Konvergenz des Proximal–Punkt–Verfahrens für inkorrekt gestellte Optimalsteuerprobleme mit partiellen Differentialgleichungen

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften

Dem Fachbereich IV der Universität Trier
vorgelegt von

Sergej Rotin

Trier, 1999

1. Gutachter: Prof. Dr. R. Tichatschke
2. Gutachter: Prof. Dr. F. Tröltzsch

Tag der Einreichung: 19.02.1999
Tag der mündlichen Prüfung: 24.03.1999

Inhaltsverzeichnis

1 Einleitung **1**

2 Regularisierte Strafverfahren in Hilberträumen **7**
2.1 Regularisierte Strafverfahren für das abstrakte Kontrollproblem . . . 8
2.2 Einige vorbereitende Resultate 12
2.3 Konvergenzsätze für Tichonov- und Proximal-Punkt-Verfahren 29
2.4 Schlußbemerkungen 37

3 Probleme mit unterschiedlich ausgeprägter Inkorrektheit **39**
3.1 Elliptische Operatoren in Sobolevräumen 40
3.2 Ein inkorrekt gestelltes Steuerproblem mit einem elliptischen Operator 2-ter Ordnung 53
3.3 Ein inkorrekt gestelltes Steuerproblem mit einem elliptischen Operator 1-ter Ordnung 75

4 Konvergenz innerer Approximationen im MSR-Verfahren **93**
4.1 Approximation der elliptischen Gleichung 2-ter Ordnung mit gemischten Randbedingungen 94
4.2 Approximation des linear-quadratischen Optimalsteuerproblems mit elliptischer Zustandsgleichung 104
4.3 Das Konvergenzkriterium der Näherungen im MSR-Verfahren 108
4.4 Numerische Behandlung der diskretisierten Ersatzprobleme 123

A **133**

B **137**

Literaturverzeichnis **143**

Kapitel 1

Einleitung

Die Entwicklung der Theorie der Extremalprobleme zeigt immer wieder die Notwendigkeit der systematischen Untersuchung von inkorrekt gestellten Aufgaben und zugehöriger numerischer Verfahren [40], [93]. Von Anfang an war es offensichtlich, daß die inkorrekten Probleme in den meisten Klassen der Variationsaufgaben auftreten. Ebenso existieren viele Variationsungleichungen und Optimalsteuerprobleme aus verschiedenen Bereichen von Physik, Technik und Ökonomie, die inkorrekt gestellt sind.

Gegenstand unserer Betrachtung sind schlecht gestellte Probleme der mathematischen Physik und insbesondere schlecht gestellte Optimalsteuerbropleme mit partiellen Differentialgleichungen. Erheblich hierbei ist, daß wir Kontrollprobleme mit aktiven Steuer- und Zustandsrestriktionen betrachten, welche sich auf Fredholmsche Gleichungen erste Art [49] nicht mehr zurückführen lassen. Um uns weiter abzugrenzen, ist zu vermerken, daß hier keine Probleme der Parameteridentifizierung durch Beobachtungsdaten [6], also keine inversen Aufgaben betrachtet werden. Am nächsten zu unserer Aufgabenklasse stehen die Optimalsteuerprobleme aus [61] und [62], falls sich das entsprechnde Zielfunktional exakt diskretisieren läßt. Im Gegensatz zur gewöhnlichen Ideologie der schlecht gestellten Optimalsteuerprobleme [23], in denen man meistens ein Evalutionssystem als Steuerobjekt betrachtet und die Mehrdeutigkeit der quasioptimalen Steuerung als eine günstige Tatsache empfindet, erforden unsere Aufgabenstellungen eine sorgfältige Konvergenzanalyse nicht nur für die Zielfunktionalswerte, sondern in erster Linie für die Zustands- und Kontrollfunktionen.

Gewöhnlich wird die Konvergenz numerischer Algorithmen für Kontrollprobleme unter der zusätzlichen Annahme untersucht, daß das Zielfunktional bezüglich der Steuerung stark konvex ist oder daß die Steuerfunktion ein klassisches Bang-Bang-Verhalten (ohne singuläre Bereiche) aufweist, vgl. hierzu Alt, Mackenroth [2], Glashoff, Sachs [28], Hackbusch, Will [33], Knowles [43], Lasiecka [53], [54], Mackenroth [63], Malanowski [64], Tröltzsch [96]. Als Hauptergebnis gilt dabei der Beweis der Konvergenz der approximativen Lösungen gegen die einzige Lösung des Originalproblems in vorgegebenen Banachräumen. Ein analoges Ergebnis für schlecht

gestellte Kontrollprobleme kann man nur unter der Anwendung spezieller Regularisierungstechniken erhalten.

Schwierig ist auch die Frage der Existenz einer Lösung, falls man eine nichtlineare Zustandsgleichung betrachtet und keine quadratischen Steuerkosten im Zielfunktional einführt. Das ist ein Grund, warum wir im Rahmen dieser Dissertation nur mit linearen Zustandsgleichungen arbeiten. Es sei beispielweise vermerkt, daß eine Reihe von Ergebnissen für nichtlineare Probleme in Kelley, Sachs [41], Heinkenschloss, Sachs [34], Kelley, Sachs [42] und Kupfer, Sachs [46] nur unter den Voraussetzungen hinreichender Bedingungen 2-ter Ordnung erhalten werden konnte.

Ein erstes spezielles Verfahren für nichtkorrekte Probleme (im Fall linearer Optimierungs- und Optimalsteuerungsprobleme) wurde von Tichonov [91], [92] im Jahre 1965 vorgeschlagen. Die untersuchten Probleme wurden durch eine Folge korrekter Probleme mit denselben Restriktionen regularisiert.

Weitere Fortschritte in den Verfahren für die Lösung inkorrekt gestellter Variationsungleichungen kamen mit der Arbeit von Mosco [73]. Anhand des Tichonov'schen Prinzips untersuchte er stabile numerische Schemata für die Approximation der Variationsungleichungen, indem er die Regularisierung gleichzeitig im Zielfunktional und im zulässigen Bereich einführte. Diese Technik wurde später als iterative Regularisierung bezeichnet.

Eine sich auf die Stabilitätseigenschaft der Proximal-Punkt-Abbildung stützende Methode wurde von Martinet [67] für die unrestrigierte Minimierung eines konvexen Funktionals in einem Hilbertraum vorgeschlagen. Rockafellar [84], [85] gab eine theoretische Begründung zur iterativen Prox-Regularisierung für Variationsungleichungen mit monotonen Operatoren und nichtkorrekten konvexen Optimierungsproblemen. Diese Ergebnisse haben in der numerischen Analysis großen Widerhall gefunden. Wegen der guten Stabilitätseigenschaften der Ersatzprobleme, die man auf jedem Iterationsschritt zu lösen hat, sind die Methoden der Prox-Regularisierung besonders vorteilhaft.

In der Monographie von Kaplan und Tichatschke [40] findet man sehr allgemeine Techniken zur Untersuchung der Konvergenzeigenschaften der iterativen Proximal-Punkt-Methoden bei der Behandlung (nicht streng) konvexer Variationsungleichungen. Das Prinzip der iterativen Regularisierung wurde durch die Verfahren mit Mehrschritt-Regularisierung (MSR) in den Ersatzproblemen verallgemeinert. Solche Verfahren erlauben es auf fixierten Diskretisierungsniveaus genauer zu rechnen und beschleunigen somit wesentlich den gesamten numerischen Vorgang. Die genannte Technik ergibt ebenso die Möglichkeit, Diskretisierung und Regularisierung unter Einbeziehung spezieller numerischer Basisalgorithmen der Optimierung gleichzeitig zu behandeln, d.h. das Verhalten des Lösungsprozesses läßt sich ganzheitlich untersuchen.

Straftechniken für Kontrollprobleme wurden zunächst von Lions [57] und Balakrishnan [7], [8] betrachtet. Weitere Anwendungen des Algorithmus findet man zum Beispiel in Lions [58]. Wenn die Zustandsgleichung in einen Strafterm eingebettet wird, d.h. eine Relaxierung der Zustandsgleichung erfolgt, können die Kontroll-

und die Zustandsvariablen als unabhängig betrachtet werden. Somit kann man ein Optimalsteuerproblem auf eine Variationsungleichung zurückführen. In Bergounioux [11], [12] wurde die Straftechnik genutzt, um die Existenz der Lagrange-Multiplikatoren für elliptische und parabolische Steuerprobleme mit Zustandsrestriktionen unter schwächeren Regularitätsbedingungen zu beweisen.

In Hettich, Kaplan, Tichatschke [35] und Kaplan, Tichatschke [39] werden lineare elliptische und parabolische Kontrollprobleme ohne die Voraussetzung der strengen Konvexität behandelt. Es werden keine Annahmen vorausgesetzt, welche die Korrektheit des Steuerproblems sichern würden. Somit kann das Problem eine eindeutige aber instabile Lösung besitzen; es kann sein, daß die Lösungsmenge aus mehreren Elementen besteht oder sogar eine unendlichdimensionale Mannigfaltigkeit im Steuerraum bildet. Die aufgezählten Fälle sind typisch für die von uns behandelten inkorrekt gestellten Kontrollprobleme, welche ihrerseits nur ein Sonderfall eines konvexen restringierten Variationsproblems [40] sind. Man wendet eine partielle Mehrschritt-Regularisierung (nur bezüglich der Kontrollvariablen) auf eine Folge von penalisierten Hilfsproblemen an. Das garantiert die Korrektheit der Hilfsprobleme und liefert sowohl die schwache Konvergenz der Lösungen gegen einen Optimalprozess als auch die Konvergenz der Zielfunktionalswerte gegen den Optimalwert des Originalproblems.

Die Ergebnisse in [35] und [39] kann man als Startpunkt für diese Dissertation bezeichnen. Die Fragen der Diskretisierung und numerischer Realisierung sind aber nicht der Gegenstand der in den genannten Publikationen beschriebenen Untersuchungen. Die inexakte Lösbarkeit der Ersatzprobleme wird nämlich vorausgestzt, man analysiert jedoch den Fehler nicht weiter und läßt ihn wärend der nachfolgenden Diskussion als ein abstraktes ε_i stehen. Es werden von den Autoren keine Abschätzungen des Diskretisierungs- und Berechnungsfehlers durchgeführt und in die Konvergenzanalyse mit einbezogen. Somit bleibt es immer noch unklar, warum der Regularisierungsparameter χ_i nicht beliebig schnell von Gitter zu Gitter verkleinert werden darf. Die erhaltenen Konvergenzbedingungen beruhen auf zwei abstrakten Konstanten d_1 und d_2, deren Berechnungen fehlen, was auf eine Kluft zwischen der Theorie und der numerischen Umsetzung hinausläuft: der Abbruchparameter des inneren Prox-Zyklus δ_i läßt sich aufgrund der unbekannten Konstanten nicht berechnen. Es bleibt außerdem offen, was das charakteristsche Verhalten der Lösungen der schlecht gestellten Optimalsteuerprobleme mit elliptischen oder parabolischen Zustandsgleichungen ist; das in [35] konstruierte Beispiel eines konkreten Kontrollproblems gibt keine erschöpfende Antwort dazu.

Im Kapitel 2 dieser Arbeit wird eine Verallgemeinerung der Ergebnisse aus [35] und [39] dahingehend vorgenommen, daß ein abstraktes Steuerproblem betrachtet wird. Das erlaubt es, die neuen Ergebnisse direkt für andere inkorrekte Kontrollprobleme in Hilberträumen anzuwenden. Es wird ein allgemeines regularisiertes Strafverfahren zur Lösung eines inkorrekt gestellten Optimalsteuerproblems mit einem abstrakten linearen System entwickelt und untersucht. Dabei werden zwei unterschiedliche Regularisierungstechniken — Tichonov- und Proximal-Punkt-

Regularisierung — betrachtet und verglichen. Das Zielfunktional als auch die Mengen der zulässigen Steuerungen und Zustände werden als konvex vorausgesetzt. Insbesondere kann auch die nichteindeutige Lösbarkeit der Zustandsgleichung im Rahmen unserer Untersuchungen zugelassen werden.

Eine möglichst präzise Auswertung der in den Konvergenzbedingungen auftretenden Konstanten nimmt im zweiten Kapitel einen besonderen Platz ein. Diese Abschätzungen im Rahmen des allgemeinen Schemas in Hilberträumen sind von Interesse sowohl für die klassische Tichonov Regularisierung, als auch für oben erwähnte Proximal-Punkt-Techniken.

In Kapitel 3 werden diese Ergebnisse für die Untersuchung der Konvergenz der inneren Approximationen in den Hilfsproblemen des regularisierten Strafverfahrens ausgenutzt. Es sei darauf hingewiesen, daß man in Abschnitt 3.2.2 die Berechenbarkeit der unteren Grenze des Abbruchparameters der inneren Prox-Schleife des MSR-Verfahrens erreicht. Dies gelingt eigentlich nur für das einfache Fuller-Problem [26], welches im Rahmen dieser Dissertation ein wichtiges Musterbeispiel für ein schlecht gestelltes Kontrollproblem mit elliptischer Zustandsgleichung und verteilter Steuerung darstellt. Das Fuller-Problem besitzt eine eindeutige, aber instabile Lösung und zählt nach der allgemeinen Definition in [40] zu inkorrekt gestellten Variationsaufgaben.

In Kapitel 3 befassen wir uns außerdem mit er wichtigen Frage der Numerik der schlecht gestellten Optimalsteuerprobleme: wie kann man den Grad der Inkorrektheit bestimmen? Es sind mehrere Fälle bekannt, wo die Kontrollprobleme selbst beim Fehlen der stabilisierenden quadratischen Steuerkosten entweder vollständig korrekt bleiben [57], oder bei ihrer numerischen Bechandlung auf hinreichend feinen Diskretisierungsniveaus keine Regularisierungstechniken erfordern [68]. Um die schlecht gestellten Optimalsteuerprobleme soweit zu klassifizieren, mußte man in Abschnitt 3.1 tiefer in die Theorie der elliptischen Operatoren in Sobolevräumen greifen. Insbesondere ist dabei die Diskussion der Sobolevräume mit nichtganzzahligen Indizes von Interesse. Man definiert durch sie eine Ordnung des elliptischen Operators und lernt diese Ordnung mittels eines Aufbaus zugehöriger Poincaré-Steklov Operatoren zu reduzieren. In Abschnitt 3.3 werden diese Vorbereitungen auf die Betrachtung eines schlecht gestellten Opimalsteuerproblems mit elliptischer Zustandsgleichung erster Ordnung angewandt.

Die numerischen Erfahrungen mit regularisierten Strafverfahren werden in Abschnitt 3.2.3 ausführlich beschrieben. Man kann sie nicht als besonders erfreulich bezeichnen. Das Verfahren erhält zwar eine Stabilität, konvergiert aber auf feineren Diskretisierungsniveaus unzulässig langsam. Dies führt dazu, daß man einen erheblichen Teil der Information über die gesuchte Lösung verliert. Die explizit geschriebenen Optimalitätsbedingungen zeigen, daß die Strafansätze auf Operatoren der Form A^*A führen. Es wird dadurch den Grad der Inkorrektheit praktisch verdoppelt und der Definitionsbereich eines derartigen Operators ungünstigerweise eingeengt.

In der Suche nach einem Ausweg verzichten wir in Kapitel 4 auf die unbe-

schränkten Steuerprozesse, die bei praktischen Anwendungen äußerst selten auftreten, und zugleich auf die Straftechniken für die Behandlung der inkorrekt gestellten Optimalsteuerprobleme. Somit werden alle Ersatzprobleme im MSR-Verfahren mit exakten linearen Zustandsrestriktionen betrachtet. Die neue Formuliernug des MSR-Verfahrens ohne jegliche Straftechniken ergibt auch günstigere Bedingungen für die Konvergenz der inneren Approximationen: nun darf der Regularisierungsparameter χ_i wesentlich schneller absteigen, was im Falle eines Proximal-Punkt-Verfahrens immer von Vorteil ist. Dies kann man leicht durch den Vergleich der Konvergenzbedingungen in Satz 11 und Satz 19 ersehen.

Die numerischen Tests des exakten MSR-Verfahrens, deren Ergebnisse man in Abschnitt 4.4 findet, scheinen auch erfolgreicher zu sein; es ist nicht nur gelungen, ein Optimalsteuerproblem auf dem zweidimensionalen Gebiet zu behandeln, sondern auch die feineren Diskretisierungsnivaus (bis zu $h = 2^{-8}$) mit wesentlich besseren Konvergenzqualitäten zu erreichen. Andererseits ist zu vermerken, daß die Fälle mit Zustandsrestriktionen und mit unbeschränktem Steuerbereich ohne entsprechende theoretische und numerische Behandlung blieben.

Kapitel 2

Regularisierte Strafverfahren in Hilberträumen

Dieses Kapitel enthält eine direkte Verallgemeinerung der in [35] und [39] erhaltenen Ergebnisse, welche man nun für beliebige Hilberträume erneut herleitet. **Es wird außerdem keine eindeutige Lösbarkeit der linearen Zustandsgleichung vorausgesetzt.** Wir führen ein System der *primären Konstanten* ein, welche als praktisch berechenbar angesehen werden, und bauen Zusammenhänge zwischen diesen und übrigen Konstanten auf, indem wir die konstruktiven Berechnungsformeln für die *abgeleiteten Konstanten* angeben. Die Listen der primären und abgeleiteten Konstanten sind im Anhang A zu finden. Die Gesamtheit der angeführten Berechnugsformeln ergibt ein Konstantenabbildung, die die im Konvergenzsatz auftretenden Größen d_1 und d_2 durch die primären Konstanten bestimmt.

Das Proximal-Punkt-Verfahren wird zusammen mit dem herkömmlichen Tichonov-Verfahren einer Konvergenzanalyse unterzogen. Nach der Formulierung beider Verfahren führen wir notwendige Vorbereitungen durch. Wir beweisen die Existenz einer abstrakten adjungierten Funktion p, welche für die Erhaltung der schwächeren Konvergenzbedingungen zu den regularisierten Strafverfahren eine wichtige Rolle spielt. Danach berechnen wir eine Konstante der starken Konvexität für Funktionale in partiell regularisierten Ersatzproblemen. Im nächsten Abschnitt werden die Fragen der Existenz und Eindeutigkeit auf einem sehr allgemeinen Niveau diskutiert. Die Abschnitte 2.2.4 und 2.2.5 sind dem Strafverfahren zugrunde liegenden Abschätzungen gewidmet.

