \frac{1}{5} \alpha_3^n - \frac{2n-1}{2} \alpha_2^n - (n-2) \alpha_2^{n-1} + \frac{2n-1}{2} (\alpha_0^n - \alpha_1^n) \right) h_2(x) \\
&\quad + \left(-\frac{2n-1}{2} \alpha_1^n - (n-2) \alpha_1^{n-1} + \frac{2n-1}{2} \alpha_1^n \right) h_1(x) \\
&\quad + \left(-\frac{2n-1}{2} \alpha_0^n - (n-2) \alpha_0^{n-1} - \frac{2n-1}{2} \alpha_0^n \right) h_0(x)
\end{aligned}$$

Division mit $n+1$ liefert dann die Behauptung. \square

Im Folgenden sei

$$z := \begin{cases} (2 + c_1, 0), & c_0 = 1, \nu = 0 \\ (1, 2 + c_1), & c_0 = 1, \nu = 1 \\ (2 + c_1, 1), & c_0 = 1, \nu = 2 \\ (0, 2 + c_1), & c_0 = 1, \nu = 3 \\ (2 + i, 2 + j), & c_0 = 2, c_1 = \alpha(i, j), \end{cases}$$

$$\chi(r', c, h) := \begin{cases} 1, & r = r' \vee (c \neq 2 = h \neq 2) \\ -1, & \text{sonst,} \end{cases}$$

$$\tilde{\chi}(r'_0, r'_1, c, h) := \begin{cases} 1, & r = r'_0 \vee \chi(r'_1, c, h) = 1 \\ -1, & \text{sonst.} \end{cases}$$

Nach einer längeren Argumentation, die über den Rahmen dieses Anhangs hinausginge, erhält man die gewünschten Kopplungsgewichte:

$$\gamma_{d,c_0,\nu,c_1}^{h_0,r,h_1} := \begin{cases} \frac{1}{2}, & d = 1, c_0 = h_0 = 0 \\ \chi(0, c_1, h_1 + 1) \alpha_0^{2+c_1}, & d = 1, c_0 = 1, h_0 = 0 \\ \chi(0, c_1, h_1) \alpha_{2+h_1}^{2+c_1}, & d = 1, c_0 = 1, h_0 = 1, \\ \frac{1}{4}, & c_0 = h_0 = 0 \\ \alpha_1^{z_0} \alpha_1^{z_1}, & d = 2, c_0 > 0, h_0 = 0 \\ \chi(0, z_1, h_1) \alpha_1^{z_0} \alpha_{2+h_1}^{z_1}, & c_0 > 0, h_0 = 1, r \in \{0, 2\} \\ \chi(3, z_0, h_1) \alpha_{2+h_1}^{z_0} \alpha_1^{z_1}, & c_0 > 0, h_0 = 1, r \in \{1, 3\} \\ \tilde{\chi}(0, 3, z_0, i) \tilde{\chi}(1, 3, z_1, j) \alpha_{2+i}^{z_0} \alpha_{2+j}^{z_1}, & c_0 = h_0 = 2, \alpha(i, j) = c_1. \end{cases}$$

Die oben aufgeführten Kopplungsgewichte wurden entsprechend im FEM-Paket SOFAR [87] implementiert und überprüft.

Der Anwendungsbereich derartiger Kopplungsgewichte erstreckt sich nicht nur auf die Behandlung von FEM-Diskretisierungen mit irregulären Gittern, es lassen sich hiermit auch Prolongations- und Restriktionsoperatoren definieren, die insbesondere für Mehrgitterlösungsverfahren eingesetzt werden können. Diesbezüglich sei ebenfalls auf [87] verwiesen.