In Abschnitt 2.3 erhalten wir die Konvergenzbedingungen für das Tichonov- und Prox-regularisierte Strafverfahren, die zur Lösung eines linear-quadratischen konvexen Optimalsteuerproblems verwendet werden. Diese Ergebnisse sind im Falle sowohl elliptischer, als auch parabolischer und hyperbolischer Zustandsgleichungen aktuell. Im parabolischen Fall muß man entsprechende Normen für den Zustandsraum wählen [39], [51]. Die erwähnte Normeinführeung beeinflüßt offenbar die in [39] bewiesene schwache L_2-Konvergenz im Steuerraum nicht. Somit erhalten wir hier im Vergleich mit [39] eine schwächere Konvergenzbedingung (2.131), das die

Verkleinerungsraten für die gegen Null strebenden Regularisierungs- und Strafparameter (χ_i und r_i) beschreibt:

$$\sum_{i=1}^{\infty} \sqrt{\frac{r_i}{\chi_i}} \quad < \quad \infty \, .$$

Den Tichonov-Algorithmus schreiben wir vorsätzlich in der Form, daß der Näherungsfehler ε_i zu den Ersatzproblemen nicht bezüglich der Zielfunktionalswerte, sondern bezüglich der Argumente gemessen wird. Dies vereinfacht einen späteren Vergleich der Konvergenzbedingungen von T- und MSR-Verfahren.

2.1 Regularisierte Strafverfahren für das abstrakte Kontrollproblem

Im Mittelpunkt unserer Untersuchung steht folgendes Minimierungsproblem im Hilbertraum $\mathcal{Y} \times \mathcal{U}$:

$$\mathbf{P}: \qquad \begin{aligned} & J(y,\, u) \xrightarrow{!} \min_{(y,\, u) \,\in\, \mathcal{Z}_{ad}} , \\ & A\, y = B\, u + f \quad \text{in } \mathcal{H} \, , \end{aligned}$$

wobei

$$J(y,\, u) \quad = \quad \frac{1}{2} \, \|C\, y\|_{\mathcal{K}}^2 \, . \tag{2.1}$$

Die linearen Operatoren $A \in \mathcal{L}(\mathcal{Y},\, \mathcal{H})$ und $B \in \mathcal{L}(\mathcal{U},\, \mathcal{H})$ dürfen im allgemeinen abgeschlossene Kerne besitzen und es mögen positive Konstanten $M_A,\, M_B$ berechenbar sein, so daß gilt:

$$\|A\, y\|_{\mathcal{H}} \leq M_A \, \|y\|_{\mathcal{Y}} \qquad \forall \, y \in \mathcal{Y} \, , \tag{2.2}$$

$$\|B\, u\|_{\mathcal{H}} \leq M_B \, \|u\|_{\mathcal{U}} \qquad \forall \, u \in \mathcal{U} \, . \tag{2.3}$$

Wir betrachten spezielle Hilfsräume $\mathcal{Y}_0$ und $\mathcal{U}_0$, welche einen Isomorphismus mit abgeschlossenen linearen Unterräumen in $\mathcal{Y}$ bzw. in $\mathcal{U}$ bilden. Wir setzen voraus, daß der lineare Operator G_y (bzw. G_u) konstruiert werden kann, der jedes Element aus dem Raume $\mathcal{Y}_0$ (bzw. $\mathcal{U}_0$) in das ihm entsprechende Element in $\mathcal{Y}$ (bzw. $\mathcal{U}$) überführt, so daß die Beziehungen

$$m_A \, \|G_y\, y_0\|_{\mathcal{Y}} \leq \|A G_y\, y_0\|_{\mathcal{H}} \qquad \forall \, y_0 \in \mathcal{Y}_0 \, , \tag{2.4}$$

$$m_B \, \|G_u\, u_0\|_{\mathcal{U}} \leq \|B G_u\, u_0\|_{\mathcal{H}} \qquad \forall \, u_0 \in \mathcal{U}_0 \tag{2.5}$$

mit positiven abschätzbaren Konstanten m_A, m_B gelten. Außerdem gelte für die Wertebereiche

$$\mathcal{R}(AG_y) = \mathcal{R}(A) = \mathcal{H}\,, \qquad \mathcal{R}(BG_u) = \mathcal{R}(B) \subseteq \mathcal{H}\,.$$

Diese Annahmen lassen sich in praktischen Fällen (z. B. nichteindeutig lösbare elliptische Zustandsgleichungen) recht einfach erfüllen und sichern die Existenz von inversen Operatoren $(AG_y)^{-1} \in \mathcal{L}(\mathcal{H},\, \mathcal{Y}_0)$ und $(BG_u)^{-1} \in \mathcal{L}(\mathcal{H},\, \mathcal{U}_0)$. Für die Definitionsbereiche gelte dabei

$$\mathcal{D}((AG_y)^{-1}) = \mathcal{H}\,, \qquad \mathcal{D}((BG_u)^{-1}) = \mathcal{R}(B) \subseteq \mathcal{H}\,.$$

Es wird angenommen, daß der Operator $C \in \mathcal{L}(\mathcal{Y},\, \mathcal{K})$ kompakt und im allgemeinen nicht injektiv ist. Das Problem **P** wird deshalb inkorrekt. Für den Operator C mögen folgende Ungleichungen

$$m_C \;\|y\|_{\mathcal{Y}} \le \sqrt{\|C\,y\|^2_{\mathcal{K}} + \|A\,y\|^2_{\mathcal{H}}} \qquad \forall\, y \in \mathcal{Y}\,, \tag{2.6}$$

$$\|C\,y\|_{\mathcal{K}} \le M_C \;\|y\|_{\mathcal{Y}} \qquad \forall\, y \in \mathcal{Y}\,, \tag{2.7}$$

mit positiven Konstanten m_C und M_C gelten. Die Bedingung (2.6) ergibt für die Nullmengen: $\mathcal{N}(A) \cap \mathcal{N}(C) = \{0\}$.

Die Menge $\mathcal{Z}_{ad} = \mathcal{Y}_{ad} \times \mathcal{U}_{ad}$, welche durch Zustands- und Steuerrestriktionen im Prozessraum $\mathcal{Z}$ gegeben ist, der sich als direktes Produkt zwischen Zustandsraum $\mathcal{Y}$ und Steuerraum $\mathcal{U}$ definieren läßt, sei konvex und abgeschlossen. Für $\mathcal{Y}_{ad}$ und $\mathcal{U}_{ad}$ gelten außerdem die Voraussetzungen:

$$\exists\, M_{ad} > 0 \;:\quad \|u\|_{\mathcal{U}} \le M_{ad} \quad \forall\, u \in \mathcal{U}_{ad}\,, \tag{2.8}$$

$$\exists\, m_{ad} > 0\,, \quad \exists\, \tilde{u} \in \mathcal{U}_{ad} \;:\quad m_{ad} \le \|\tilde{y}(\tilde{u}) - y\|_{\mathcal{Y}} \quad \forall\, y \in \partial\mathcal{Y}_{ad}\,, \tag{2.9}$$

wobei $\tilde{y} = G_y(AG_y)^{-1}(B\,\tilde{u} + f) \in \mathcal{Y}_{ad}$ die normale Lösung der Zustandsgleichung des Problems **P** darstellt. Die Bedingung (2.9) sichert somit die Erreichbarkeit eines inneren Punkts in den Zustandsrestriktionen. Es sei vorausgesetzt, daß **P** lösbar ist und mit $z^* = (y^*,\, u^*) \in \mathcal{Z}_{ad}$ bezeichnen wir eine der Lösungen von **P** im Prozessraum $\mathcal{Z}$.

Das Problem **P** ist i.a. nicht eindeutig lösbar. Wir untersuchen zu seiner Lösung folgende regularisierte Strafverfahren. Wir betrachten drei Folgen von positiven gegen Null strebenden Parametern: $\{\chi_i\}_{i=1}^{\infty}$, $\{r_i\}_{i=1}^{\infty}$ und $\{\varepsilon_i\}_{i=1}^{\infty}$. Sie beschreiben die Regularisierungsparameter, die Strafparameter und die Genauigkeiten, mit denen die Hilfsprobleme gelöst werden. Darüber hinaus betrachten wir zwei Folgen von Prozessen $\{(y^i,\, u^i)\}_{i=1}^{\infty}$ und $\{(\bar{y}^i,\, \bar{u}^i)\}_{i=1}^{\infty}$. Das Strafverfahren mit der Tichonov (T) Regularisierung läßt sich dann auf folgende Weise formulieren:

T-Verfahren:

Gegeben $t_d \in \mathcal{T}$; $i := 1$;

(a) Berechne eine Näherung $z^i = (y^i, u^i) \in \mathcal{Y} \times \mathcal{U}$

der exakten Lösung $\bar{z}^i = (\bar{y}^i, \bar{u}^i)$ des Hilfsproblems $\bar{\mathbf{P}}_i^{(T)}$,

so daß $\quad \| z^i - \bar{z}^i \|_{\mathcal{Z}} \leq \varepsilon_i$;

$i := i + 1$; gehe zu (a) .

Die Folge der Hilfsprobleme im T-Verfaren ist

$$\bar{\mathbf{P}}_i^{(T)} : \qquad \bar{\Psi}_i(y, u) \xrightarrow{!} \min_{(y,u) \in \mathcal{Z}_{ad}} ,$$

wobei

$$\begin{aligned} \bar{\Psi}_i(y, u) &= J(y, u) + \frac{1}{2\, r_i} \, \| W(A\, y - B\, u - f) \|_{\mathcal{S}}^2 + \\ &\quad + \frac{\chi_i}{2} \, \| N u - t_d \|_{\mathcal{T}}^2 \ . \end{aligned} \tag{2.10}$$

Um die zugehörigen Hilfsfunktionale (2.10) exakt zu definieren, braucht man im allgemeinen neue Hilberträume $\mathcal{S}$, $\mathcal{T}$ und neue Operatoren W , N . Es wird dabei die Existenz der entsprechenden Lipschitz-Konstanten

$$m_W \, \| h \|_{\mathcal{H}} \ \leq \| W\, h \|_{\mathcal{S}} \ \leq M_W \, \| h \|_{\mathcal{H}} \ , \tag{2.11}$$

$$m_N \, \| u \|_{\mathcal{U}} \ \leq \| N\, u \|_{\mathcal{T}} \ \leq M_N \, \| u \|_{\mathcal{U}} \tag{2.12}$$

gefordert. Für den Wertebereich des Operators N gelte auch $\mathcal{R}(N) = \mathcal{T}$. Die Hilberträume $\mathcal{S}$, $\mathcal{T}$ und die linearen Operatoren W , N werden auch bei der Zusammensetzung von Straf- und Regularisierungsgliedern im Proximal-Punkt-Verfahren verwendet. Dazu betrachten wir wieder die Folgen von positiven Parametern $\{\chi_i\}_{i=1}^{\infty}$, $\{r_i\}_{i=1}^{\infty}$, und $\{\varepsilon_i\}_{i=1}^{\infty}$. Um den Abbruch der inneren Schleife des neuen Verfahrens zu kontrollieren, benötigen wir außerdem die positive Folge $\{\delta_i\}_{i=1}^{\infty}$. Das Strafverfahren mit einer Mehrschritt-Regularisierung (MSR) läßt sich dann auf folgende Weise formulieren:

MSR-Verfahren:

Gegeben $u^{0,0} \in \mathcal{U}$; $s(0) := 0$; $i := 1$;

(a) $u^{i,0} := u^{i-1,s(i-1)}$; $s := 0$;

(b) Gegeben $u^{i,s} \in \mathcal{U}$; $s := s+1$;

Berechne eine Näherung $z^{i,s} = (y^{i,s},\, u^{i,s}) \in \mathcal{Y} \times \mathcal{U}$

der exakten Lösung $\bar{z}^{i,s} = (\bar{y}^{i,s},\, \bar{u}^{i,s})$ des Hilfsproblems $\bar{\mathbf{P}}_{i,s}^{(MSR)}$,

so daß $\| z^{i,s} - \bar{z}^{i,s} \|_{\mathcal{Z}} \leq \varepsilon_i$;

(c) Wenn $\| u^{i,s} - u^{i,s-1} \|_{\mathcal{U}} \geq \delta_i$,

dann gehe zu (b)

sonst $s(i) := s$; $i := i+1$; gehe zu (a)

Die Folge der Hilfsprobleme im MSR-Verfahren hat die Struktur

$$\bar{\mathbf{P}}_{i,s}^{(MSR)}: \qquad \bar{\Psi}_{i,s}(y,\, u) \xrightarrow{!} \min_{(y,\, u) \in \mathcal{Z}_{ad}},$$

wobei

$$\begin{aligned} \bar{\Psi}_{i,s}(y,\, u) &= J(y,\, u) + \frac{1}{2\, r_i} \, \|W(A\, y - B\, u - f)\|_{\mathcal{S}}^2 + \\ &\quad + \frac{\chi_i}{2} \, \left\|N(u - u^{i,s-1})\right\|_{\mathcal{T}}^2 . \end{aligned} \tag{2.13}$$

Wenn der Abbruchparameter hinreichend groß (z. B. $\delta_i > 2M_{ad}$) gewählt wird, geht das MSR-Verfahren in das OSR-Verfahren[1] über.

Die weiteren Darlegungen erfordern die Einführung von zwei neuen Folgen von Hilfsproblemen:

$$\begin{aligned} \bar{\bar{\mathbf{P}}}_i^{(T)}: \qquad & \bar{\bar{\Psi}}_i(y,\, u) \xrightarrow{!} \min_{(y,\, u) \in \mathcal{Z}_{ad}}, \\ & A\, y = B\, u + f \quad \text{in } \mathcal{H}, \end{aligned}$$

wobei

$$\bar{\bar{\Psi}}_i(y,\, u) = J(y,\, u) + \frac{\chi_i}{2} \, \|Nu - t_d\|_{\mathcal{T}}^2 \tag{2.14}$$

[1] One step regularization (vgl. [40])

und

$$\mathbf{P}_i : \qquad J_i(y,\, u) \xrightarrow{!} \min_{(y,\, u) \in \mathcal{Z}_{ad}} ,$$

wobei

$$J_i(y,\, u) \;=\; J(y,\, u) + \frac{1}{2\, r_i}\, \|W(A\, y - B\, u - f)\|_{\mathcal{S}}^2 \ . \tag{2.15}$$

2.2 Einige vorbereitende Resultate

2.2.1 Abschätzung der adjungierten Funktion

Jedem Element $y \in \mathcal{Y}$ kann ein Element

$$p \;=\; (G_y^* A^*)^{-1} G_y^* C^* \Lambda_{\mathcal{K}} C\, y \qquad \text{in } \ \mathcal{H}' \tag{2.16}$$

zugeordnet werden, das die normale Lösung der adjungierten Zustandsgleichung darstellt. Dabei werden adjungierte Operatoren mit Sternen und duale Räume mit Strichen versehen. Wir bezeichnen p als *adjungierte Funktion* und weisen nach, daß die oben stehende Definition korrekt erfolgt und daß die adjungierte Funktion beschränkt ist, wenn eine obere Schranke für die Norm des Elements y vorliegt. Dazu benötigen wir folgendes Lemma [102] .

Lemma 1 *Seien $\mathcal{X}_1$ und $\mathcal{X}_2$ Hilberträume mit $\|\cdot\|_{\mathcal{X}_1}$ und $\|\cdot\|_{\mathcal{X}_2}$. Für einen Operator $L \in \mathcal{L}(\mathcal{X}_1,\, \mathcal{X}_2)$ gilt*

(i) *$\mathcal{R}(L) = \mathcal{X}_2$ dann und nur dann, wenn L^* stetig invertierbar ist ;*

(ii) *$\mathcal{R}(L^*) = \mathcal{X}_1'$ dann und nur dann, wenn L stetig invertierbar ist .*

Den Beweis für Banachräume im Falle, daß L ein unbeschränkter abgeschlossener Operator mit in $\mathcal{X}_1$ dichtem Definitionsbereich ist, findet man in Yosida [102, Kap. VII, §5] . Für eine genaue Konstantenabschätzung brauchen wir später eine Erweiterung dieses Lemmas.

Lemma 2 *Seien $\mathcal{X}_1$ und $\mathcal{X}_2$ Hilberträume mit $\|\cdot\|_{\mathcal{X}_1}$ und $\|\cdot\|_{\mathcal{X}_1}$. Für einen Operator $L \in \mathcal{L}(\mathcal{X}_1,\, \mathcal{X}_2)$ gelte:*

(i) *$\mathcal{R}(L) = \mathcal{X}_2$;*

(ii) *$\exists\, m_L,\, M_L$ mit $\quad 0 < m_L \le M_L < \infty$, so daß*

$$m_L\, \|x_1\|_{\mathcal{X}_1} \le \|L\, x_1\|_{\mathcal{X}_2} \le M_L\, \|x_1\|_{\mathcal{X}_1} \qquad \forall\, x_1 \in \mathcal{X}_1 \ . \tag{2.17}$$

Dann existiert ein adjungierter Operator $L^* \in \mathcal{L}(\mathcal{X}_2',\, \mathcal{X}_1')$ und es gilt :

(i) $\mathcal{R}(L^*) = \mathcal{X}_1'$;

(ii) *mit denselben Konstanten wie für den Operator L ist*

$$m_L \, \|\tilde{x}_2\|_{\mathcal{X}_2'} \leq \|L^* \, \tilde{x}_2\|_{\mathcal{X}_1'} \leq M_L \, \|\tilde{x}_2\|_{\mathcal{X}_2'} \qquad \forall \, \tilde{x}_2 \in \mathcal{X}_2' \, . \tag{2.18}$$

Beweis: Aus den Voraussetzungen (i) und (ii) folgt die stetige Invertierbarkeit von L (s. z.B. [60, Kap. III, §5, Satz 1]). Man findet auch eine obere Schranke für die Norm des inversen Operators:

$$\left\| L^{-1} x_2 \right\|_{\mathcal{X}_1} \leq \frac{1}{m_L} \, \|x_2\|_{\mathcal{X}_2} \qquad \forall \, x_2 \in \mathcal{X}_2 \, . \tag{2.19}$$

Nach Lemma 1 gilt, daß $\mathcal{R}(L^*) = \mathcal{X}_1'$ und daß $(L^*)^{-1} \in \mathcal{L}(\mathcal{X}_1', \, \mathcal{X}_2')$ existiert. Somit ist die Aussage (i) bewiesen und wir haben insgesamt $\mathcal{D}(L^{-1}) = \mathcal{X}_2$ und $\mathcal{D}((L^*)^{-1}) = \mathcal{X}_1'$. Die Definition des adjungierten Operators ist:

$$\left\langle \, L^* \, \tilde{x}_2, \, x_1 \, \right\rangle_{\mathcal{X}_1', \, \mathcal{X}_1} \;\; = \;\; \langle \, \tilde{x}_2, \, L \, x_1 \, \rangle_{\mathcal{X}_2', \, \mathcal{X}_2} \qquad \forall \, \tilde{x}_2 \in \mathcal{X}_2' \, , \;\; \forall \, x_1 \in \mathcal{X}_1 \, . \tag{2.20}$$

Für beliebige $\tilde{x}_1 \in \mathcal{X}_1'$ und $x_2 \in \mathcal{X}_2$ wählen wir $\tilde{x}_2 = (L^*)^{-1} \tilde{x}_1$ und $x_1 = L^{-1} \, x_2$ und setzen diese in (2.20) ein. Nach Elimination bekommt man:

$$\langle \, (L^*)^{-1} \tilde{x}_1, \, x_2 \, \rangle_{\mathcal{X}_2', \, \mathcal{X}_2} \;\; = \;\; \left\langle \, \tilde{x}_1, \, L^{-1} x_2 \, \right\rangle_{\mathcal{X}_1', \, \mathcal{X}_1} \qquad \forall \, \tilde{x}_1 \in \mathcal{X}_1' \, , \;\; \forall \, x_2 \in \mathcal{X}_2 \, . \tag{2.21}$$

Aus der Gleichung (2.21) erhält man

$$(L^{-1})^* = (L^*)^{-1} \qquad \text{in} \;\; \mathcal{L}(\mathcal{X}_1', \, \mathcal{X}_2') \, . \tag{2.22}$$

Mit Hilfe von (2.19) und (2.22) gilt folgende Abschätzung:

$$\left\| (L^*)^{-1} \tilde{x}_1 \right\|_{\mathcal{X}_2'} = \left\| (L^{-1})^* \tilde{x}_1 \right\|_{\mathcal{X}_2'} = \sup_{x_2 \in \mathcal{X}_2 \setminus \{0\}} \frac{\langle \, (L^{-1})^* \tilde{x}_1, \, x_2 \, \rangle_{\mathcal{X}_2', \, \mathcal{X}_2}}{\|x_2\|_{\mathcal{X}_2}} \leq$$

$$\|\tilde{x}_1\|_{\mathcal{X}_1'} \sup_{x_2 \in \mathcal{X}_2 \setminus \{0\}} \frac{\|L^{-1} x_2\|_{\mathcal{X}_1}}{\|x_2\|_{\mathcal{X}_2}} \leq \frac{1}{m_L} \, \|\tilde{x}_1\|_{\mathcal{X}_1'} \qquad \forall \, \tilde{x}_1 \in \mathcal{X}_1' \, . \tag{2.23}$$

Analog beweist man

$$\|L^* \tilde{x}_2\|_{\mathcal{X}_1'} \leq M_L \, \|\tilde{x}_2\|_{\mathcal{X}_2'} \qquad \forall \, \tilde{x}_2 \in \mathcal{X}_2' \, . \tag{2.24}$$

Da $\mathcal{R}((L^*)^{-1}) = \mathcal{D}(L^*) = \mathcal{X}_2'$, folgt aus (2.23)

$$m_L \, \|\tilde{x}_2\|_{\mathcal{X}_2'} \leq \|L^* \tilde{x}_2\|_{\mathcal{X}_1'} \qquad \forall \, \tilde{x}_2 \in \mathcal{X}_2' \, . \tag{2.25}$$

Mit den Ungleichungen (2.24) und (2.25) folgt die Aussage (ii) des Lemmas.