A.3 Funktionalanalytische Grundlagen

Normierte Räume

Satz A.2 *Es sei V ein normierter Raum. Dann ist $F : V \rightarrow \overline{\mathbb{R}}$ genau dann (schwach) unterhalb stetig, wenn $\text{epi}(F)$ (schwach) abgeschlossen ist.*

Beweis: Prop.I.2.3 in [49]. \square

Satz A.3 *Es sei V ein reflexiver Raum. Dann besitzt jede beschränkte Folge eine schwach konvergente Teilfolge.*

Beweis: Th.III.3.7 in [106]. \square

Satz A.4 (Trennungssatz von Hahn-Banach) *Es sei V ein normierter Raum, $K \subset V$ sei abgeschlossen und konvex, und es sei $v \notin K$. Dann existiert ein $\mu \in V'$ mit $\langle \mu, v \rangle < \inf_{w \in K} \langle \mu, w \rangle$.*

Beweis: Th.III.2.5 in [106]. \square

Satz A.5 *Es sei V ein normierter Raum und $K \subset V$ abgeschlossen und konvex. Dann ist K schwach abgeschlossen.*

Beweis: Es sei $\{v_n\} \subset K$ mit $v_n \rightarrow v \in V$. Angenommen $v \notin K$, dann existiert ein $\mu \in V'$ mit

$$\langle \mu, v \rangle < \inf_{w \in K} \langle \mu, w \rangle \leq \liminf_{n \rightarrow \infty} \langle \mu, v_n \rangle = \lim_{n \rightarrow \infty} \langle \mu, v_n \rangle = \langle \mu, v \rangle.$$

Widerspruch! \square

Satz A.6 *Es seien V ein normierter Raum, $G \subset V$ ein abgeschlossener und konvexer Kegel und $G' := \{\mu \in V' \mid \forall v \in G : \langle \mu, v \rangle \geq 0\}$. Dann gilt:*

$$v \in G \Leftrightarrow \forall \mu \in G' : \langle \mu, v \rangle \geq 0.$$

Beweis: Es sei $v \in V$ und für alle $\mu \in G'$ sei $\langle \mu, v \rangle \geq 0$. Angenommen, $v \notin G$, dann existiert nach dem Trennungssatz von Hahn-Banach A.4 ein $\tilde{\mu} \in V'$ mit

$$\langle \tilde{\mu}, v \rangle < \inf_{w \in G} \langle \tilde{\mu}, w \rangle. \quad (\text{A.1})$$

Da $0 \in G$ ist, folgt

$$\langle \tilde{\mu}, v \rangle < 0. \quad (\text{A.2})$$

Für $t \geq 0$ und $w \in G$ ist $tw \in G$. Angenommen, $\inf_{w \in G} \langle \tilde{\mu}, w \rangle < 0$, dann folgt $\inf_{w \in G} \langle \tilde{\mu}, tw \rangle = t \inf_{w \in G} \langle \tilde{\mu}, w \rangle \rightarrow -\infty$ für $t \rightarrow \infty$ im Widerspruch zu (A.1). Also muss $\tilde{\mu} \in G'$ sein, im Widerspruch zu (A.2). \square

Satz A.7 *Es sei V ein reflexiver Banachraum. Ferner seien $M \subset V$ und $M \subset V'$ Unterräume. Dann ist $(M^\perp)^\perp = \overline{M}$ und $(N_\perp)^\perp = \overline{N}$.*

Beweis: Da $(N_\perp)^\perp$ abgeschlossen ist, ist $\overline{N} \subset (N_\perp)^\perp$. Es sei $\phi \notin \overline{N}$, dann existiert nach dem Trennungssatz A.4 von Hahn-Banach ein $v \in V \simeq V''$, so dass $v \in N_\perp$ und $\langle \phi, v \rangle < \inf_{\mu \in \overline{N}} \langle \mu, v \rangle \leq 0$ ist. Also ist $\phi \notin (N_\perp)^\perp$. Der Beweis von $(M^\perp)^\perp = \overline{M}$ ist entsprechend. \square

Satz A.8 *Es sei V ein normierter Raum, $U \subset V$ sei ein abgeschlossener Unterraum. Dann ist U' isometrisch isomorph zu V'/U^\perp .*

Beweis: Die Abbildung $\phi + U^\perp \mapsto \phi|_U$ ist ein isometrischer Isomorphismus von V'/U^\perp nach U' .¹ \square

Satz A.9 *Es seien V_0 und V_1 Banachräume mit $V_1 \subset V_0$. Für $0 < t < \infty$ sei $K(t, v) := \inf_{v=v_0+v_1} \|v_0\|_{V_0} + t\|v_1\|_{V_1}$, und für $0 < \theta < 1$ und $1 \leq q \leq \infty$ sei $\|v\|^q := \int_0^\infty (t^{-\theta} K(t, v))^q t^{-1} dt$. Dann ist $[V_0, V_1]_{\theta, q} := \{v \in V_0 \mid \|v\| < \infty\}$ ein Banachraum, und es ist*

$$\forall v \in [V_0, V_1]_{\theta, q} : \|v\| \leq C(\theta, q) \|v\|_{V_0}^{1-\theta} \|v\|_{V_1}^\theta.$$

Beweis: Ch. 1.3.3 in [103]. \square

¹vgl. Th.4.9 in [86]

Operatoren

Satz A.10 *Es seien V und W Banachräume, $B \in L(V, W)$ und $\text{rg}(B)$ sei abgeschlossen. Dann existiert ein $\beta > 0$, so dass gilt:*

$$\forall w \in \text{rg}(B) : \exists v \in V : B(v) = w, \|v\| \leq \beta \|w\|.$$

Beweis: Lemma IV.5.2 in [106]. □

Satz A.11 (Satz vom abgeschlossenen Bild) *Es seien V und W Banachräume und $B \in L(V, W)$. Dann sind äquivalent: $\text{rg}(B)$ ist abgeschlossen, $\text{rg}(B) = \ker B'_\perp$, $\text{rg}(B')$ ist abgeschlossen, $\text{rg}(B') = (\ker B)^\perp$.*

Beweis: S.205 in [107]. □

Satz A.12 *Es seien V und W Banachräume, $U \subset V$ ein Unterraum und $B : V \supset U \rightarrow W$ abgeschlossen. Dann ist U versehen mit der Norm $\|\cdot\|_U^2 := \|\cdot\|_V^2 + \|B(\cdot)\|_W^2$ ein Banachraum, und es ist $B \in L(U, W)$.*

Beweis: Lemma IV.4.3 in [106] □

Satz A.13 (Fortsetzungssatz von Hahn-Banach) *Es sei W ein linearer Raum und $V \subset W$ ein Unterraum. Ferner sei $p : W \rightarrow \mathbb{R}$ sublinear und $l : V \rightarrow \mathbb{R}$ linear mit*

$$\forall v \in V : l(v) \leq p(v).$$

Dann existiert eine lineare Fortsetzung $L : W \rightarrow \mathbb{R}$, $L|_V = l$, mit

$$\forall w \in W : L(w) \leq p(w).$$

Beweis: Satz III.1.2 in [106]. □

Satz A.14 *Es seien W ein normierter Raum und $V \subset W$ ein Unterraum. Dann ist V genau dann dicht in W , wenn aus $\mu \in W'$ und $\mu|_V = 0$ folgt, dass $\mu = 0$ ist.*