□

Nun wenden wir das Lemma 2 auf den Systemoperator A des Problems **P** an. Das Ergebnis wird im folgenden Satz formuliert:

Satz 1 *Sei $G_y\,\mathcal{Y}_0 \subseteq \mathcal{Y}$ und $G_y \in \mathcal{L}(\mathcal{Y}_0,\,\mathcal{Y})$ eine entsprechende Einbettung. Für den Operator $A \in \mathcal{L}(\mathcal{Y},\,\mathcal{H})$ gelte:*

(i) $\mathcal{R}(AG_y) = \mathcal{H}$;

(ii) $\exists\, m_A,\, M_A$ *mit* $\quad 0 < m_A \leq M_A < \infty$, *so daß*

$$m_A\,\|G_y\,y_0\|_{\mathcal{Y}} \leq \|AG_y\,y_0\|_{\mathcal{H}} \leq M_A\,\|G_y\,y_0\|_{\mathcal{Y}} \qquad \forall\, y_0 \in \mathcal{Y}_0\,.$$

Dann existieren adjungierte Operatoren $A^* \in \mathcal{L}(\mathcal{H}',\,\mathcal{Y}')$, $G_y^* \in \mathcal{L}(\mathcal{Y}',\,\mathcal{Y}_0')$ und es gilt :

(i) $\mathcal{R}(G_y^*A^*) = \mathcal{Y}_0'$;

(ii) $\exists\,(G_y^*A^*)^{-1} \in \mathcal{L}(\mathcal{Y}_0',\,\mathcal{H}')$, *so daß*

$$\left\|(G_y^*A^*)^{-1}G_y^*\tilde{y}\right\|_{\mathcal{H}'} \leq m_A^{-1}\,\|\tilde{y}\|_{\mathcal{Y}'} \qquad \forall\,\tilde{y} \in \mathcal{Y}'\,. \tag{2.26}$$

Beweis: Der Beweis folgt direkt aus Lemma 2 mit den Substitutionen: $L := AG_y$, $L^* := G_y^*A^*$, $\mathcal{X}_1 := \mathcal{Y}_0$, $\mathcal{X}_2 := \mathcal{H}$, $x_1 := y_0 \in \mathcal{Y}_0$, $\tilde{x}_2 := (G_y^*A^*)^{-1}G_y^*\,\tilde{y} \in \mathcal{H}'$, $m_L := m_A$, $M_L := M_A$. Die Existenz von $(G_y^*A^*)^{-1}$ läßt sich analog zu dem vorherigen Beweis zeigen unter Anwendung von (2.19) und (2.22). Mit (2.18) bekommen wir

$$m_A\,\left\|(G_y^*A^*)^{-1}G_y^*\tilde{y}\right\|_{\mathcal{H}'} \leq \left\|G_y^*A^*(G_y^*A^*)^{-1}G_y^*\tilde{y}\right\|_{\mathcal{Y}_0'} = \left\|G_y^*\tilde{y}\right\|_{\mathcal{Y}_0'} =$$

$$\sup_{y_0\in\mathcal{Y}_0\setminus\{0\}} \frac{\langle\, G_y^*\,\tilde{y},\, y_0\,\rangle_{\mathcal{Y}_0',\mathcal{Y}_0}}{\|G_y\,y_0\|_{\mathcal{Y}}} \leq \sup_{y\in\mathcal{Y}\setminus\{0\}} \frac{\langle\,\tilde{y},\, y\,\rangle_{\mathcal{Y}',\mathcal{Y}}}{\|y\|_{\mathcal{Y}}} = \|\tilde{y}\|_{\mathcal{Y}'}\ .$$

□

Mit (2.26) und (2.18) (mit $L^* := C^*$) erhalten wir eine Abschätzung für die adjungierte Funktion:

$$\|p\|_{\mathcal{H}'} \;\leq\; \frac{M_C^2}{m_A}\,\|y\|_{\mathcal{Y}}\ . \tag{2.27}$$

2.2.2 Starke Konvexität von $\bar{\Psi}_i(z)$ und $\bar{\Psi}_{i,s}(z)$

Wir betrachten zunächst einen abstrakten Operator $F \in \mathcal{L}(\mathcal{T}, \mathcal{H})$ und setzen voraus, daß es einen Isomorphismus zwischen dem orthogonalen Komplement des Kerns von F , der eine abgeschlossene Menge in $\mathcal{T}$ darstellt, und einem Hilbertraum $\mathcal{T}_0$ gebe. Sei $G \in \mathcal{L}(\mathcal{T}_0, \mathcal{T})$ der entsprechende Operator des kanonischen Isomorphismus:

$$\|t_0\|_{\mathcal{T}_0} \;=\; \|G\,t_0\|_{\mathcal{T}} \qquad \forall\, t_0 \in \mathcal{T}_0$$

Voraussetzung 1 *Für den Operator F gelte:*

$$m_0\ \|G\,t_0\|_{\mathcal{T}} \le \|FG\,t_0\|_{\mathcal{H}} \qquad \forall\, t_0 \in \mathcal{T}_0\ , \tag{2.28}$$

$$\|F\,t\|_{\mathcal{H}} \le M_0\ \|t\|_{\mathcal{T}} \qquad \forall\, t \in \mathcal{T}\ . \tag{2.29}$$

Aus dieser Voraussetzung folgt die Existenz des inversen Operators $(FG)^{-1} \in \mathcal{L}(\mathcal{H}, \mathcal{T}_0)$ mit $\mathcal{D}((FG)^{-1}) = \mathcal{R}(F)$.

Lemma 3 *Die Operatoren $(FG)^{-1}F : \mathcal{T} \to \mathcal{T}_0$ und $((FG)^{-1}F)^* : \mathcal{T}_0' \to \mathcal{T}'$ sind beschränkt.*

Beweis: Für beliebiges $t \in \mathcal{T}$ folgt aus (2.28) und (2.29)

$$m_0\ \left\|(FG)^{-1}F\,t\right\|_{\mathcal{T}_0} \le \;\; \|FG(FG)^{-1}F\,t\|_{\mathcal{H}} \;\; \le M_0\ \|t\|_{\mathcal{T}}\ .$$

Somit haben wir

$$\left\|(FG)^{-1}F\,t\right\|_{\mathcal{T}_0} \le \frac{M_0}{m_0}\ \|t\|_{\mathcal{T}} \qquad \forall\, t \in \mathcal{T}\ . \tag{2.30}$$

Unter Anwendung von Lemma 2 mit $M_L = M_0/m_0$ schätzen wir mit (2.18) ab:

$$\left\|\left((FG)^{-1}F\right)^*\,\tilde{t}_0\right\|_{\mathcal{T}'} \le \frac{M_0}{m_0}\ \left\|\tilde{t}_0\right\|_{\mathcal{T}_0'} \qquad \forall\, \tilde{t}_0 \in \mathcal{T}_0'\ . \tag{2.31}$$

□

Lemma 4 *Sei $\Lambda_{\mathcal{T}}$ ein Operator des kanonischen Isomorphismus zwischen $\mathcal{T}$ und $\mathcal{T}'$. Der Operator $(FG)^{-1}F\Lambda_{\mathcal{T}}^{-1}\,((FG)^{-1}F)^* : \mathcal{T}_0' \to \mathcal{T}_0 = \mathcal{T}_0''$ ist beschränkt und positiv definit.*

Beweis: Die Beschränktheit des Operators folgt aus der Beschränktheit der einzelnen Komponenten. Für beliebiges $\tilde{t}_0 \in \mathcal{T}_0'$ gelten die Abschätzungen:

$$\left\|\tilde{t}_0\right\|_{\mathcal{T}_0'}^2 = \left\|\left((FG)^{-1}FG\right)^*\,\tilde{t}_0\right\|_{\mathcal{T}_0'}^2 = \left\|G^*\left((FG)^{-1}F\right)^*\,\tilde{t}_0\right\|_{\mathcal{T}_0'}^2 \le$$

$$\left\|\left((FG)^{-1}F\right)^*\,\tilde{t}_0\right\|_{\mathcal{T}'}^2 = \left\langle\, \left((FG)^{-1}F\right)^*\,\tilde{t}_0,\, \Lambda_{\mathcal{T}}^{-1}\left((FG)^{-1}F\right)^*\,\tilde{t}_0\,\right\rangle_{\mathcal{T}',\mathcal{T}} =$$

$$\left\langle\, \tilde{t}_0,\, (FG)^{-1}F\Lambda_{\mathcal{T}}^{-1}\left((FG)^{-1}F\right)^*\,\tilde{t}_0\,\right\rangle_{\mathcal{T}_0',\mathcal{T}_0}\ .$$

□

Aus dem Lax-Milgram-Lemma (s. z.B. [20, Satz 1.1.3] oder [30, Lemma 3.6]) folgt die Existenz des inversen Operators und die zugehörige obere Schranke:

$$\left\| \left\{ (FG)^{-1} F \Lambda_{\mathcal{T}}^{-1} \left((FG)^{-1} F \right)^* \right\}^{-1} t_0 \right\|_{\mathcal{T}_0'} \leq \|t_0\|_{\mathcal{T}_0} \qquad \forall\, t_0 \in \mathcal{T}_0 \,. \tag{2.32}$$

Wir ordnen nun jedem Element $t \in \mathcal{T}$ drei andere Elemente $t_{\mathrm{Ker}}^{\perp} \in \mathcal{T}$, $t_{\mathrm{Ker}} \in \mathcal{T}$ und $t_0 \in \mathcal{T}_0$ zu:

$$t_{\mathrm{Ker}}^{\perp} = L\, t\,; \qquad t_{\mathrm{Ker}} = (I - L)\, t\,; \tag{2.33}$$

$$t_0 = (FG)^{-1} F\, t = (FG)^{-1} F\, t_{\mathrm{Ker}}^{\perp}\,, \tag{2.34}$$

wobei

$$L = \Lambda_{\mathcal{T}}^{-1} \left((FG)^{-1} F \right)^* \left\{ (FG)^{-1} F \Lambda_{\mathcal{T}}^{-1} \left((FG)^{-1} F \right)^* \right\}^{-1} (FG)^{-1} F \,. \tag{2.35}$$

Die in (2.33) definierten Elemente besitzen die folgenden Eigenschaften:

1. $F\, t_{\mathrm{Ker}} = 0 \;\text{ in } \mathcal{H} \qquad \forall\, t \in \mathcal{T}$;
2. $(t_{\mathrm{Ker}},\, t_{\mathrm{Ker}}^{\perp})_{\mathcal{T}} = 0 \qquad \forall\, t \in \mathcal{T}$;
3. $t = t_{\mathrm{Ker}} + t_{\mathrm{Ker}}^{\perp}$.

Folgende Beziehungen sind auch leicht zu prüfen:

$$\left\| t_{\mathrm{Ker}}^{\perp} \right\|_{\mathcal{T}} \leq \frac{M_0}{m_0}\, \|t_0\|_{\mathcal{T}_0} \qquad \forall\, t \in \mathcal{T}\,, \tag{2.36}$$

$$\left\| F\, t_{\mathrm{Ker}}^{\perp} \right\|_{\mathcal{H}} \geq \frac{m_0^2}{M_0}\, \left\| t_{\mathrm{Ker}}^{\perp} \right\|_{\mathcal{T}} \qquad \forall\, t \in \mathcal{T}\,. \tag{2.37}$$

Nun erläutern wir den Zusammenhang von F, G und $\mathcal{T}_0$ mit den im Abschnitt 2.1 eingeführten Räumen und Operatoren. Zuerst betrachten wir die zweite Gâteaux-Ableitung des Hilfsfunktionals $\bar{\Psi}_i(z)$ im Punkte z_0 . Durch Differentiation erhält man:

$$\nabla_{zz}^2 \bar{\Psi}_i(y_0,\, u_0)(y,\, u)(y,\, u) = \|C\, y\|_{\mathcal{K}}^2 + \frac{1}{r_i}\, \|W(A\, y - B\, u)\|_{\mathcal{S}}^2 + \chi_i\, \|N\, u\|_{\mathcal{T}}^2 \,. \tag{2.38}$$

Um die Konstante der starken Konvexität von $\bar{\Psi}_i(z)$ zu gewinnen, müssen wir (2.38) nach unten abschätzen. Dabei benutzen wir die obigen Lemmata. Zunächst erklären wir eine Zerlegung des Elements $u \in \mathcal{U}$ mit folgenden Eigenschaften:

1. $B\,u_1 = 0 \;\text{ in } \mathcal{H}$;

2. $(N\,u_1,\, N\,u_2)_{\mathcal{T}} = 0$;

3. $u = u_1 + u_2$.

Wir wenden nun die erhaltenen Lemmata 3 und 4 auf die folgenden Substitutionen an: $F := BN^{-1}$, $G := NG_u$, $\mathcal{T}_0 := \mathcal{U}_0$ und somit $t_0 := u_0$, $t_{\text{Ker}} := N\,u_1$, $t^{\perp}_{\text{Ker}} := N\,u_2$, $m_0 := m_B/m_N$, $M_0 := M_B/M_N$. Die Existenz eines inversen Operators $N^{-1} \in \mathcal{L}(\mathcal{T}, \mathcal{U})$ ist mit den Annahmen in Abschnitt 2.1 gesichert. Die folgenden Relationen bezüglich der Konstanten gelten entsprechend Voraussetzung 1 und (2.12), (2.5), (2.3):

$$\begin{aligned} \forall\, u_0 \in \mathcal{U}_0 \;: \qquad & \|BG_u\,u_0\|_{\mathcal{H}} = \left\|BN^{-1}NG_u\,u_0\right\|_{\mathcal{H}} \geq \\ & m_0\;\|NG_u\,u_0\|_{\mathcal{T}} \geq m_0 m_N\;\|G_u\,u_0\|_{\mathcal{T}}\;; \end{aligned} \tag{2.39}$$

$$\forall\, u \in \mathcal{U} \;: \qquad \|B\,u\|_{\mathcal{H}} = \left\|BN^{-1}N\,u\right\|_{\mathcal{H}} \leq M_0\;\|N\,u\|_{\mathcal{T}} \leq M_0 M_N\;\|u\|_{\mathcal{U}}\;. \tag{2.40}$$

Für (2.38) kann man jetzt schreiben:

$$\nabla^2_{zz}\bar{\Psi}_i = \|C\,y\|^2_{\mathcal{K}} + \frac{1}{r_i}\;\|W(A\,y - B\,u_2)\|^2_{\mathcal{S}} + \chi_i\,\|N\,u_1\|^2_{\mathcal{T}} + \chi_i\,\|N\,u_2\|^2_{\mathcal{T}}\;. \tag{2.41}$$

Nach (2.11), (2.12) und (2.3) gilt

$$\|Nu_2\|^2_{\mathcal{T}} \geq m_N^2\;\|u_2\|^2_{\mathcal{U}} \geq \frac{m_N^2}{M_B^2}\;\|B\,u_2\|^2_{\mathcal{H}} \geq \frac{m_N^2}{M_B^2 M_W^2}\;\|WB\,u_2\|^2_{\mathcal{S}}\;.$$

Mit Hilfe der letzten Ungleichung kann man nun (2.41) analysieren, indem man für jedes i eine gewöhnliche quadratische Form von drei Variablen nach unten abschätzt:

$$\begin{aligned} k(y,\,u_1,\,u_2) \;&=\; \frac{1}{r_i}\,(y-u_2)^2 + \chi_i\,u_1^2 + \frac{m_N^2}{M_B^2 M_W^2}\,\chi_i\,u_2^2 \geq \\ &\geq \lambda_i\,(y^2 + u_1^2 + u_2^2)\;. \end{aligned} \tag{2.42}$$

Somit bekommen wir folgende Abschätzung:

$$\nabla^2_{zz}\bar{\Psi}_i \;\geq\; \|C\,y\|^2_{\mathcal{K}} + \lambda_i\,\left(\|WA\,y\|^2_{\mathcal{S}} + \|WB\,u_2\|^2_{\mathcal{S}} + \|N\,u_1\|^2_{\mathcal{T}}\right)\;, \tag{2.43}$$

wobei

$$\lambda_i = \min\left\{\chi_i\,,\ \frac{\chi}{r_i\chi + 2}\right\}\ , \qquad \chi = \frac{m_N^2}{M_B^2 M_W^2}\,\chi_i\ . \tag{2.44}$$