Beweis: Kor.III.1.9 in [106]. □

Literaturverzeichnis

- [1] R. A. Adams. *Sobolev spaces*. Pure and Applied Mathematics. Academic Press, New York, 1975.
- [2] M. Ainsworth, J. T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics. Wiley-Interscience, Chichester, 2000.
- [3] M. Ainsworth, B. Senior. Aspects of an adaptive hp -finite element method: Adaptive strategy, conforming approximation and efficient solvers. *Comput. Methods Appl. Mech. Eng.*, 150(1-4):65–87, 1997.
- [4] M. Ainsworth, B. Senior. An adaptive refinement strategy for hp -finite element computations. *Appl. Numer. Math.*, 26(1-2):165–178, 1998.
- [5] J. P. Aubin. *Approximation of elliptic boundary-value problems*. Pure and Applied Mathematics. Wiley-Interscience., New York, 1972.
- [6] I. Babuška, A. Craig, J. Mandel, J. Pitkäranta. Efficient preconditioning for the p -version finite element method in two dimensions. *SIAM J. Numer. Anal.*, 28(3):624–661, 1991.
- [7] I. Babuska, J. P. de S.R.Gago, D. W. Kelly, O. C. Zienkiewicz. A posteriori error analysis and adaptive processes in the finite element method. ii. adaptive mesh refinement. *Int. J. Numer. Methods Eng.*, 19:1621–1656, 1983.
- [8] I. Babuska, H. C. Elman. Performance of the $h - p$ version of the finite element method with various elements. *Int. J. Numer. Methods Eng.*, 36(15):2503–2523, 1993.

- [9] I. Babuska, W. Gui. The h , p and h - p versions of the finite element method in 1 dimension. I,II,III. The error analysis of the p -version. *Numer. Math.*, 49:577–683, 1986.
- [10] I. Babuska, B. Guo. The h - p version of the finite element method. i. the basic approximation results. *Comput. Mech.*, 1:21–41, 1986.
- [11] I. Babuska, B. Guo. The h - p version of the finite element method. ii. general results and applications. *Comput. Mech.*, 1:203–220, 1986.
- [12] I. Babuska, B. Q. Guo. The h - p version of the finite element method for domains with curved boundaries. *SIAM J. Numer. Anal.*, 25(4):837–861, 1988.
- [13] I. Babuska, R. B. Kellogg, J. Pitkäranta. Direct and inverse error estimates for finite elements with mesh refinements. *Numer. Math.*, 33:447–471, 1979.
- [14] I. Babuska, W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15:736–754, 1978.
- [15] I. Babuska, T. Strouboulis, K. Copps. hp optimization of finite element approximations: Analysis of the optimal mesh sequences in one dimension. *Comput. Methods Appl. Mech. Eng.*, 150(1-4):89–108, 1997.
- [16] I. Babuska, M. Suri. The h - p version of the finite element method with quasiouniform meshes. *M²AN*, 21:199–238, 1987.
- [17] I. Babuska, M. Suri. The p - and h - p versions of the finite element method. an overview. *Comput. Methods Appl. Mech. Eng.*, 80(1-3):5–26, 1990.
- [18] I. Babuska, M. Suri. The p and $h - p$ versions of the finite element method, basic principles and properties. *SIAM Rev.*, 36(4):578–632, 1994.
- [19] W. Bangerth, R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics, ETH Zürich. Birkhäuser, Basel, 2003.
- [20] K. J. Bathe. *Finite-Elemente-Methoden*. Springer-Verlag, Berlin, 2002.