Um (2.43) weiter nach unten abzuschätzen, benutzen wir die Eigenschaft (2.37) mit den oben erklärten Substitutionen. Es gilt

$$\|WB\,u_2\|_{\mathcal{S}}^2 \geq m_W^2\ \left\|BN^{-1}N\,u_2\right\|_{\mathcal{H}}^2 \geq m_W^2\ \left[\left(\frac{m_B}{m_N}\right)^2 \Big/ \frac{M_B}{M_N}\right]^2\ \|N\,u_2\|_{\mathcal{T}}^2\ . \tag{2.45}$$

Mit

$$c_u = \min\left\{\,1\,,\ m_W^2\ \left[\left(\frac{m_B}{m_N}\right)^2 \Big/ \frac{M_B}{M_N}\right]^2\right\} \tag{2.46}$$

setzen wir (2.43) fort:

$$\begin{aligned}\nabla_{zz}^2\bar{\Psi}_i \ \ &\geq\ \ \|C\,y\|_{\mathcal{K}}^2 + \lambda_i\ \left(\|WA\,y\|_{\mathcal{S}}^2 + c_u\ \left(\|N\,u_2\|_{\mathcal{T}}^2 + \|N\,u_1\|_{\mathcal{T}}^2\right)\right) = \\ &\|C\,y\|_{\mathcal{K}}^2 + \lambda_i\ \left(\|WA\,y\|_{\mathcal{S}}^2 + c_u\ \|N\,u\|_{\mathcal{T}}^2\right)\ .\end{aligned} \tag{2.47}$$

Bezeichne

$$c_\lambda = \min\left\{\,1\,,\ m_W^2\,\lambda_i\,\right\}\ . \tag{2.48}$$

Dann erhalten wir mit (2.6) und (2.12)

$$\begin{aligned}\nabla_{zz}^2\bar{\Psi}_i \ \ &\geq\ \ c_\lambda\ \left(\|C\,y\|_{\mathcal{K}}^2 + \frac{1}{m_W^2}\ \|WA\,y\|_{\mathcal{S}}^2\right) + \lambda_i\,c_u\ \|N\,u\|_{\mathcal{T}}^2 \geq \\ &m_C^2\,c_\lambda\ \|y\|_{\mathcal{Y}}^2 + \lambda\,m_N^2\,c_u\ \|u\|_{\mathcal{U}}^2\ ,\end{aligned} \tag{2.49}$$

d.h. die Hessematrix des Funktionals $\bar{\Psi}_i$ ist positiv definit. Das Endergebnis läßt sich auch wie folgt formulieren:

$$\begin{aligned}&\forall\, z_1 \in \mathcal{Z},\ \forall\, z_2 \in \mathcal{Z}\ : \\ &\bar{\Psi}_i(z_2) - \bar{\Psi}_i(z_1) \geq \nabla_z\bar{\Psi}_i(z_1)(z_2 - z_1) + c_\kappa\ \|z_2 - z_1\|_{\mathcal{Z}}^2\ ,\end{aligned} \tag{2.50}$$

wobei

$$c_\kappa(\chi_i,\,r_i) = \frac{1}{2}\ \min\left\{\,m_C^2\,c_\lambda\,,\ m_N^2\,c_u\,\lambda_i\,\right\}\ . \tag{2.51}$$

Offensichtlich gilt dasselbe Ergebnis für das Funktional $\bar{\Psi}_{i,s}$ in (2.13).

2.2.3 Existenz und Eindeutigkeit der Lösungen

Die Existenz einer Lösung der Probleme $\mathbf{P}$, $\bar{\mathbf{P}}_i^{(T)}$, $\bar{\bar{\mathbf{P}}}_i^{(T)}$, $\bar{\mathbf{P}}_{i,s}^{(MSR)}$ oder $\mathbf{P}_i$ beweisen wir am Beispiel des Problems $\mathbf{P}$, da die Beweise für die anderen Probleme analog geführt werden können.

Satz 2 *Das Problem* $\mathbf{P}$ *hat wenigstens eine Lösung, d.h.* $\mathcal{Z}^* \neq \emptyset$.

Beweis: Mit Rücksicht auf (2.9) definieren wir die nichtleere Menge

$$\mathcal{Z}_0 = \{z = (y,\, u) \in \mathcal{Z}_{ad} :\; A\,y = B\,u + f \;\text{ in } \mathcal{H}\} \neq \emptyset \,. \tag{2.52}$$

Da der Operator A einen nicht nullidentischen Kern hat, kann die Menge $\mathcal{Z}_0$ unbeschränkt sein, obwohl die Menge der zulässigen Steuerungen $\mathcal{U}_{ad}$ in $\mathcal{U}$ nach (2.8) beschränkt ist. Wir betrachten eine minimierende Folge $\{z^n\}_{n=1}^{\infty} \in \mathcal{Z}_0$:

$$J(z^0) \geq J(z^n) = \frac{1}{2}\,\|C\,y^n\|_{\mathcal{K}}^2 \to \inf_{(y,u)\in\mathcal{Z}_0}\,\frac{1}{2}\,\|C\,y\|_{\mathcal{K}}^2 \;. \tag{2.53}$$

Nach (2.6) kann man feststellen, daß die eingeführte Folge beschränkt ist.

$$\|y^n\|_{\mathcal{Y}}^2 \leq \frac{1}{m_C^2}\,\left(\|C\,y^n\|_{\mathcal{K}}^2 + \|A\,y^n\|_{\mathcal{H}}^2\right) \leq$$

$$\frac{1}{m_C^2}\,\left(2\,J(y^n,\,u^n) + \left(\|B\,u^n\|_{\mathcal{H}} + \|f\|_{\mathcal{H}}\right)^2\right) \leq$$

$$\frac{1}{m_C^2}\,\left(2\,J(y^0,\,u^0) + \left(M_B M_{ad} + \|f\|_{\mathcal{H}}\right)^2\right) \;.$$

Somit gilt $\|z^n\|_{\mathcal{Z}} \leq R_0$ mit

$$R_0 = \sqrt{\frac{1}{m_C^2}\,\left(2\,J(z^0) + \left(M_B M_{ad} + \|f\|_{\mathcal{H}}\right)^2\right) + M_{ad}^2} \;. \tag{2.54}$$

Jetzt können wir eine schwach konvergierende Teilfolge $z^{n_k} \xrightarrow{w} z_1$ in $\mathcal{Z}$ finden. Da die Menge $\mathcal{Z}_0 \cap \overline{U_{R_0}(0)}$ beschränkt, abgeschlossen und konvex ist, ist sie auch schwach abgeschlossen [97, Kap. 1, §3, Satz 4] . Daraus folgt $z_1 \in \mathcal{Z}_0$. Wir weisen nun nach, daß $z_1 \in \mathcal{Z}^*$, d.h. auf z_1 ein Infimum des Zielfunktionals $J(z)$ erreicht wird. Das Funktional $J(z)$ ist konvex und stetig auf der konvexen Menge $\mathcal{Z}_0$, da der lineare Operator $C \in \mathcal{L}(\mathcal{Y},\,\mathcal{K})$ beschränkt ist. Die Unterhalbstetigkeit folgt unmittelbar aus der Stetigkeit und nach [97, Kap. 1, §3, Satz 5] ist $J(z)$ schwach unterhalbstetig. Somit gilt

$$J(z_1) \leq \liminf_{k\to\infty}\, J(z^{n_k}) = \inf_{z\in\mathcal{Z}_0}\, J(z) \,. \tag{2.55}$$

Nach der Definition des Infimums muß in (2.55) ein Gleichheitszeichen stehen. Wir nehmen also $z^* = z_1$.

□

Die Lösungen von $\mathbf{P}$ und $\mathbf{P}_i$ sind im allgemeinen *nichteindeutig*. Die Eindeutigkeit der Lösungen von $\bar{\mathbf{P}}_i^{(T)}$ und $\bar{\mathbf{P}}_{i,s}^{(MSR)}$ begründet sich aus der im Abschnitt 2.2.2 bewiesenen starken Konvexität des entsprechenden Hilfsfunktionals.

Satz 3 *Das Problem $\bar{\bar{\mathbf{P}}}_i^{(T)}$ hat höchstens eine Lösung z^* mit dem Zielfunktionswert $\bar{\bar{\Psi}}_i^*$.*

Beweis: Seien $z^{(1)}$ und $z^{(2)}$ unterschiedliche Lösungen von $\bar{\bar{\mathbf{P}}}_i^{(T)}$. Es gelten die Gleichheiten

$$\bar{\Psi}_i(z^{(1)}) = \bar{\Psi}_i(z^{(2)}) = \bar{\bar{\Psi}}_i(z^{(1)}) = \bar{\bar{\Psi}}_i(z^{(2)}) = \bar{\bar{\Psi}}_i^* \; . \tag{2.56}$$

Das Element $(z^{(1)} + z^{(2)})/2$ ist in der Menge $\mathcal{Z}_0$ (s. (2.52)) enthalten. Das Zielfunktional $\bar{\Psi}_i(z)$ ist stark konvex und man kann mit Hilfe von (2.56) schreiben:

$$\bar{\bar{\Psi}}_i\left(\frac{z^{(1)} + z^{(2)}}{2}\right) = \bar{\Psi}_i\left(\frac{z^{(1)} + z^{(2)}}{2}\right) < \frac{1}{2}\,\bar{\Psi}_i(z^{(1)}) + \frac{1}{2}\,\bar{\Psi}_i(z^{(2)}) = \bar{\bar{\Psi}}_i^* \; . \tag{2.57}$$

Die letzte Aussage bedeutet, daß weder $z^{(1)}$ noch $z^{(2)}$ Lösung von $\bar{\bar{\mathbf{P}}}_i^{(T)}$ ist.

□

Aus Satz 2 und Satz 3 erhält man, daß die beiden Hilfsprobleme $\bar{\mathbf{P}}_i^{(T)}$ und $\bar{\bar{\mathbf{P}}}_i^{(T)}$ eine eindeutige Lösung besitzen. Nun schätzen wir den Radius dieser Lösung ab. Wir benötigen zwei neue Konstanten, deren Existenz bei der Formulierung der Konvergenzsätze in Abschnitt 2.3 gesichert wird:

$$\bar{\chi} = \sup_i \chi_i \; , \qquad \bar{r} = \sup_i r_i \; . \tag{2.58}$$

Satz 4 *Für die Elemente $\bar{y}^i$ und $\bar{\bar{y}}^i$ der Hilfsprobleme $\bar{\mathbf{P}}_i^{(T)}$ und $\bar{\bar{\mathbf{P}}}_i^{(T)}$ gilt:*

$$\left\|\bar{y}^i\right\|_{\mathcal{Y}} \le c_y \; , \qquad \left\|\bar{\bar{y}}^i\right\|_{\mathcal{Y}} \le c_y \; , \tag{2.59}$$

wobei

$$c_y = \frac{1}{m_C}\sqrt{\left(c_r\,\frac{M_C^2}{m_A^2} + \frac{2\,M_W^2}{m_W^2}\right)(M_B M_{ad} + \|f\|_{\mathcal{H}})^2 + c_r\,\bar{\chi}\,(M_N M_{ad} + \|t_d\|_{\mathcal{K}})^2} \; , \tag{2.60}$$

$$c_r = \max\left\{\,1\,,\;\frac{2\,\bar{r}}{m_W^2}\,\right\} \; .$$

Beweis: Die ersten Abschätzungen sind ähnlich wie im Beweis des Satzes 2 .

$$\left\|\bar{y}^i\right\|_{\mathcal{Y}}^2 \leq \frac{1}{m_C^2}\ \left(\left\|C\,\bar{y}^i\right\|_{\mathcal{K}}^2 + \frac{1}{m_W^2}\ \left\|W A\,\bar{y}^i\right\|_{\mathcal{S}}^2\right) \leq$$

$$\frac{1}{m_C^2}\ \left(\left\|C\,\bar{y}^i\right\|_{\mathcal{K}}^2 + \frac{2}{m_W^2}\ \left(\left\|W(A\,\bar{y}^i - B\,\bar{u}^i - f)\right\|_{\mathcal{S}}^2 + M_W^2\ \left\|B\,\bar{u}^i + f\right\|_{\mathcal{H}}^2\right)\right) \leq$$

$$\frac{1}{m_C^2}\ \left(c_r\ \left(\left\|C\,\bar{y}^i\right\|_{\mathcal{K}}^2 + \frac{1}{r_i}\ \left\|W(A\bar{y}^i - B\bar{u}^i - f)\right\|_{\mathcal{S}}^2 + \right.\right.$$

$$\left.\left.\chi_i\ \left\|N\bar{u}^i - t_d\right\|_{\mathcal{T}}^2\right) + \frac{2\,M_W^2}{m_W^2}\ \left(M_B M_{ad} + \|f\|_{\mathcal{H}}\right)^2\right) \leq$$

$$\frac{1}{m_C^2}\ \left(2c_r\ \bar{\bar{\Psi}}_i(\tilde{y},\,\tilde{u}) + \frac{2\,M_W^2}{m_W^2}\ \left(M_B M_{ad} + \|f\|_{\mathcal{H}}\right)^2\right)\ .$$

Bei der letzten Abschätzung haben wir die offenbare Relation $\bar{\Psi}_i(\bar{y}^i,\,\bar{u}^i) \leq \bar{\Psi}_i(\tilde{y},\,\tilde{u}) = \bar{\bar{\Psi}}_i(\tilde{y},\,\tilde{u})$ mit $\tilde{y} = G_y(AG_y)^{-1}(B\tilde{u} + f)$ aus (2.9) ausgenutzt. Um jetzt das Endergebnis (2.60) zu erhalten, müssen wir erst den Wert $\bar{\bar{\Psi}}_i(\tilde{y},\,\tilde{u})$ nach oben abschätzen. Da $(\tilde{y},\,\tilde{u})$ die Zustandsgleichung genau erfüllt, gilt

$$2\,\bar{\bar{\Psi}}_i(\tilde{y},\,\tilde{u}) = \left\|CG_y(AG_y)^{-1}(B\tilde{u} + f)\right\|_{\mathcal{K}}^2 + \chi_i\ \left\|N\bar{u}^i - t_d\right\|_{\mathcal{T}}^2 \leq$$

$$\frac{M_C^2}{m_A^2}\ \left(M_B M_{ad} + \|f\|_{\mathcal{H}}\right)^2 + \bar{\chi}\ \left(M_N M_{ad} + \|t_d\|_{\mathcal{T}}\right)^2\ . \tag{2.61}$$

Mit dem Element $\bar{\bar{y}}^i$ kann man analog verfahren.

□

2.2.4 Der Fall ohne Zustandsrestriktionen

In den folgenden Abschnitten untersuchen wir einige Eigenschaften des Strafverfahrens. Die Ergebnisse ohne Zustandsrestriktionen werden gesondert betrachtet, weil man in diesem Fall stärkere Abschätzungen erreichen kann. Auf der Lösung $(\bar{y}^i,\,\bar{u}^i)$ des Hilfsproblems $\bar{\mathbf{P}}_i^{(T)}$ definieren wir ein neues Element:

$$\bar{q}^i \ = \ \frac{1}{r_i}\ \left(A\,\bar{y}^i - B\,\bar{u}^i - f\right) \qquad \text{in } \ \mathcal{H}\ . \tag{2.62}$$

Lemma 5 *Es gilt*

$$\exists\, c_0 > 0 \;:\qquad \left\| r_i \bar{q}^i \right\|_{\mathcal{H}} \;=\; \left\| A\,\bar{y}^i - B\,\bar{u}^i - f \right\|_{\mathcal{H}} \le c_0\,\sqrt{r_i} \quad \forall\, i\;, \tag{2.63}$$

wobei

$$c_0 = \sqrt{\frac{M_C^2}{m_A^2}\;\left(M_B M_{ad} + \|f\|_{\mathcal{H}}\right)^2 + \bar{\chi}\;\left(M_N M_{ad} + \|t_d\|_{\mathcal{T}}\right)^2}\;. \tag{2.64}$$

Beweis: Das Funktional $\bar{\bar{\Psi}}_i(z)$ besitzt folgende Eigenschaften:

$$\bar{\bar{\Psi}}_i(\bar{y}^i,\,\bar{u}^i) \;=\; \bar{\Psi}_i(\bar{y}^i,\,\bar{u}^i) - \frac{1}{2r_i}\;\left\| A\,\bar{y}^i - B\,\bar{u}^i - f \right\|_{\mathcal{H}}^2\;, \tag{2.65}$$

$$\bar{\bar{\Psi}}_i(\bar{\bar{y}}^i,\,\bar{\bar{u}}^i) \;=\; \bar{\Psi}_i(\bar{\bar{y}}^i,\,\bar{\bar{u}}^i)\;, \tag{2.66}$$

wobei $\bar{\bar{z}}^i = (\bar{\bar{y}}^i,\,\bar{\bar{u}}^i)$ die eindeutige Lösung des Hilfsproblems $\bar{\bar{\mathbf{P}}}_i^{(T)}$ ist. Subtrahieren wir (2.65) von (2.66) ,

$$\begin{aligned}\bar{\bar{\Psi}}_i(\bar{\bar{y}}^i,\,\bar{\bar{u}}^i) - \bar{\bar{\Psi}}_i(\bar{y}^i,\,\bar{u}^i) \;&=\; \bar{\Psi}_i(\bar{\bar{y}}^i,\,\bar{\bar{u}}^i) - \bar{\Psi}_i(\bar{y}^i,\,\bar{u}^i) + \\ &\qquad \frac{1}{2r_i}\;\left\| A\,\bar{y}^i - B\,\bar{u}^i - f \right\|_{\mathcal{H}}^2\;,\end{aligned}$$

so folgt

$$\left\| A\,\bar{y}^i - B\,\bar{u}^i - f \right\|_{\mathcal{H}} \le \sqrt{2r_i\,\bar{\bar{\Psi}}_i(\bar{\bar{y}}^i,\,\bar{\bar{u}}^i)} \le \sqrt{2r_i\,\bar{\bar{\Psi}}_i(\tilde{y},\,\tilde{u})} \le \sqrt{r_i}\,c_0\;.$$

Den letzten Übergang erhalten wir mit (2.61) .

□

Lemma 6 *Die dem Zustand $\bar{y}^i$ entsprechende adjungierte Funktion ist beschränkt:*

$$\exists\, c_p > 0 \;:\qquad \left\| \bar{p}^i \right\|_{\mathcal{H}'} \;\le\; c_p \quad \forall\, i\;, \tag{2.67}$$

wobei

$$c_p \;=\; \frac{M_C^2}{m_A}\,c_y\;. \tag{2.68}$$

Beweis: Das Ergebnis (2.68) folgt unmittelbar aus (2.27) und aus dem Satz 4 .