- [21] R. Becker, R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East-west J. Numer. Math.*, 4:237–264, 1996.
- [22] R. Becker, R. Rannacher. Weighted a posteriori error control in fe methods. In *Bock, Hans Georg (ed.) et al., ENUMATH 97. Proceedings of the 2nd European conference on numerical mathematics and advanced applications held in Heidelberg, Germany, September 28-October 3, 1997. Including a selection of papers from the 1st conference (ENUMATH 95) held in Paris, France, September 1995.*, pages 621–637. World Scientific, Singapore, 1998.
- [23] H. Blum. Numerical treatment of corner and crack singularities. In *Finite element and boundary element techniques from mathematical and engineering point of view*, volume 301 of *CISM Courses Lect.*, pages 171–212. Springer-Verlag, New-York, 1988.
- [24] H. Blum, D. Braess, F. T. Suttmeier. A cascadic multigrid algorithm for variational inequalities. *Comput. Vis. Sci.*, 7(3-4):153–157, 2004.
- [25] D. Braess. *Finite Elemente. Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer-Verlag, Berlin, 1997.
- [26] D. Braess. A posteriori error estimators for obstacle problems - another look. *Numer. Math.*, 101(3):415-421, 2005.
- [27] S. C. Brenner, L. R. Scott. *The mathematical theory of finite element methods*. Texts in Applied Mathematics. Springer-Verlag, New York, 1994.
- [28] F. Brezzi, W. W. Hager, P. A. Raviart. Error estimates for the finite element solution of variational inequalities. part i. primal theory. *Numer. Math.*, 28:431–443, 1977.
- [29] F. Brezzi, W. W. Hager, P. A. Raviart. Error estimates for the finite element solution of variational inequalities. part ii. mixed methods. *Numer. Math.*, 31:1–16, 1978.
- [30] M. Cabaravdic, K. Kneupner, B. Kuhlenkoetter, W. Kreis, T. Schueppstuhl, X. Zhang. Belt-grinding models for sculptured surfaces. In *Conference Proceedings of Simulation Aided Offline Process Design and Optimization in Manufacturing Sculptured Surface*, pages 31–38, Dortmund, 2003. ISF, Universität Dortmund.

- [31] M. Cabaravdic, K. Kneupner, B. Kuhlenkoetter, A. Schröder. Force-distribution model by the belt grinding with elastic contact wheel. In *8th. International Research/Expert Conference "Trends in the Development of Machinery and Associated Technology"*, 2004.
- [32] C. Carstensen, S. Bartels. Each averaging technique yields reliable a posteriori error control in fem on unstructured grids. i: Low order conforming, nonconforming, and mixed fem. *Math. Comput.*, 71(239):945–969, 2002.
- [33] C. Carstensen, S. Bartels. Each averaging technique yields reliable a posteriori error control in fem on unstructured grids. ii: Higher order fem. *Math. Comput.*, 71(239):971–994, 2002.
- [34] J. Cea. *Lectures on optimization - theory and algorithms*. Tata Institute of Fundamental Research Lectures on Mathematics and Physics. Mathematics. Springer-Verlag, Berlin, 1978.
- [35] A. B. Chaudhary, K. J. Bathe. A solution method for static and dynamic analysis of three-dimensional contact problems with friction. *Comput. Struct.*, 24:855–873, 1986.
- [36] Ph. Clément. Approximation by finite element functions using local regularization. *R.A.I.R.O.*, 9:77–84, 1975.
- [37] A. Czekanski, S.A. Meguid. Analysis of dynamic frictional contact problems using variational inequalities. *Finite Elem. Anal. Des.*, 37(11):861–879, 2001.
- [38] A. Czekanski, S.A. Meguid, N. El-Abbasi, M.H. Refaat. On the elastodynamic solution of frictional contact problems using variational inequalities. *Int. J. Numer. Methods Eng.*, 50(3):611–627, 2001.
- [39] M. Dauge. *Elliptic boundary value problems on corner domains. Smoothness and asymptotics of solutions*. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1988.
- [40] T. A. Davis. Umfpack version 4.4 user guide. Technical report, Dept. of Computer and Information Science and Engineering, Univ. of Florida, 2005.
- [41] L. Demkowicz, J. T. Oden, W. Rachowicz, O. Hardy. Toward a universal h-p adaptive finite element strategy. i: Constrained approximation and data structure. *Comp. Meth. Appl. Mech. Engrg.*, 77:79–112, 1989.

- [42] L. Demkowicz, W. Rachowicz, Ph. Devloo. A fully automatic *hp*-adaptivity. *J. Sci. Comput.*, 17(1-4):117–142, 2002.
- [43] Din 8580, Fertigungsverfahren - Begriffe, Einteilung., 2003.
- [44] Din 8589, Fertigungsverfahren Spanen - Teil 12: Bandschleifen; Einordnung, Unterteilung., 2003.
- [45] W. Domke. *Werkstoffkunde und Werkstoffprüfung*. Cornelsen Verlag, 1987.
- [46] Z. Dostál, J. Haslinger, R. Kučera. Implementation of the fixed point method in contact problems with coulomb friction based on a dual splitting type technique. *J. Comput. Appl. Math.*, 140(1-2):245–256, 2002.
- [47] G. Duvaut, J. L. Lions. *Inequalities in mechanics and physics*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1976.
- [48] H. Ehlich, K. Zeller. Schwankungen von Polynomen zwischen Gitterpunkten. *Math. Z.*, 86:41–44, 1964.
- [49] I. Ekeland, R. Temam. *Convex analysis and variational problems*. Studies in Mathematics and its Applications. North-Holland Publishing Company, Amsterdam, 1976.
- [50] C. Geiger, Ch. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer-Verlag, Berlin, 2002.
- [51] P.E. Gill, W. Murray, M. A. Saunders. User's guide for sqopt 5.3: A fortran package for large-scale linear and quadratic programming. Technical report, Department of Mathematics, University of California, San Diego, 1999.
- [52] P.E. Gill, W. Murray, M. A. Saunders. User's guide for snopt version 6: A fortran package for large-scale nonlinear programming. Technical report, Department of Mathematics, University of California, San Diego, 2002.
- [53] Ph. E. Gill, W. Murray, M. A. Saunders. Snopt: An sqp algorithm for large-scale constrained optimization. *SIAM J. Optim.*, 12(4):979–1006, 2002.

- [54] Ph. E. Gill, W. Murray, M. H. Wright. *Practical optimization*. Academic Press, London, 1981.
- [55] R. Glowinski, J. L. Lions, R. Trémoierès. *Numerical analysis of variational inequalities*. Studies in Mathematics and its Applications. North-Holland Publishing Company, Amsterdam, 1981.
- [56] Roland Glowinski. *Numerical methods for nonlinear variational problems*. Springer Series in Computational Physics. Springer-Verlag, New York, 1984.
- [57] Ch. Großmann, H. G. Roos. *Numerik partieller Differentialgleichungen*. Teubner Studienbücher: Mathematik. B. G. Teubner, Stuttgart, 1992.
- [58] G. Grube, M. Heetmann, T. Schüppstuhl, S. Zinke. *Offenlegungsschrift DE 10013340 A1*. Deutsches Patent- und Markenamt, 2001.
- [59] G. Grube, M. Heetmann, T. Schüppstuhl, S. Zinke. *Patentschrift DE 10013340 C2*. Deutsches Patent- und Markenamt, 2003.
- [60] W. Gui, I. Babuska. The h, p and h-p versions of the finite element method in 1 dimension. iii. the adaptive h-p version. *Numer. Math.*, 49:659–683, 1986.
- [61] W. Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen*. Teubner Studienbücher: Mathematik. B. G. Teubner, Stuttgart, 1996.
- [62] J. Haslinger. Mixed formulation of elliptic variational inequalities and its approximation. *Apl. Mat.*, 26:462–475, 1981.
- [63] J. Haslinger, Z. Dostál, R. Kučera. On a splitting type algorithm for the numerical realization of contact problems with coulomb friction. *Comput. Methods Appl. Mech. Eng.*, 191(21-22):2261–2281, 2002.
- [64] J. Haslinger, I. Hlaváček. Approximation of the signorini problem with friction by a mixed finite element method. *J. Math. Anal. Appl.*, 86:99–122, 1982.
- [65] J. Haslinger, J. Lovíšek. Mixed variational formulation of unilateral problems. *Commentat. Math. Univ. Carol.*, 21:231–246, 1980.