□

Nun versuchen wir, ein stärkeres Resultat zu erhalten. Der nächste Satz stützt sich dabei wesentlich auf das Nichtvorhandensein von Zustandsrestriktionen, d.h. $\mathcal{Y}_{ad} = \mathcal{Y}$. Der Fall mit Zustandsrestriktionen wird im Punkt 2.2.5 betrachtet.

Satz 5 *Das Element $\bar{q}^i$ ist beschränkt:*

$$\exists\, \hat{c}_q > 0 \;:\qquad \left\|\bar{q}^i\right\|_{\mathcal{H}} \;\leq\; \hat{c}_q \quad \forall\, i\,, \tag{2.69}$$

wobei

$$\hat{c}_q = c_p \left(\frac{1}{m_W^2} + c_0\,\sqrt{\bar{r}}\right)\,. \tag{2.70}$$

Beweis : Für das Problem $\bar{\mathbf{P}}_i^{(T)}$ lauten die Optimalitätsbedingungen der ersten Ordnung im Optimalpunkt $\bar{z}^i$:

$$\begin{aligned}
&\forall\, u \in \mathcal{U}_{ad}\;\forall\, y \in \mathcal{Y}\;:\\
&\left(C\,\bar{y}^i,\, C\,\left(y-\bar{y}^i\right)\right)_{\mathcal{K}} + \frac{1}{r_i}\,\left(W\,\left(A\,\bar{y}^i - B\,\bar{u}^i - f\right),\, WA\,\left(y-\bar{y}^i\right)\right)_{\mathcal{S}} \geq 0\,,\\
&-\frac{1}{r_i}\,\left(W\,\left(A\,\bar{y}^i - B\,\bar{u}^i - f\right),\, WB\,\left(u-\bar{u}^i\right)\right)_{\mathcal{S}} +\\
&\qquad\qquad \chi_i\,\left(N\,\bar{u}^i - t_d,\, N\,\left(u-\bar{u}^i\right)\right)_{\mathcal{T}} \geq 0\,,
\end{aligned}$$

oder mit Rücksicht auf (2.62)

$$\left(C\,\bar{y}^i,\, C\,\left(y-\bar{y}^i\right)\right)_{\mathcal{K}} + \left(W\,\bar{q}^i,\, WA\,\left(y-\bar{y}^i\right)\right)_{\mathcal{S}} \geq 0\,, \tag{2.71}$$

$$\chi_i\,\left(N\,\bar{u}^i - t_d,\, N\,\left(u-\bar{u}^i\right)\right)_{\mathcal{T}} - \left(W\,\bar{q}^i,\, WB\,\left(u-\bar{u}^i\right)\right)_{\mathcal{S}} \geq 0\,. \tag{2.72}$$

Wir wählen in (2.71)

$$y = G_y(AG_y)^{-1}\left(B\,\bar{u}^i - \bar{h}^i + f\right) + \left(I_y - G_y(AG_y)^{-1}A\right)\,\bar{y}^i$$

mit

$$\bar{h}^i = \bar{q}^i\,\frac{\|\bar{q}^i\|_{\mathcal{H}}}{\|W\,\bar{q}^i\|_{\mathcal{S}}^2}\,, \tag{2.73}$$

und gehen zu den Dualitätsklammern über

$$\begin{aligned}
&\left\langle\, G_y^* C^* \Lambda_{\mathcal{K}} C\,\bar{y}^i,\, (AG_y)^{-1}\left[\left(B\,\bar{u}^i - \bar{h}^i + f\right) - A\,\bar{y}^i\right]\,\right\rangle_{\mathcal{Y}_0',\,\mathcal{Y}_0} \;+\\
&\qquad\left\langle\, W^*\Lambda_{\mathcal{S}} W\,\bar{q}^i,\, A\,\left[\left(B\,\bar{u}^i - \bar{h}^i + f\right) - A\,\bar{y}^i\right]\,\right\rangle_{\mathcal{H}',\,\mathcal{H}} \;\geq\; 0\,.
\end{aligned} \tag{2.74}$$

Nach (2.16) gilt der Zusammenhang $G_y^*\,(C^*\Lambda_{\mathcal{K}} C\,\bar{y}^i) = G_y^* A^*\,\bar{p}^i$ in $\mathcal{Y}_0'$. Somit kann man (2.74) weiter entwickeln:

$$
\begin{aligned}
-\left(\,\Lambda_{\mathcal{H}}^{-1}\,\left(\bar{p}^i + W^*\Lambda_{\mathcal{S}} W\,\bar{q}^i\right),\,\bar{h}^i + r_i\bar{q}^i\,\right)_{\mathcal{H}} \;&\geq\; 0\,, \\
-\left(\,\Lambda_{\mathcal{H}}^{-1}\,\bar{p}^i,\,\bar{q}^i\,\right)_{\mathcal{H}}\,\frac{\|\bar{q}^i\|_{\mathcal{H}}}{\|W\,\bar{q}^i\|_{\mathcal{S}}^2} - r_i\,\left(\,\Lambda_{\mathcal{H}}^{-1}\,\bar{p}^i,\,\bar{q}^i\,\right)_{\mathcal{H}} - r_i\,\left\|W\,\bar{q}^i\right\|_{\mathcal{S}}^2 \;&\geq\; \left\|\bar{q}^i\,\right\|_{\mathcal{H}}\,, \\
\left(\frac{1}{m_W^2} + \left\|r_i\bar{q}^i\,\right\|_{\mathcal{H}}\right)\,\left\|\bar{p}^i\,\right\|_{\mathcal{H}'} \geq \left(\frac{\|\bar{q}^i\|_{\mathcal{H}}}{\|W\,\bar{q}^i\|_{\mathcal{S}}^2} + r_i\right)\,\left\|\bar{p}^i\,\right\|_{\mathcal{H}'}\left\|\bar{q}^i\,\right\|_{\mathcal{H}} \;&\geq\; \left\|\bar{q}^i\,\right\|_{\mathcal{H}}\,.
\end{aligned}
$$

Nach den Vorbereitungen in Lemma 5 und Lemma 6 erhalten wir das gewünschte Endergebnis.

□

Um ein Probeelement aus der Menge $\mathcal{Z}_0$ zu bestimmen, betrachten wir

$$
\hat{\bar{y}}^i = \bar{y}^i - r_i\,G_y(AG_y)^{-1}\bar{q}^i \qquad \text{in } \mathcal{Y}\,. \tag{2.75}
$$

Es ist nicht schwierig zu sehen: das Element $\hat{\bar{y}}^i$

(i) erfüllt die Zustandsgleichung

$$
A\,\hat{\bar{y}}^i = B\,\bar{u}^i + f \qquad \text{in } \mathcal{H}\,, \tag{2.76}
$$

(ii) liegt "nah" bei dem Element $\bar{y}^i$

$$
\left\|\hat{\bar{y}}^i - \bar{y}^i\right\|_{\mathcal{Y}} \leq \frac{\hat{c}_q}{m_A}\,r_i\,. \tag{2.77}
$$

Die Ungleichung (2.77) folgt aus (2.4). Nun definieren wir das Probeelement auf folgende Weise

$$
\hat{\bar{z}}^i = (\hat{\bar{y}}^i,\,\bar{u}^i) \in \mathcal{Z}_0: \qquad \left\|\hat{\bar{z}}^i - \bar{z}^i\right\|_{\mathcal{Z}} \leq \frac{\hat{c}_q}{m_A}\,r_i\,. \tag{2.78}
$$

Die letzte Bedingung gilt nach (2.77) .

2.2.5 Der Fall mit Zustandsrestriktionen

Lemmata 5 und 6 gelten auch im Falle von Zustandsrestriktionen, das heißt $\mathcal{Y}_{ad} \subset \mathcal{Y}$ sei eine beliebige konvexe Menge, so daß die Bedingung (2.9) erfüllt ist. Wir benutzen die Definitionen (2.62) und (2.73) wie im Punkt 2.2.4 und beweisen unter neuen Voraussetzungen den Satz 5 .

Satz 6 *Das Element $\bar{q}^i$ ist beschränkt:*

$$\exists\ \hat{\hat{c}}_q > 0\ :\qquad \left\|\bar{q}^i\right\|_{\mathcal{H}}\ \leq\ \hat{\hat{c}}_q\qquad \forall\, i\ , \tag{2.79}$$

wobei

$$\begin{aligned}\hat{\hat{c}}_q\ &=\ \frac{1}{\tau}\left(c_y^2 M_C^2\left(1+\frac{M_A}{m_A}\right)+c_p\left(2M_B M_{ad}+\frac{\tau}{m_W^2}+c_0\sqrt{\bar{r}}\right)+\right.\\ &\qquad\left.2\bar{\chi}M_N M_{ad}\left(M_N M_{ad}+\|t_d\|_{\mathcal{T}}\right)\right)\ ,\end{aligned} \tag{2.80}$$

$$\tau\ =\ m_W^2 m_A m_{ad}\ . \tag{2.81}$$

Beweis : Wir addieren (2.71) und (2.72) und ersetzen dabei $u := \tilde{u}$ und

$$y := G_y(AG_y)^{-1}\left(B\,\tilde{u}+f-\tau\bar{h}^i\right)\ . \tag{2.82}$$

Die Größe τ läßt sich so klein wählen, daß $y \in \mathcal{Y}_{ad}$. Wir wollen jetzt diese Größe nach oben abschätzen. Es gilt gemäß der Definition von $\tilde{u}$ in (2.9) folgende Implikation

$$\left\|G_y(AG_y)^{-1}\left(B\,\tilde{u}+f\right)-y\right\|_{\mathcal{Y}}\leq m_{ad}\quad\Rightarrow\quad y\in\mathcal{Y}_{ad}\ .$$

Wir setzen in den letzten Ausdruck das Element y aus (2.82) ein und erhalten

$$\tau\,\left\|G_y(AG_y)^{-1}\bar{h}^i\right\|_{\mathcal{Y}}\leq m_{ad}\ . \tag{2.83}$$

Das Ergebnis (2.81) folgt aus (2.83) mit (2.4) und (2.11) . Wir führen nun die bereits angekündigte Addition aus und erhalten die folgende Ungleichung:

$$\begin{aligned}&\left(C\,\bar{y}^i,\,C\,\left(G_y(AG_y)^{-1}\left[B\,\tilde{u}+f-\tau\bar{h}^i\right]-\bar{y}^i\right)\right)_{\mathcal{K}}+\\ &\left(W\,\bar{q}^i,\,WA\,\left(G_y(AG_y)^{-1}\left[B\,\tilde{u}+f-\tau\bar{h}^i\right]-\bar{y}^i\right)\right)_{\mathcal{S}}+\\ &\chi_i\,\left(N\,\bar{u}^i-t_d,\,N\,\left(u-\bar{u}^i\right)\right)_{\mathcal{T}}-\left(W\,\bar{q}^i,\,WB\,\left(u-\bar{u}^i\right)\right)_{\mathcal{S}}\geq 0\ .\end{aligned} \tag{2.84}$$

Für $\bar{y}^i$ im ersten Summanden von (2.84) benutzen wir folgende Identität:

$$\bar{y}^i = G_y(AG_y)^{-1}A\,\bar{y}^i+\left(I_y-G_y(AG_y)^{-1}A\right)\,\bar{y}^i\ . \tag{2.85}$$

Mit (2.4) und (2.2), zusammen mit Satz 4, gilt

$$\left\|\left(I_y-G_y(AG_y)^{-1}A\right)\,\bar{y}^i\right\|_{\mathcal{Y}}\leq\left\|\bar{y}^i\right\|_{\mathcal{Y}}+\frac{M_A}{m_A}\,\left\|\bar{y}^i\right\|_{\mathcal{Y}}\leq c_y\,\left(1+\frac{M_A}{m_A}\right)\ . \tag{2.86}$$

Jetzt setzen wir (2.85) in (2.84) ein und berücksichtigen (2.86). Gleichzeitig gehen wir in (2.84) zu den Dualitätsklammern über

$$
\begin{aligned}
& c_y^2 M_C^2 \left(1 + \frac{M_A}{m_A}\right) + \\
& \left\langle G_y^* C^* \Lambda_{\mathcal{K}} C\, \bar{y}^i,\, (AG_y)^{-1} \left[B\,\tilde{u} + f - \tau \bar{h}^i - A\,\bar{y}^i\right] \right\rangle_{\mathcal{Y}_0', \mathcal{Y}_0} + \\
& \left\langle W^* \Lambda_{\mathcal{S}} W\, \bar{q}^i,\, B\,\tilde{u} + f - \tau \bar{h}^i - A\,\bar{y}^i \right\rangle_{\mathcal{H}', \mathcal{H}} + \\
\chi_i & \left\langle N^* \Lambda_{\mathcal{T}} \left(N\,\bar{u}^i - t_d\right),\, u - \bar{u}^i \right\rangle_{\mathcal{U}', \mathcal{U}} - \\
& \left\langle B^* W^* \Lambda_{\mathcal{S}} W\, \bar{q}^i,\, u - \bar{u}^i \right\rangle_{\mathcal{U}', \mathcal{U}} \geq 0 \,. \qquad (2.87)
\end{aligned}
$$

Nach (2.16) gilt der Zusammenhang $G_y^* \left(C^* \Lambda_{\mathcal{K}} C\, \bar{y}^i\right) = G_y^* A^*\, \bar{p}^i$ in $\mathcal{Y}_0'$. Somit kann man (2.87) weiter vereinfachen:

$$
\begin{aligned}
& c_y^2 M_C^2 \left(1 + \frac{M_A}{m_A}\right) + \\
& \left(\Lambda_{\mathcal{H}}^{-1} \left(\bar{p}^i + W^* \Lambda_{\mathcal{S}} W\, \bar{q}^i\right),\, B\,(\tilde{u} - \bar{u}^i) - \tau \bar{h}^i - r_i \bar{q}^i \right)_{\mathcal{H}} + \\
& \left(\Lambda_{\mathcal{U}}^{-1} \left(\chi_i\, N^* \Lambda_{\mathcal{T}} \left(N\,\bar{u}^i - t_d\right) - B^* W^* \Lambda_{\mathcal{S}} W\, \bar{q}^i\right),\, u - \bar{u}^i \right)_{\mathcal{U}} \geq 0 \,. \qquad (2.88)
\end{aligned}
$$

Die Ausdrücke mit $(W\,\bar{q}^i,\, WB\,(u - \bar{u}^i))_{\mathcal{H}}$ heben sich in (2.88) gegenseitig auf. Es bleibt übrig

$$
\begin{aligned}
& c_y^2 M_C^2 \left(1 + \frac{M_A}{m_A}\right) + \left(\Lambda_{\mathcal{H}}^{-1}\, \bar{p}^i,\, B\,(\tilde{u} - \bar{u}^i) - \tau \bar{h}^i - r_i \bar{q}^i \right)_{\mathcal{H}} - \\
& \tau \left(\Lambda_{\mathcal{H}}^{-1} W^* \Lambda_{\mathcal{S}} W\, \bar{q}^i,\, \bar{h}^i \right)_{\mathcal{H}} - r_i \left(\Lambda_{\mathcal{H}}^{-1} W^* \Lambda_{\mathcal{S}} W\, \bar{q}^i,\, \bar{q}^i \right)_{\mathcal{H}} + \\
& \chi_i \left(\Lambda_{\mathcal{U}}^{-1}\, N^* \Lambda_{\mathcal{T}} \left(N\,\bar{u}^i - t_d\right),\, u - \bar{u}^i \right)_{\mathcal{U}} \geq 0 \,. \qquad (2.89)
\end{aligned}
$$

Weitere Umformungen von (2.89) erfolgen mit (2.73) und der Cauchy–Schwarz Ungleichung

$$
\begin{aligned}
& c_y^2 M_C^2 \left(1 + \frac{M_A}{m_A}\right) + 2 \left\|\bar{p}^i\right\|_{\mathcal{H}'} M_B M_{ad} - \\
& \tau \left(\Lambda_{\mathcal{H}}^{-1} \bar{p}^i,\, \bar{q}^i\right)_{\mathcal{H}} \frac{\|\bar{q}^i\|_{\mathcal{H}}}{\|W\,\bar{q}^i\|_{\mathcal{S}}^2} - r_i \left(\Lambda_{\mathcal{H}}^{-1}\, \bar{p}^i,\, \bar{q}^i\right)_{\mathcal{H}} + \\
& \chi_i \left(N\,\bar{u}^i - t_d,\, N\,(u - \bar{u}^i)\right)_{\mathcal{T}} \geq \tau \left\|\bar{q}^i\right\|_{\mathcal{H}} \,,
\end{aligned}
$$

$$\left\|\bar{q}^i\right\|_{\mathcal{H}} \le \frac{1}{\tau}\left(c_y^2 M_C^2\left(1+\frac{M_A}{m_A}\right)+2c_p\,M_B M_{ad}+\right.$$

$$\left.\left(\frac{\tau}{m_W^2}+\left\|r_i\bar{q}^i\right\|_{\mathcal{H}}\right)\left\|\bar{p}^i\right\|_{\mathcal{H}'}+2\bar{\chi}M_N M_{ad}\,(M_N M_{ad}+\|t_d\|_{\mathcal{T}})\right)\ .$$

Zum Abschluß wendet man die Lemmata 5 und 6 an.

□

Wir betrachten erneut das Element $\hat{y}^i \in \mathcal{Y}$ nach (2.75) und bemerken, daß die Gültigkeit von (2.76) und (2.77) erhalten bleibt. Falls das Element $\hat{z}^i$ nach (2.78) zur Menge $\mathcal{Z}_0$ (2.52) gehört, können wir es wie zuvor als Probeelement wählen. Dabei gelte aber gemäß (2.79)

$$\hat{z}^i = (\hat{y}^i,\,\bar{u}^i) \in \mathcal{Z}_0 : \qquad \left\|\hat{z}^i - \bar{z}^i\right\|_{\mathcal{Z}} \le \frac{\hat{\hat{c}}_q}{m_A}\,r_i\ . \tag{2.90}$$

Nun nehmen wir an, daß $\hat{z}^i$ nicht zur Menge $\mathcal{Z}_0$ gehört. Wir führen dann zwei neue Elemente $\hat{\hat{z}}^i$ und w^i in $\mathcal{Z}$ ein:

$$\hat{\hat{z}}^i \in \{\tilde{z}+\lambda(\hat{z}^i-\tilde{z}),\ \lambda\ge 0\}\cap\{\partial\mathcal{Y}_{ad}\times\mathcal{U}_{ad}\}\ , \tag{2.91}$$

$$w^i = \operatorname*{argmin}_{z\in\mathcal{Z}_{ad}}\left\|z-\hat{z}^i\right\|_{\mathcal{Z}}\ , \tag{2.92}$$

mit dem der Steuerung $\tilde{u}$ in (2.9) entsprechenden Element aus $\mathcal{Z}$

$$\tilde{z} = \left(G_y(AG_y)^{-1}(B\,\tilde{u}+f),\,\tilde{u}\right)\ . \tag{2.93}$$

Da die Menge $\mathcal{Y}_{ad}$ konvex ist und $G_y(AG_y)^{-1}(B\,\tilde{u}+f)$ nach (2.9) ein innerer Punkt ist, ist die Definition von $\hat{\hat{z}}^i$ eindeutig. Wenn $\hat{\hat{z}}^i = w^i$ in $\mathcal{Z}$, dann erhalten wir mit (2.92) und (2.90) folgende Beziehungen:

$$\left\|\hat{\hat{z}}^i-\bar{z}^i\right\|_{\mathcal{Z}} \le \left\|\hat{\hat{z}}^i-\hat{z}^i\right\|_{\mathcal{Z}}+\left\|\hat{z}^i-\bar{z}^i\right\|_{\mathcal{Z}} =$$

$$\left\|w^i-\hat{z}^i\right\|_{\mathcal{Z}}+\left\|\bar{z}^i-\hat{z}^i\right\|_{\mathcal{Z}} \le 2\left\|\bar{z}^i-\hat{z}^i\right\|_{\mathcal{Z}} \le \frac{2\ \hat{\hat{c}}_q}{m_A}\,r_i\ . \tag{2.94}$$

Sei jetzt $\hat{\hat{z}}^i \ne w^i$ in $\mathcal{Z}$. Dann führen wir ein drittes Element v^i in $\mathcal{Z}$ ein:

$$v^i \in \{\tilde{z}+\lambda(\hat{z}^i-w^i),\ \lambda\ge 0\}\cap\{\hat{\hat{z}}^i+\mu(\hat{\hat{z}}^i-w^i),\ \mu\ge 0\}\ .$$

Man sieht leicht, daß $v^i = (v_y^i,\,v_u^i)$ das einzige Element in der nichtleeren Menge ist. Für die fünf Elemente : $\hat{z}^i$, $\hat{\hat{z}}^i$, w^i, v^i, $\tilde{z}$ gilt

$$\frac{\|\hat{z}^i-w^i\|_{\mathcal{Z}}}{\|v^i-\tilde{z}\|_{\mathcal{Z}}} = \frac{\|\hat{z}^i-\hat{\hat{z}}^i\|_{\mathcal{Z}}}{\|\hat{\hat{z}}^i-\tilde{z}\|_{\mathcal{Z}}}\ . \tag{2.95}$$

Für vier beliebige positive Zahlen gilt $\frac{\alpha}{\beta} = \frac{\gamma}{\delta} \Rightarrow \frac{\alpha}{\alpha+\beta} = \frac{\gamma}{\gamma+\delta}$, d.h. aus (2.95) folgt

$$\left\| \hat{z}^i - \hat{\hat{z}}^i \right\|_{\mathcal{Z}} = \frac{\|\hat{z}^i - \tilde{z}\|_{\mathcal{Z}}}{\|v^i - \tilde{z}\|_{\mathcal{Z}} + \|\hat{z}^i - w^i\|_{\mathcal{Z}}} \left\| \hat{z}^i - w^i \right\|_{\mathcal{Z}} . \tag{2.96}$$

Auf der rechten Seite von (2.96) schätzen wir den Zähler nach oben ab. Dabei werden die Beziehungen (2.75), (2.4), (2.8), (2.3) und (2.60) berücksichtigt:

$$\left\| \hat{z}^i - \tilde{z} \right\|_{\mathcal{Z}} \leq \left\| \hat{y}^i - G_y(AG_y)^{-1}(B\,\tilde{u} + f) \right\|_{\mathcal{Y}} + \left\| \bar{u}^i - \tilde{u} \right\|_{\mathcal{U}} \leq$$

$$\left\| \bar{y}^i \right\|_{\mathcal{Y}} + \frac{1}{m_A} \left\| B\,(\tilde{u} - \bar{u}^i) + A\,\bar{y}^i \right\|_{\mathcal{Y}} + 2\,M_{ad} \leq$$

$$c_y + \frac{1}{m_A}\,(2\,M_B M_{ad} + c_y M_A) + 2\,M_{ad} . \tag{2.97}$$

Wir schätzen nun noch den Nenner von (2.96) nach unten ab. Da $\mathcal{Y}_{ad}$ konvex ist, gilt $v_y^i \notin \operatorname{int} \mathcal{Y}_{ad}$. Nach (2.9) ist

$$\left\| v^i - \tilde{z} \right\|_{\mathcal{Z}} \geq \left\| v_y^i - G_y(AG_y)^{-1}(B\,\tilde{u} + f) \right\|_{\mathcal{Y}} \geq m_{ad} . \tag{2.98}$$

Wenn man (2.96), (2.97) und (2.98) zusammenbringt, so erhält man unter Berücksichtigung von (2.92)

$$\left\| \hat{z}^i - \hat{\hat{z}}^i \right\|_{\mathcal{Z}} \leq c_w \left\| \hat{z}^i - w^i \right\|_{\mathcal{Z}} \leq c_w \left\| \hat{z}^i - \bar{z}^i \right\|_{\mathcal{Z}} , \tag{2.99}$$

mit

$$c_w \;=\; \frac{1}{m_{ad}} \left(c_y + 2\,M_{ad} + \frac{1}{m_A}\,(c_y M_A + 2\,M_B M_{ad}) \right) . \tag{2.100}$$

Wiederholen wir die Überlegen in (2.94) und nutzen dabei (2.90) und (2.99) aus, so folgt

$$\left\| \hat{\hat{z}}^i - \bar{z}^i \right\|_{\mathcal{Z}} \leq \left\| \hat{\hat{z}}^i - \hat{z}^i \right\|_{\mathcal{Z}} + \left\| \hat{z}^i - \bar{z}^i \right\|_{\mathcal{Z}} \leq$$

$$(1 + c_w) \left\| \bar{z}^i - \hat{z}^i \right\|_{\mathcal{Z}} \leq (1 + c_w)\,\frac{\hat{\hat{c}}_q}{m_A}\,r_i . \tag{2.101}$$

Sofern das Verfahren keine Information über die Lage der Iterationspunkte bezüglich der Menge $\mathcal{Y}_{ad} \times \mathcal{U}_{ad}$ ausnutzt, müssen wir die Ergebnisse (2.90), (2.94) und (2.101) zusammensetzen, um eine "universale" Abschätzung für das Probeelement $\hat{\hat{z}}^i \in \mathcal{Z}_0$ zu erhalten:

$$\left\| \hat{\hat{y}}^i - \bar{y}^i \right\|_{\mathcal{Y}} \leq \left\| \hat{\hat{z}}^i - \bar{z}^i \right\|_{\mathcal{Z}} \leq r_i\,\frac{\hat{\hat{c}}_q}{m_A}\,\max\{\,2\,,\;(1 + c_w)\}\,. \tag{2.102}$$

2.3 Konvergenzsätze für Tichonov- und Proximal-Punkt-Verfahren

2.3.1 Verfahren der Tichonov Regularisierung

Zunächst untersuchen wir die Konvergenz der Lösungen der Hilfsprobleme $\overline{\overline{\mathbf{P}}}_i^{(T)}$ (s. Seite 12).

Lemma 7 *Die Folge $\overline{\overline{z}}^i$ der exakten Lösungen der Hilfsprobleme $\overline{\overline{\mathbf{P}}}_i^{(T)}$ konvergiert stark gegen das Element*

$$z_N^* = (y_N^*,\, u_N^*) = \underset{(y,\,u)\in\mathcal{Z}^*}{\operatorname{argmin}} \|N\,u - t_d\|_{\mathcal{T}} \qquad in \;\; \mathcal{Z} \tag{2.103}$$

aus der Optimalmenge $\mathcal{Z}^*$, falls

$$\lim_{i\to\infty} \chi_i = 0\,. \tag{2.104}$$

Beweis : Zuerst weisen wir nach, daß das Element z_N^* durch die Definition (2.103) eindeutig bestimmt ist. Seien $(y_1^*,\, u_1^*) \in \mathcal{Z}^*$ und $(y_2^*,\, u_2^*) \in \mathcal{Z}^*$ unterschiedliche Lösungen von $\mathbf{P}$. Es gilt die Tatsache [57, Kap. I, §5, Satz 5.1] , daß die vorliegende Lösungsmenge $\mathcal{Z}^*$ konvex ist. Für das zu minimierende Funktional auf der rechten Seite in (2.103) gilt:

$$\|N\,u_1^* - t_d\|_{\mathcal{T}}^2 - \left\|N\,\left(\frac{u_1^*+u_2^*}{2}\right) - t_d\right\|_{\mathcal{T}}^2 = \frac{1}{2}\,\|N\,u_1^* - t_d\|_{\mathcal{T}}^2 + \frac{1}{2}\,\|N\,u_2^* - t_d\|_{\mathcal{T}}^2 -$$

$$\left\|N\,\left(\frac{u_1^*+u_2^*}{2}\right) - t_d\right\|_{\mathcal{T}}^2 = \frac{1}{4}\,\|N\,(u_1^* - u_2^*)\|_{\mathcal{T}}^2\;,$$

woraus folgt, daß $u_1^* = u_2^* = u^*$. Für die Zielfunktionalswerte erhält man:

$$2J^* = \|C\,(y_1^* + y_2^*)/2\|_{\mathcal{K}}^2 = \|C\,y_1^*\|_{\mathcal{K}}^2 = \|C\,y_2^*\|_{\mathcal{K}}^2 = \frac{3}{4}\,\|C\,y_1^*\|_{\mathcal{K}}^2 + \frac{1}{4}\,\|C\,y_2^*\|_{\mathcal{K}}^2\;;$$

$$\frac{1}{4}\,\|C\,y_1^*\|_{\mathcal{K}}^2 + \frac{1}{2}\,(C\,y_1^*,\, C\,y_2^*)_{\mathcal{K}} + \frac{1}{4}\,\|C\,y_2^*\|_{\mathcal{K}}^2 = \frac{3}{4}\,\|C\,y_1^*\|_{\mathcal{K}}^2 + \frac{1}{4}\,\|C\,y_2^*\|_{\mathcal{K}}^2\;;$$

$$(C\,y_1^*,\, C\,(y_2^* - y_1^*))_{\mathcal{K}} = 0\,. \tag{2.105}$$

Die folgenden Überlegungen stützen sich auf (2.105), (2.6) und auf die Tatsache, daß $A\,(y_2^* - y_1^*) = 0$ in $\mathcal{H}$.

$$\|C\,y_2^*\|_{\mathcal{K}}^2 = \|C\,y_1^* + C\,(y_2^* - y_1^*)\|_{\mathcal{K}}^2 = \|C\,y_1^*\|_{\mathcal{K}}^2 + \|C\,(y_2^* - y_1^*)\|_{\mathcal{K}}^2 \geq$$

$$\|C\,y_1^*\|_{\mathcal{K}}^2 + m_C^2\,\|y_2^* - y_1^*\|_{\mathcal{Y}}^2 - \|A\,(y_2^* - y_1^*)\|_{\mathcal{H}}^2\;.$$

Damit haben wir folgenden Widerspruch erhalten

$$J(y_2^*,\, u^*) \geq J(y_1^*,\, u^*) + m_C^2 \, \|y_2^* - y_1^*\|_{\mathcal{Y}}^2 \ .$$

Somit folgt die Eindeutigkeit von z_N^* aus der von u_N^* .

Gemäß Satz 4 ist die Folge $\overline{\overline{z}}^i$ in $\mathcal{Z}$ beschränkt. Dies bedeutet: es existiert mindestens eine schwach konvergierende Teilfolge $\overline{\overline{z}}^{i_k} \xrightarrow{w} z_2$ in $\mathcal{Z}$. Die Menge $\mathcal{Z}_0 \cap \overline{U_R(0)}$ (s. (2.52)) ist schwach abgeschlossen [97, Kap. 1, §3, Satz 4] . Daraus folgt

$$z_2 \in \mathcal{Z}_0 \cap \overline{U_R(0)} \ , \qquad R = \sqrt{c_y^2 + M_{ad}^2} \ .$$

Da $\overline{\overline{z}}^{i_k}$ ein Optimalpunkt von $\overline{\overline{\mathbf{P}}}_{i_k}^{(T)}$ ist, gilt

$$\overline{\overline{\Psi}}_{i_k}(\overline{\overline{z}}^{i_k}) \leq \overline{\overline{\Psi}}_{i_k}(z_N^*) \ . \tag{2.106}$$

Da z_N^* ein Optimalpunkt von $\mathbf{P}$ ist, gilt

$$J(z_N^*) \leq J(\overline{\overline{z}}^{i_k}) \ . \tag{2.107}$$

Aus (2.106) und (2.107) mit (2.14) folgt

$$\frac{\chi_{i_k}}{2} \ \left\| N\, \overline{\overline{u}}^{i_k} - t_d \right\|_{\mathcal{T}}^2 \leq \frac{\chi_{i_k}}{2} \ \|N\, u_N^* - t_d\|_{\mathcal{T}}^2 \tag{2.108}$$

Da das Funktional $\|N \cdot - t_d\|_{\mathcal{T}}^2$ schwach unterhalbstetig in $\mathcal{U}$ ist [97, Kap. 1, §3, Satz 5] , gilt

$$\|N\, u_2 - t_d\|_{\mathcal{T}}^2 \leq \liminf_{k\to\infty} \left\| N\, \overline{\overline{u}}^{i_k} - t_d \right\|_{\mathcal{T}}^2 \ . \tag{2.109}$$

Nun eliminieren wir in (2.108) die χ's und bilden unter Berücksichtigung von (2.109) den Grenzwert

$$\|N\, u_2 - t_d\|_{\mathcal{T}}^2 \leq \|N\, u_N^* - t_d\|_{\mathcal{T}}^2 \ . \tag{2.110}$$

Weil auch das Funktional $J(\,\cdot\,)$ schwach unterhalbstetig in $\mathcal{Z}$ ist, gilt

$$J(z_2) \leq \liminf_{k\to\infty} J(\overline{\overline{z}}^{i_k}) \qquad \forall\, k \ . \tag{2.111}$$

Falls $\chi_i \to 0$, können wir in (2.106) zum Grenzwert übergehen. Es wird dabei (2.111) berücksichtigt :

$$J(z_2) \leq J(z_N^*) \ . \tag{2.112}$$

Aus (2.112) folgt $z_2 \in \mathcal{Z}^*$; aus (2.110) und (2.103) folgt $z_2 = z_N^*$. Somit erreicht man $\overline{\overline{z}}^{i_k} \xrightarrow{w} z_N^*$ in $\mathcal{Z}$. Als nächstes zeigen wir die starke Konvergenz derselben Teilfolge. Mit (2.6) und mit (2.108) gilt

$$\left\|\overline{\overline{z}}^{i_k} - z_N^*\right\|_{\mathcal{Z}}^2 = \left\|\overline{\overline{y}}^{i_k} - y_N^*\right\|_{\mathcal{Y}}^2 + \left\|\overline{\overline{u}}^{i_k} - u_N^*\right\|_{\mathcal{U}}^2 \leq$$

$$\frac{1}{m_C}\left(\left\|C\left(\overline{\overline{y}}^{i_k} - y_N^*\right)\right\|_{\mathcal{Y}}^2 + \left\|A\left(\overline{\overline{y}}^{i_k} - y_N^*\right)\right\|_{\mathcal{Y}}^2\right) + \left\|\overline{\overline{u}}^{i_k} - u_N^*\right\|_{\mathcal{U}}^2 \leq$$

$$\frac{1}{m_C}\left\|C\left(\overline{\overline{y}}^{i_k} - y_N^*\right)\right\|_{\mathcal{Y}}^2 + \frac{1}{m_N}\left(1 + \frac{M_B}{m_C}\right)\left\|N\left(\overline{\overline{u}}^{i_k} - u_N^*\right)\right\|_{\mathcal{T}}^2 =$$

$$\frac{1}{m_C}\left\|C\left(\overline{\overline{y}}^{i_k} - y_N^*\right)\right\|_{\mathcal{Y}}^2 + \frac{1}{m_N}\left(1 + \frac{M_B}{m_C}\right) \cdot$$

$$\left(2\left\|N\,\overline{\overline{u}}^{i_k} - t_d\right\|_{\mathcal{T}}^2 + 2\left\|N\,u_N^* - t_d\right\|_{\mathcal{T}}^2 - \left\|N\left(\overline{\overline{u}}^{i_k} + u_N^*\right) - 2t_d\right\|_{\mathcal{T}}^2\right) \leq$$

$$\frac{1}{m_C}\left\|C\left(\overline{\overline{y}}^{i_k} - y_N^*\right)\right\|_{\mathcal{Y}}^2 + \tag{2.113}$$

$$\frac{1}{m_N}\left(1 + \frac{M_B}{m_C}\right)\left(4\left\|N\,u_N^* - t_d\right\|_{\mathcal{T}}^2 - \left\|N\left(\overline{\overline{u}}^{i_k} + u_N^*\right) - 2t_d\right\|_{\mathcal{T}}^2\right) .$$

Der erste Term des letzten Ausdrucks in (2.113) konvergiert gegen Null, weil der Operator C als kompakt vorausgesetzt worden ist (s. Abschnitt 2.1). Auch der zweite Term desselben Ausdrucks muß gegen Null tendieren, weil das Funktional $\|N\,(\,\cdot + u_N^*) - 2t_d\|_{\mathcal{T}}^2$ schwach unterhalbstetig ist. Da $\overline{\overline{z}}^{i_k}$ eine beliebige schwach konvergierende Teilfolge von $\overline{\overline{z}}^{i}$ ist, gilt auch

$$\lim_{i\to\infty}\left\|\overline{\overline{z}}^{i} - z_N^*\right\|_{\mathcal{Z}} = 0 . \tag{2.114}$$

□

Mit der Konstanten der starken Konvexität c_κ aus Abschnitt 2.2.2, (2.51) beweisen wir jetzt den Konvergenzsatz für das T-Verfahren.