- [66] J. Haslinger, T. Sassi. Mixed finite element approximation of 3d contact problems with given friction: error analysis and numerical realization. *Math. Mod. Numer. Anal.*, 38:563–578, 2004.
- [67] J. Haslinger, M. Tvrđy. Approximation and numerical solution of contact problems with friction. *Apl. Mat.*, 28:55–71, 1983.
- [68] I. Hlaváček, J. Haslinger, J. Nečas, J. Lovsek. *Solution of variational inequalities in mechanics*. Applied Mathematical Sciences. Springer-Verlag, New York, 1988.
- [69] E. Hornbogen. *Werkstoffe*. Springer-Verlag, 1991.
- [70] F. Jarre, J. Stoer. *Optimierung*. Springer-Verlag, Berlin, 2004.
- [71] J. Jarušek. Contact problems with bounded friction. coercive case. *Czechoslovak Mathematical Journal*, 33:237–261, 1983.
- [72] N. Kikuchi, J.T. Oden. *Contact problems in elasticity: A study of variational inequalities and finite element methods*. SIAM Studies in Applied Mathematics. SIAM, Society for Industrial and Applied Mathematics, Philadelphia, 1988.
- [73] P. Knabner, L. Angermann. *Numerik partieller Differentialgleichungen. Eine anwendungsorientierte Einführung*. Springer-Verlag, Berlin, 2000.
- [74] A. Krebs. *On Solving Nonlinear Variational Inequalities by p -Version Finite Elements*. PhD thesis, Fachbereich Mathematik der Universität Hannover, 2004.
- [75] A. Kroó, S. Révész. On bernstein and markov-type inequalities for multivariate polynomials on convex bodies. *J. Approximation Theory*, 99(1):134–152, 1999.
- [76] J. L. Lions, E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I. Translated from the French by P. Kenneth*. Die Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1972.
- [77] J. F. Maitre, O. Pourquier. About the conditioning of matrices in the p -version of the finite element method for second order elliptic problems. *J. Comput. Appl. Math.*, 63(1-3):341–348, 1995.

- [78] J. E. Marsden, T. J. R. Hughes. *Mathematical foundations of elasticity*. Prentice-Hall Civil Engineering and Engineering Mechanics Series. Prentice-Hall, New Jersey, 1983.
- [79] G. Matthies, L. Tobiska. The inf-sup condition for the mapped $q_k - p_{k-1}^{\text{disc}}$ element in arbitrary space dimensions. *Computing*, 69(2):119–139, 2002.
- [80] J.M. Melenk. hp-interpolation of nonsmooth functions and an application to hp-a posteriori error estimation. *SIAM J. Numer. Anal.*, 43(1):127–155, 2005.
- [81] J.M. Melenk, B.I. Wohlmuth. On residual-based a posteriori error estimation in hp-fem. *Adv. Comput. Math.*, 15(1-4):311–331, 2001.
- [82] J. Nečas, J. Jarušek, J. Haslinger. On the solution of the variational inequality to the signorini problem with small friction. *Boll. Unione Mat. Ital., V. Ser., B*, 17:796–811, 1980.
- [83] J. Nocedal, St. J. Wright. *Numerical optimization*. Springer Series in Operations Research. Springer-Verlag, New York, 1999.
- [84] A. Patra, A. Gupta. A systematic strategy for simultaneous adaptive hp finite element mesh modification using nonlinear programming. *Comput. Methods Appl. Mech. Eng.*, 190(29-30):3797–3818, 2001.
- [85] W. Rachowicz, J. T. Oden, L. Demkowicz. Toward a universal h-p adaptive finite element strategy. iii: Design of h-p meshes. *Comp. Meth. Appl. Mech. Engrg.*, 77:181–212, 1989.
- [86] W. Rudin. *Functional analysis*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Comp, New York, 1973.
- [87] A. Schröder, H. Blum. Sofar: Small object oriented finite element library for application and research. Technical report, Lehrstuhl Wissenschaftliches Rechnen, Fachbereich Mathematik, Universität Dortmund, 2002.
- [88] T. Schueppstuhl. *Beitrag zum Bandschleifen komplexer Freiformgeometrien mit Industrieroboter*. Shaker-Verlag, 2002.
- [89] Ch. Schwab. *p- and hp-finite element methods. Theory and applications in solid and fluid mechanics*. Numerical Mathematics and Scientific Computation. Clarendon Press, Oxford, 1998.