Satz 7 *Die Folge der Näherungen z^i aus dem T-Verfahren konvergiert stark gegen das Element $z_N^* \in \mathcal{Z}^*$, wenn*

$$\lim_{i\to\infty}\chi_i = \lim_{i\to\infty}\left(\varepsilon_i + \sqrt{\frac{d_1\,r_i}{c_\kappa(\chi_i, r_i)}}\right) = 0 , \tag{2.115}$$

wobei

$$d_1 = M_C^2\,\frac{\hat{c}_q}{m_A}\left(c_y + \frac{c_0}{2\,m_A}\sqrt{\bar{r}}\right) . \tag{2.116}$$

Beweis : Wir benutzen im folgenden das Probeelement $\hat{\bar{z}}^i$ und seine Eigenschaften. Wir wenden (2.75), (2.77) und (2.7), (2.4) an:

$$\begin{aligned}
\bar{\Psi}_i(\bar{\bar{z}}^i) - \bar{\Psi}_i(\bar{z}^i) = \bar{\Psi}_i(\bar{\bar{z}}^i) - \bar{\Psi}_i(\bar{z}^i) - \bar{\Psi}_i(\hat{\bar{z}}^i) + \bar{\Psi}_i(\hat{\bar{z}}^i) &\leq \\
\bar{\bar{\Psi}}_i(\bar{\bar{z}}^i) - \bar{\bar{\Psi}}_i(\hat{\bar{z}}^i) + \bar{\Psi}_i(\hat{\bar{y}}^i,\, \bar{u}^i) - \bar{\Psi}_i(\bar{y}^i,\, \bar{u}^i) &\leq \\
J(\hat{\bar{y}}^i,\, \bar{u}^i) - J(\bar{y}^i,\, \bar{u}^i) - \frac{1}{2\, r_i} \left\| A\, \bar{y}^i - B\, \bar{u}^i - f \right\|_{\mathcal{H}}^2 &\leq \\
\frac{1}{2} \left\| C\, \hat{\bar{y}}^i \right\|_{\mathcal{K}}^2 - \frac{1}{2} \left\| C\, \bar{y}^i \right\|_{\mathcal{K}}^2 = \frac{1}{2} \left(C\, \hat{\bar{y}}^i - C\, \bar{y}^i\,,\, C\, \hat{\bar{y}}^i + C\, \bar{y}^i \right)_{\mathcal{K}} &\leq \\
\frac{1}{2} \left\| C\, \hat{\bar{y}}^i - C\, \bar{y}^i \right\|_{\mathcal{K}} \left\| C\, \hat{\bar{y}}^i + C\, \bar{y}^i \right\|_{\mathcal{K}} &\leq \\
M_C^2 \left\| \hat{\bar{y}}^i - \bar{y}^i \right\|_{\mathcal{Y}} \left\| \bar{y}^i - \frac{r_i}{2}\, G_y (A G_y)^{-1} \bar{q}^i \right\|_{\mathcal{Y}} &\leq \\
M_C^2\, \frac{\hat{\bar{c}}_q}{m_A} \left(c_y + \frac{c_0}{2\, m_A} \sqrt{\bar{r}} \right) r_i &= d_1\, r_i\,. \qquad (2.117)
\end{aligned}$$

Da $\bar{\bar{z}}^i$ ein Optimalelement des Problems $\bar{\mathbf{P}}_i^{(T)}$ ist, gilt für die erste Gâteaux-Ableitung von $\bar{\Psi}_i(z)$

$$\nabla_z \bar{\Psi}_i(\bar{\bar{z}}^i)(\bar{\bar{z}}^i - \bar{\bar{z}}^i) \geq 0\,.$$

Somit haben wir folgende Abschätzung:

$$\bar{\Psi}_i(\bar{\bar{z}}^i) - \bar{\Psi}_i(\bar{\bar{z}}^i) \geq c_t(\chi_i,\, r_i) \left\| \bar{\bar{z}}^i - \bar{\bar{z}}^i \right\|_{\mathcal{Z}}^2\,. \qquad (2.118)$$

Wir fassen jetzt die Ergebnisse (2.117) und (2.118) zusammen und erhalten mit Hilfe der Definition (2.116)

$$\left\| \bar{\bar{z}}^i - \bar{\bar{z}}^i \right\|_{\mathcal{Z}} \leq \sqrt{\frac{d_1\, r_i}{c_\kappa(\chi_i, r_i)}}\,. \qquad (2.119)$$

Mit Lemma 7 und mit den Konvergenzbedingungen (2.115) folgt die starke Konvergenz des T-Verfahrens.

□

Sind Zustandsrestriktionen vorhanden, muß die Konstante d_1 gemäß (2.102) geändert werden:

$$d_1 = M_C^2\, \frac{\hat{\hat{\bar{c}}}_q}{m_A} \left(c_y + \frac{c_0}{2\, m_A} \sqrt{\bar{r}} \right) \max\{\, 2\,,\, (1 + c_w)\}\,. \qquad (2.120)$$

Als Probeelement im Beweis des Satzes 7 tritt dann $\hat{\hat{\bar{z}}}^i \in \mathcal{Z}_0$ aus (2.91) auf.

2.3.2 Verfahren der Prox-Regularisierung

Bevor wir zum Beweis des Konvergenzsatzes für das MSR-Verfahren kommen, zitieren wir ein allgemeines Ergebnis zur Proximal-Punkt-Regularisierung eines konvexen Extremalproblems (vgl. Lemma 13.11 in Kaplan, Tichatschke [40]) . Sei $\mathcal{X}$ ein Hilbertraum mit $\|\cdot\|_{\mathcal{X}}$, ; $\mathcal{X}_1$ sei ein abgeschlossener Unterraum von $\mathcal{X}$ und $P : \mathcal{X} \to \mathcal{X}_1$ sei der Orthoprojektor. Wir betrachten das Problem

$$\text{Minimiere } \Phi(x) = a(x,x) - \ell(x) \quad \text{bei} \quad x \in \mathcal{G}. \tag{2.121}$$

In (2.121) ist $a(\cdot,\cdot)$ eine stetige, symmetrische und nicht negativ definite Bilinearform auf $\mathcal{X} \times \mathcal{X}$, ℓ ist ein stetiges lineares Funktional auf $\mathcal{X}$ und $\mathcal{G} \subset \mathcal{X}$ ist eine konvexe abgeschlossene Menge. Sei $b(\cdot,\cdot)$ eine zweite symmetrische Bilinearform auf $\mathcal{X} \times \mathcal{X}$ mit

$$0 \le b(x,x) \le a(x,x) \qquad \forall\, x \in \mathcal{X} \tag{2.122}$$

und

$$\exists\, \beta > 0 \ : b(x,x) + ||P\,x||^2_{\mathcal{X}} \ge \beta ||x||^2_{\mathcal{X}} \qquad \forall\, x \in \mathcal{X}\ . \tag{2.123}$$

Wir führen in $\mathcal{X}$ eine neue Norm

$$|x|^2_{\mathcal{X}} = b(x,x) + ||P\,x||^2_{\mathcal{X}} \tag{2.124}$$

ein, die zu der alten $||\cdot||_{\mathcal{X}}$ gleichwertig ist:

$$(M+1)||x||^2_{\mathcal{X}} \ge |x|^2_{\mathcal{X}} \ge \beta ||x||^2_{\mathcal{X}}\ ,$$

mit $M \ge \sup\limits_{x \in \mathcal{X}\setminus\{0\}} \dfrac{b(x,x)}{||x||^2_{\mathcal{X}}}$.

Lemma 8 *Für beliebige $x \in \mathcal{G}$, $a^0 \in \mathcal{X}$ und*

$$a^1 = \operatorname{argmin}\ \{\Phi(x) + \frac{\chi}{2}||P\,x - P\,a^0||^2_{\mathcal{X}}\ : x \in \mathcal{G}\} \tag{2.125}$$

($\chi > 0$) gelten die folgenden Ungleichungen:

$$|a^1 - x|^2_{\mathcal{X}} - |a^0 - x|^2_{\mathcal{X}} \le -||P\,a^1 - P\,a^0||^2_{\mathcal{X}} + \frac{2}{\chi}\,[\Phi(x) - \Phi(a^1)] \tag{2.126}$$

und

$$|a^1 - x|_{\mathcal{X}} \ \le\ |a^0 - x|_{\mathcal{X}} + \eta(x), \tag{2.127}$$

mit

$$\eta(x) = \begin{cases} \left(\dfrac{2}{\chi}\,(\Phi(x) - \Phi(a^1))\right)^{1/2} & \textit{wenn}\ \ \Phi(x) > \Phi(a^1) \\ 0 & \textit{sonst} \end{cases}\ .$$

Wenn außerdem $||P\,a^1 - P\,a^0||_{\mathcal{X}} \ge \delta \ge \eta(x)$, so gilt

$$|a^1 - x|_{\mathcal{X}} \le |a^0 - x|_{\mathcal{X}} + \frac{\eta^2(x) - \delta^2}{2|a^0 - x|_{\mathcal{X}}}. \tag{2.128}$$

Der folgende Konvergenzsatz stützt sich auf die Resultate der vorangegangenen Abschnitte. Man muß lediglich die gewonnenen Aussagen auf einfache Substitutionen anwenden und erreicht somit, daß diese auch für das MSR-Verfahren herangezogen werden können. Zum Beispiel:

$$\hat{z}^i := \hat{z}^{i,s},\ \bar{q}^i := \bar{q}^{i,s},\ t_d := N\,u^{i,s-1} \ \text{ usw.}$$

Zur Formulierung des Theorems führen wir noch Schlupfparameter $\vartheta_i > 0$ ein, welche die Länge des inneren Prox-Zyklus des MSR-Verfahrens abstecken und die Größe der Abbruchparameter δ_i beeinflußen.

Satz 8 *Die Folge der Näherungen $z^{i,s}$ aus dem MSR-Verfahren hat immer endlich viele Glieder in s , d. h.*

$$s(i) \le (d_2 + \varepsilon_i)/\vartheta_i \tag{2.129}$$

und konvergiert schwach gegen ein Element aus der Optimalmenge $\mathcal{Z}^*$, wenn

$$0 \le \underline{\chi} < \chi_i \le \bar{\chi}\,, \qquad 0 < r_i \le \bar{r}\,, \qquad 0 < \vartheta_i \le \bar{\vartheta}\,,$$

$$\delta_i \ge \varepsilon_i + \frac{1}{m_N}\sqrt{2\bar{c}^2\,(d_2+\varepsilon_i)\;(\varepsilon_i+\vartheta_i) + \frac{4d_1\,r_i}{\chi_i}}\ , \tag{2.130}$$

$$\sum_{i=1}^{\infty}\left(\bar{c}\,\varepsilon_i + \sqrt{\frac{2d_1\,r_i}{\chi_i}}\ \right) < \infty\ , \tag{2.131}$$

wobei

$$d_1 = M_C^2\,\frac{\hat{c}_q}{m_A}\,\left(c_y + \frac{c_0}{2\,m_A}\,\sqrt{\bar{r}}\right)\ , \tag{2.116}$$

$$d_2 = 2\,\sqrt{c_y^2 + M_{ad}^2}\ , \tag{2.132}$$

$$\bar{c} = \sqrt{\max\left\{M_C^2 + \frac{2\,m_W^2}{\bar{r}}\,M_A^2\,,\ \frac{2\,m_W^2}{\bar{r}}\,M_B^2 + M_N^2\right\}}\ . \tag{2.133}$$

Beweis : Der Beginn dieses Beweises verläuft analog zu dem vom Beweis des Satzes 7 . Statt des Elements $\bar{\bar{z}}^i$ betrachten wir nun aber z^* — ein beliebiges Element der Optimalmenge $\mathcal{Z}^*$ des Originalproblems $\mathbf{P}$ und statt des Funktionals $\bar{\Psi}_i(z)$ untersuchen wir $J_i(z)$ aus (2.15) . Hierbei sei $(\bar{y}^{i,s},\,\bar{u}^{i,s})$ eine Lösung von $\bar{\mathbf{P}}_{i,s}^{(MSR)}$. Dann gilt

$$\begin{aligned} J_i(z^*) - J_i(\bar{z}^{i,s}) = J_i(z^*) - J_i(\bar{z}^{i,s}) - J_i(\hat{z}^{i,s}) + J_i(\hat{z}^{i,s}) \ &\le \\ J(z^*) - J(\hat{z}^{i,s}) + J_i(\hat{y}^{i,s},\,\bar{u}^{i,s}) - J_i(\bar{y}^{i,s},\,\bar{u}^{i,s}) \ &\le \end{aligned}$$

$$J(\hat{\bar{y}}^{i,s},\, \bar{u}^{i,s}) - J(\bar{y}^{i,s},\, \bar{u}^{i,s}) - \frac{1}{2\,r_i}\, \left\| A\,\bar{y}^{i,s} - B\,\bar{u}^{i,s} - f \right\|_{\mathcal{H}}^2 \;\leq$$

$$M_C^2\, \frac{\hat{c}_q}{m_A}\, \left(c_y + \frac{c_0}{2\,m_A}\, \sqrt{\bar{r}} \right)\, r_i \;=\; d_1\, r_i\,. \tag{2.134}$$

Die letzte Ungleichung erhält man ganz analog zu (2.117) . Um das Lemma 8 anwenden zu können, müßen wir noch zwei Abschätzungen vornehmen. Wenn man den Parameter δ_i im MSR-Verfahren hinreichend groß (z.B. wie in (2.130)) wählt, so gilt

$$\left\| N\,(\bar{u}^{i,s} - u^{i,s-1}) \right\|_{\mathcal{T}} \geq m_N\, \left(\left\| u^{i,s} - u^{i,s-1} \right\|_{\mathcal{U}} - \left\| u^{i,s} - \bar{u}^{i,s} \right\|_{\mathcal{U}} \right) \geq$$

$$m_N\, \left(\left\| u^{i,s} - u^{i,s-1} \right\|_{\mathcal{U}} - \left\| z^{i,s} - \bar{z}^{i,s} \right\|_{\mathcal{Z}} \right) \geq m_N\,(\delta_i - \varepsilon_i) \geq \sqrt{\frac{4 d_1\, r_i}{\chi_i}} \;> 0\,. \tag{2.135}$$

Aus (2.134) und (2.135) folgt

$$\left\| N\,(\bar{u}^{i,s} - u^{i,s-1}) \right\|_{\mathcal{T}} \geq \left(\frac{4}{\chi_i}\, \left(\, J_i(z^*) - J_i(\bar{z}^{i,s}) \, \right) \right)^{1/2}\,. \tag{2.136}$$

Wir wenden jetzt das Lemma 8 auf die i-te Iteration des MSR-Verfahrens an. Als Raum $\mathcal{X}$ wählen wir unseren Prozessraum $\mathcal{Z}$ mit einer neuen Norm

$$\|\cdot\|_{\mathcal{X}} := \sqrt{\|\cdot\|_{\mathcal{Y}}^2 + \|N\,\cdot\|_{\mathcal{T}}^2}\,. \tag{2.137}$$

Nach (2.12) sind (2.137) und (A.1) (s. S. 133) gleichwertig. Man führe desweiteren folgende Substitutionen durch:

$$\begin{aligned}
&\mathcal{X}_1 := \{\, z = (y,\, u) \in \mathcal{Z} : y = 0 \;\text{ in } \mathcal{Y} \,\}\,, \quad \Phi(z) := 2\,(J_i(z) - J_i(0))\,,\\
&b(z_1,\, z_2) := (C\,y_1,\, C\,y_2)_{\mathcal{K}} + \frac{m_W^2}{\bar{r}}\, ((A\,y_1 - B\,u_1),\, (A\,y_2 - B\,u_2))_{\mathcal{H}}\,,\\
&a(z_1,\, z_2) := (C\,y_1,\, C\,y_2)_{\mathcal{K}} + \frac{1}{r_i}\, (W(A\,y_1 - B\,u_1),\, W(A\,y_2 - B\,u_2))_{\mathcal{S}}\,,\\
&\ell(z) := a(z,\, z) - \Phi(z)\,, \quad \mathcal{G} := \mathcal{Y}_{ad} \times \mathcal{U}_{ad}\,,\\
&a^0 := z^{i,s-1}\,, \quad a^1 := \bar{z}^{i,s}\,, \quad x := z^*\,, \quad \delta := m_N(\delta_i - \varepsilon_i)\,, \quad \chi := \chi_i\,.
\end{aligned}$$

Es läßt sich auf Grund von (2.124) folgende Normgleichwertigkeit mit den Konstanten aus (2.51) und (2.133) feststellen:

$$c_\kappa \left(1,\, \bar{r}\,(M_W/m_W)^2 \right) \|z\|_{\mathcal{Z}}^2 \leq |z|_{\mathcal{X}}^2 \leq \bar{c}^2\, \|z\|_{\mathcal{Z}}^2\,. \tag{2.138}$$

Nach (2.128) gilt

$$|\bar{z}^{i,s} - z^*|_{\mathcal{X}} \leq |z^{i,s-1} - z^*|_{\mathcal{X}} + \frac{1}{2\bar{c}\,(d_2 + \varepsilon_i)} \left(\frac{2d_1 r_i}{\chi_i} - m_N^2 \,(\delta_i - \varepsilon_i)^2 \right) . \tag{2.139}$$

Unter Berücksichtigung von (2.138) erhält man für $s = 1, .. \, s(i) - 1$

$$\begin{aligned} |z^{i,s} - z^*|_{\mathcal{X}} - |z^{i,s-1} - z^*|_{\mathcal{X}} \quad &\leq \\ \bar{c} \left\| z^{i,s} - \bar{z}^{i,s} \right\|_{\mathcal{Z}} + \frac{1}{2\bar{c}\,(d_2 + \varepsilon_i)} \left(\frac{4d_1 r_i}{\chi_i} - m_N^2 \,(\delta_i - \varepsilon_i)^2 \right) \quad &\leq \\ \bar{c}\,\varepsilon_i + \frac{1}{2\bar{c}\,(d_2 + \varepsilon_i)} \left(\frac{4d_1 r_i}{\chi_i} - m_N^2 \,(\delta_i - \varepsilon_i)^2 \right) \leq -\bar{c}\,\vartheta_i \quad &< \quad 0 \, . \end{aligned} \tag{2.140}$$

Daraus folgt, daß die Werte $|z^{i,s} - z^*|_{\mathcal{X}}$ mit einer absteigenden arithmetischen Reihe (in s) majorisiert werden. Da die zu betrachtenden Größen ϑ_i positiv sind, ist die Anzahl der s-Schritte immer endlich und man erhält dabei die Abschätzung (3.67). Gleichzeitig bekommt man aus (2.140) eine untere Schranke für den Parameter δ_i , d.h. (2.130) . Nach (2.127) ist

$$|\bar{z}^{i,s} - z^*|_{\mathcal{X}} \leq |z^{i,s-1} - z^*|_{\mathcal{X}} + \sqrt{\frac{4d_1 r_i}{\chi_i}} \tag{2.141}$$

oder analog zu (2.140)

$$|z^{i,s} - z^*|_{\mathcal{X}} - |z^{i,s-1} - z^*|_{\mathcal{X}} \leq \bar{c}\,\varepsilon_i + \sqrt{\frac{4d_1 r_i}{\chi_i}} \ . \tag{2.142}$$

Da die Werte $|z^{i,s} - z^*|_{\mathcal{X}}$ nach (2.140) bei $s = 0, .. \, s(i) - 1$ monoton fallend sind, können wir (2.142) über alle $s = 1, ..s(i)$ fortsetzen. Unter Beachtung der Identität $z^{i+1,0} = z^{i,s(i)}$ für alle i , erhält man eine von s unabhängige Abschätzung:

$$|z^{i+1,0} - z^*|_{\mathcal{X}} - |z^{i,0} - z^*|_{\mathcal{X}} \leq \bar{c}\,\varepsilon_i + \sqrt{\frac{4d_1 r_i}{\chi_i}} \ . \tag{2.143}$$

Falls die Beziehung (2.131) gilt, so folgt aus Lemma 2.2.2 in [80] die Konvergenz (in i) von $|z^{i,0} - z^*|_{\mathcal{X}}$ gegen einen positiven endlichen Wert. Die Folge für exakte Lösungen $|\bar{z}^{i,0} - z^*|_{\mathcal{X}}$ konvergiert gegen denselben Wert, denn nach (2.131) gilt $\varepsilon_i \to 0$.