- [90] Ch. Schwab, M. Suri. The p and hp versions of the finite element method for problems with boundary layers. *Math. Comput.*, 65(216):1403–1429, 1996.
- [91] M. Seager. A slap for the masses. In *Parallel Supercomputing: Methods, Algorithms and Applications*, pages 135–155. Wiley Parallel Computing Series, 1989.
- [92] P. Šolin, K. Segeth, I. Deležel. *Higher-order finite element methods*. Studies in Advanced Mathematics. CRC Press, Boca Raton, 2004.
- [93] E. Süli, P. Houston, Ch. Schwab. hp -finite element methods for hyperbolic problems. In *Whiteman, J. R. (ed.), The mathematics of finite elements and applications X, MAFELAP 1999. Proceedings of the 10th conference, Brunel Univ., Uxbridge, Middlesex, GB, June 22-25, 1999.*, pages 143–162. Elsevier, Amsterdam, 2000.
- [94] F. T. Suttmeier. Consistent error estimation of fe-approximations of variational inequalities. Technical Report 261, Ergebnisberichte Angewandte Mathematik, Univ. Dortmund, 2004.
- [95] F. T. Suttmeier. Finite element analysis of the torsion problem. submitted, 2004.
- [96] F. T. Suttmeier. Goal-oriented error-estimation and postprocessing for fe-discretisations of variational inequalities. submitted, 2004.
- [97] F.T. Suttmeier. Error analysis for finite element solutions of variational inequalities. Technical report, Universität Dortmund, 2001. Professorial dissertation for obtaining the *venia legendi* at the Fachbereich Mathematik.
- [98] F.T. Suttmeier. Reliable, goal-oriented postprocessing for fe-discretizations. *Numer. Methods Partial Differ. Equations*, 21(2):387–396, 2005.
- [99] B. Szabó, I. Babuska. *Finite element analysis*. Wiley-Interscience Publication. John Wiley & Sons Ltd., New York, 1991.
- [100] B. A. Szabó. Mesh design for the p -version of the finite element method. *Comput. Methods Appl. Mech. Eng.*, 55:181–197, 1986.

- [101] B. A. Szabó. The p- and h-p versions of the finite element method in solid mechanics. *Comput. Methods Appl. Mech. Eng.*, 80(1-3):185–195, 1990.
- [102] F. G. Tricomi. *Vorlesungen über Orthogonalreihen*. Die Grundlehre der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1955.
- [103] H. Triebel. *Interpolation theory, function spaces, differential operators*. North-Holland Mathematical Library. North-Holland Publishing Company, Amsterdam, 1978.
- [104] A. Veiser. Efficient and reliable a posteriori error estimators for elliptic obstacle problems. *SIAM J. Numer. Anal.*, 39(1):146–167, 2001.
- [105] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley-Teubner Series Advances in Numerical Mathematics. John Wiley & Sons, B. G. Teubner, Chichester, Stuttgart, 1996.
- [106] D. Werner. *Funktionalanalysis*. Springer-Verlag, Berlin, 1995.
- [107] K. Yosida. *Functional analysis*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, Berlin, 1980.
- [108] X. Zhang, M. Cabaravdic, K. Kneupner, Kuhlenkoetter B. Real-time simulation of robot controlled belt grinding processes of sculptured surfaces. *International Journal of Advanced Robotic Systems*, 1(2):109–114, 2004.