Sei $\{\bar{z}^{i_k,0}\}$ eine beliebige schwach konvergierende Teilfolge von $\{\bar{z}^{i,0}\}$ mit einem schwachen Grenzelement $z_3 \in \mathcal{Z}$. Angesichts (2.78) und anhand der Konvexität von $\mathcal{Z}_0$ (2.52) kann man schließen, daß

$$\rho_{\mathcal{Z}}(\bar{z}^{i+1,0},\, \mathcal{Z}_0) = \rho_{\mathcal{Z}}(\bar{z}^{i,s(i)},\, \mathcal{Z}_0) \leq \left\| \bar{z}^{i,s(i)} - \hat{z}^{i,s(i)} \right\|_{\mathcal{Z}} \leq \frac{\hat{c}_q}{m_A}\, r_i \ , \tag{2.144}$$

wobei $\rho_{\mathcal{Z}}(\cdot,\cdot)$ den Abstand von einem Element zu einer Menge in Hilbertraum $\mathcal{Z}$ bezeichnet. Daraus folgt $z_3 \in \mathcal{Z}_0$. Wir benutzen jetzt die Beziehung (2.126) mit $s := s(i)$

$$\begin{aligned} J(z^*) - J(\bar{z}^{i+1,0}) \geq J_i(z^*) - J_i(\bar{z}^{i,s(i)}) \quad \geq \\ \frac{\chi_i}{4} \left(|\bar{z}^{i,s(i)} - z^*|^2_{\mathcal{X}} - |z^{i,s(i)-1} - z^*|^2_{\mathcal{X}} \right) \quad \geq \\ \frac{\chi_i}{4} \left(|\bar{z}^{i+1,0} - z^*|^2_{\mathcal{X}} - |z^{i,0} - z^*|^2_{\mathcal{X}} \right) \quad . \end{aligned} \tag{2.145}$$

In der Kette (2.145) wurde der Index s diesmal rückwärts fortgesetzt. Da χ_i beschränkt und $J(z)$ schwach unterhalbstetig ist, erhalten wir in (2.145) mit $i_k \to \infty$, daß $J(z^*) \geq J(z_3)$. Daraus folgt $z_3 \in \mathcal{Z}^*$. Die Teilfolge $\{z^{i_k,0}\}$ konvergiert schwach gegen dasselbe Element z_3 , denn nach (2.131) muß $\varepsilon_{i_k} \to 0$. Zuletzt wenden wir auf die Folge $z^{i,0}$ und auf die Menge $\mathcal{Z}^*$ Lemma 1 in [79] an, woraus man die schwache Konvergenz der Iterierten des MSR-Verfahrens in $\mathcal{X}$ gegen das Element z_3 folgert. Nach (2.138) sind die Räume $\mathcal{X}$ und $\mathcal{Z}$ äquivalent.

Wenn Zustandsrestriktionen auftreten, muß die Konstante d_1 im Satz 8 entsprechend (2.120) geändert werden.

□

2.4 Schlußbemerkungen

Im letzten Abschnitt wurden zwei Konvergenzsätze formuliert und bewiesen. Die Abhängigkeit der in diesen Sätzen auftretenden Verfahrenskonstanten von den primären Konstanten konnte exakt hergeleitet werden. Diese Zusammenhänge erlauben es, die Konvergenzbedingungen für T- und MSR-Verfahren "quantitativ" zu erfassen, falls die Auswertung der primären Konstanten im konkreten Fall möglich ist. Für elliptische und parabolische Gleichungen zweiter Ordnung verweisen wir auf klassische Ergebnisse in [50] und [51] , wo die Existenz der notwendigen Lipschitz-Konstanten m_A und M_A untersucht wird. In der Praxis trifft man auch andere Arten von Systemoperatoren: thermo-elastischer Operator [48], Poincaré-Steklov Operator [55], hyperbolische Operatoren erster und zweiter Ordnung [52] u.a. Für einfache Gebiete kann man durch eine Fourier-Analyse die primären Konstanten für entsprechende Hilberträume a priori berechnen.

Die erhaltenen Zusammenhänge ermöglichen eine zusätzliche theoretische Analyse. Man kann jetzt nachvollziehen, wie stark die Werte des Strafparameters r_i geändert werden müßen , falls eine singuläre Störung in der Zustandsgleichung $(M_A/m_A \to \infty)$ vorliegt. Die recht komplizierte Abhängigkeit der Konstanten $c_\kappa(\chi_i, r_i)$ der starken Konvexität des Hilfsfunktionals $\bar{\Psi}_i$ aus (2.10) von ihren Argumenten veranschaulicht es, warum die Verhältnisse zwischen ε_i, χ_i und r_i im

T-Verfahren von der Normierung der Operatoren des Originalproblems abhängen können.

Die Einführung der Operatoren W und N ergibt die Möglichkeit, das regularisierte Srafverfahren auf die Steuerung mit elliptischen Gleichungen in verallgemeinerter Form ($\mathcal{H} = \mathcal{Y}'$) anzupassen, sowie die Fälle mit nicht punktweise beschränkten Steuerrestiktionen zu behandeln. Die Hilberträume $\mathcal{H}$ und $\mathcal{S}$, sowie $\mathcal{U}$ und $\mathcal{T}$, werden dann nicht mehr topologisch äquivalent.

Kapitel 3

Probleme mit unterschiedlich ausgeprägter Inkorrektheit

Unter Steuerspezialisten entsteht immer wieder die Frage: unter welchen Bedingungen zeigt ein schlecht gestelltes Optimalsteuerproblem, in dessen Zielfunktional keine quadratischen Steuerkosten auftreten, keine oder erst eine "unwesentliche" Inkorrektheit und kann somit in der Praxis ohne Regularisierungstechniken behandelt werden? In der Literatur findet man zunächst recht unterschiedliche Auffassungen zu dieser Frage. Offenbar gibt es viele Aufgaben, welche eine Regularisierung brauchen, weil selbst sehr kleine Störungen, die durch verschiedene Fehlerarten hervorgebracht werden, zu einem instabilen Rechenprozess führen [6], [23]. Für etliche Aufgaben braucht man sich aber nicht um die Regularisierung zu kümmern [28], [68], weil die betrachteten Kontrollprobleme durch Steuerrestriktionen eine Korrektheit bekommen, oder auf relativ groben Gittern diskretisiert werden.

Die zweite umstrittene Frage ist mit dem *verallgemeinerten Bang-Bang-Prinzip* verbunden, das für eine Spezialklasse der Kontrollprobleme mit parabolischen Systemen erstmals von F. Tröltzsch [95] formuliert und bewiesen wurde. Spätere Ausarbeitungen findet man auch in [15] und [16]. Grob aufgefasst, behauptet das genannte Prinzip, daß das ganze Gebiet, auf dem die Zustands- und Steuerfunktionen definiert sind, durch exakt drei Teilgebiete zerlegt werden kann: im ersten Teilgebiet findet man notwendigerweise aktive Zustandsrestriktionen, im zweiten müssen die Steuerrestriktionen vom box-constraint-Typ aktiv sein und im dritten fallen die Graphen von Zustands- und Zielelement fast überall zusammen. Selbstverständlich kann das eine oder andere Teilgebiet eine Leermenge darstellen. Unklar ist, ob nicht jedes linear-quadratische Kontrollproblem vom tracking-Typ mit fehlenden Steuerkosten dem verallgemeinerten Bang-Bang-Prinzip genügt?

Die beiden oben gestellten Fragen versuchen wir in diesem Kapitel im Falle elliptischer Zustandsgleichungen aufzuklären. Dazu müssen wir auf die Theorie der Sobolevräume und elliptischen Operatoren näher eingehen. Diesem Ziel wird Abschnitt 3.1 gewidmet. In Abschnitten 3.2 und 3.3 untersuchen wir zwei konkrete Optimalsteuerprobleme mit elliptischen Gleichungen 2-ter bzw. 1-ter Ordnung, wel-

che für die von uns diskutierten Fragen gewissermaßen typisch und charakteristisch sind. Es wird das bekannte Fuller-Problem und ein Steuerproblem mit dem Poincaré-Operator erster Ordnung betrachtet. Wir stellen fest, daß das erste Problem eine stark ausgeprägte Inkorrektheit besitzt und bei numerischen Berechnungen eine Regularisierung erfordert. Im Gegensatz dazu weist das zweite sehr ähnlich gebaute Steurproblem keine sichtbaren Merkmale einer Inkorrektheit auf: der Rechenprozess ohne Regularisierung bleibt bis zu sehr feinen Gittern stabil.

Das Fuller-Problem genügt dem verallgemeinerten Bang-Bang-Prinzip, weil die optimale Zustandsfunktion in einem vorliegenden singulären Bereich nullidentisch ist, wohingegen das zweite Steuerproblem mit elliptischer Gleichung erster Ordnung kein Bang-Bang-Prinzip erfüllt.

Am Beispiel des Fuller-Problems beweisen wir den Konvergenzsatz für das prox-regularisierte Strafverfahren, in dem wir nun einen Diskretisierungsfehler berücksichtigen und mit ihm informativere Konvergenzbedingungen formulieren. Wir erheben dabei keinen Anspruch auf die Allgemeinheit der in Abschnitt 3.2.2 erhaltenen Resultate. Man kann nämlich sowohl die Problemstellung ausdehnen, als auch die gewonnene Asymtotik (3.74) verbessern. (Man könnte mit

$$\sum_{i=1}^{\infty} \frac{\sqrt{h_i}}{\chi_i} < \infty \, .$$

davonkommen, was schon eine scharfe Abschätzung darstellen mag). Diese Abschätzung werden wir in Kapitel 4 aber weiter verbessern, indem wir auf die Verwendung von Straftechniken völlig verzichten.

3.1 Elliptische Operatoren in Sobolevräumen

3.1.1 Sobolevräume

Sei Ω ein offenes, beschränktes und zusammenhängendes Gebiet im $\mathbf{R}^n$. Die Menge $\partial\Omega = \overline{\Omega} \setminus \Omega$, welche den Rand des betrachteten Gebiets darstellt, sei eine unendlich differenzierbare Mannigfaltigkeit der Dimension $n-1$; dabei liege lokal in der Umgebung jedes Randpunkts das Gebiet Ω auf einer Seite von $\partial\Omega$. Wir bezeichnen mit $L_2(\Omega)$ den Raum der Funktionenklassen, deren Betrag zum Quadrat in Ω summierbar ist, d. h. der meßbaren Funktionen, für welche

$$\|u\|_\Omega = \left(\int\limits_\Omega |u|^2 \, dx \right)^{1/2} < \infty$$

gilt, wobei das Integral im Sinne von Lebesgue [18, Kap. 12.9] zu verstehen ist. Diese Funktionenmenge wird zu einem Hilbertraum, wenn man die stetige bilineare Form

$$(u,\, v)_\Omega = \int\limits_\Omega u\, \overline{v} \, dx$$

als Skalarprodukt über den Elementen aus $L_2(\Omega)$ erklärt [1], [10], [44].

Wir betrachten nun die Menge $\mathring{C}^\infty(\Omega)$ der Funktionen, welche im klassischen Sinne unendlich oft differenzierbar sind und einen kompakten Träger [100, S. 16] in Ω besitzen, d. h. supp $u \subset \Omega \;\; \forall\, u \in \mathring{C}^\infty(\Omega)$. Auf $\mathring{C}^\infty(\Omega)$ führen wir ein System von Halbnormen

$$p_{j,\mathrm{K}}(u) = \max_{|\alpha|\le j} \sup_{x\in \mathrm{K}} |D^\alpha u(x)| \;, \qquad j \in \mathbf{Z}_+ \;, \quad \overline{\mathrm{K}} \subset \Omega \;,$$

ein, wobei

$$D^\alpha = \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1} ... \partial x_n^{\alpha_n}} \;, \quad \alpha = (\alpha_1, ..., \alpha_n) \in \mathbf{Z}_+^n \;, \quad |\alpha| = \alpha_1 + \cdots + \alpha_n \;,$$

mit dessen Hilfe die angegebene Menge der Funkionen in einen linearen topologischen Raum umgewandelt werden kann. Jede offene Untermenge dieses topologischen Raumes kann als Vereinigung einer beliebigen Anzahl von Teilmengen

$$U(u, \varepsilon, \mathrm{K}, j) = \Big\{ v \in \mathring{C}^\infty(\Omega) : \; p_{i,\mathrm{K}}(v-u) < \varepsilon, \; i = 1, .., \; j \Big\} \;, \tag{3.1}$$

dargestellt werden, wobei u ein beliebiges Element aus $\mathring{C}^\infty(\Omega)$ und ε eine beliebige positive Zahl ist. Nach Definition [97, S. 60] bedeutet dann die Konvergenz einer Folge $\{u_k\} \in \mathring{C}^\infty(\Omega)$ gegen ein Element $u \in \mathring{C}^\infty(\Omega)$, daß es für jede Umgebung U der Form (3.1) eine natürliche Zahl $N = N(U)$ gibt, so daß $u_k \in U$, falls $k \ge N$.

Der Raum kann mit einem Teilraum in $L_2(\Omega)$ identifiziert werden, wenn man jeder Funktion $u(x) \in \mathring{C}^\infty(\Omega)$ ihre Äquivalenzklasse $u \in L_2(\Omega)$ zuordnet. Die Zuordnung erfolgt korrekt, da sich nur die Nullfunktion von $\mathring{C}^\infty(\Omega)$ in der Nullklasse von $L_2(\Omega)$ befindet. Der so definierte Teilraum besitzt außerdem eine sehr wichtige Eigenschaft bezüglich des ganzen Raumes: die Funktionen aus $\mathring{C}^\infty(\Omega)$ liegen in $L_2(\Omega)$ dicht [100].

Jetzt betrachten wir die Menge $C^{-\infty}(\Omega)$ aller linearen Funktionale über dem Raum $\mathring{C}^\infty(\Omega)$. Diese Menge wird zu einem linearen topologischen Raum, wenn sie mit einer schwachen Topologie [44], [97] ausgestattet wird. Das heißt, jede offene Untermenge in $C^{-\infty}(\Omega)$ kann als die Vereinigung einer beliebigen Anzahl von Mengen

$$U(f, \varepsilon, u_1, ..., u_m) = \Big\{ g \in C^{-\infty}(\Omega) : \; |\langle g, \, u_i\rangle - \langle f, \, u_i\rangle| < \varepsilon, \; i = 1, .., \, m \Big\} \;,$$

dargestellt werden, wobei f ein beliebiges Element aus $C^{-\infty}(\Omega)$, ε eine beliebige positive Zahl und $u_1, ..., u_m$ beliebige Elemente aus $\mathring{C}^\infty(\Omega)$ sind. Hierbei bezeichnen die Dualitätsklammern wie im Kapitel 2 die Wirkung des linearen Funktionals. Die Funktionale aus dem erhaltenen dualen Raum $C^{-\infty}(\Omega)$, welche wir als *verallgemeinerte Funktionen in* Ω bezeichnen, haben einige Eigenschaften der Elemente des Raumes $L_2(\Omega)$. So können die Werte einer verallgemeinerten Funktion in einzelnen Punkten des Gebiets auch nicht bestimmt werden, man kann aber beurteilen,

ob zwei verallgemeinerte Funktionen in jedem beliebig kleinen offenen Teilgebiet $\Omega' \subset \Omega$ gleichwertig sind. Die verallgemeinerte Funktion f heißt nämlich *nulliden-tisch in* Ω' , wenn $\langle f,\, u\rangle = 0\,,\ \ \forall\, u \in \{v \in \mathring{C}^{\infty}(\Omega) :\ \text{supp}\ v \subset \Omega'\}$. Auf Grund der beschriebenen Eigenschaft werden die verallgemeinerten Funktionen in der Literatur [29], [59], [100] häufig als *Distributionen in* Ω genannt.[1]

Der letzte vorbereitende Schritt besteht in der Identifizierung des Raumes $L_2(\Omega)$ mit einem Teilraum von $C^{-\infty}(\Omega)$. Dafür ordnen wir jedem Element $u \in L_2(\Omega)$ ein lineares Funktional $f \in C^{-\infty}(\Omega)$ nach folgender Regel zu:

$$\langle f,\, v\rangle = (u,\, v)_{\Omega} \qquad \forall\, v \in \mathring{C}^{\infty}(\Omega)\ .$$

Die Zuordnung erfolgt korrekt, da aus $f = 0$ in $C^{-\infty}(\Omega)$ die Bedingung $u = 0$ in $L_2(\Omega)$ folgt. Dies läßt sich unmittelbar aus der erwähnten Eigenschaft der Dichtheit $\mathring{C}^{\infty}(\Omega)$ in $L_2(\Omega)$ ersehen. Somit kommen wir nun zu einer wichtigen Einbettungs-kette

$$\mathring{C}^{\infty}(\Omega) \subset L_2(\Omega) \subset C^{-\infty}(\Omega)\ , \tag{3.2}$$

wobei die Inklusionen mengenmäßig und topologisch verstanden werden. Die Funk-tionen aus $L_2(\Omega)$ liegen in $C^{-\infty}(\Omega)$ dicht [100].

Von dieser Stelle an werden die meisten Funktionen als Distributionen aus $C^{-\infty}(\Omega)$ betrachtet. Trotzdem haben die Funktionen verschiedene Regularität, je nach dem, welchem Teilraum der Kette (3.2) sie angehören. Zum Beispiel für die Distributio-nen aus $L_2(\Omega)$ können die Operationen der Multiplikation und unter Umständen auch der Division eingeführt werden. Für die Distributionen aus $\mathring{C}^{\infty}(\Omega)$ kann man außerdem die Werte (die Spuren) in einzelnen Punkten definieren. Die wei-tere Zielsetzung ist nun eine ausführlichere Unterteilung der Enbettungskette (3.2) durch neue Hilberträume, welche größer oder kleiner als $L_2(\Omega)$ sind und somit eine vernünftige Klassifikation der existierenden Distributionen ermöglichen. Die entste-henden Hilberträume werden von uns desweiteren unter dem Gesichtspunkt der Ana-lyse der Lösungen partieller Differentialgleichungen und der Konvergenz zugehöriger numerischer Verfahren betrachtet. Eine vernünftige Klassifikation ergibt sich des-halb durch Auswahl der Distributionen nach ihrer Differenzierbarkeit.

Der oben eingeführte Differentialoperator D^{α} überführt offensichtlich den linea-ren Teilraum $\mathring{C}^{\infty}(\Omega)$ immer in sich selbst, d. h. für den Graph des Operators gilt $D^{\alpha} \in \mathcal{L}\left(\mathring{C}^{\infty}(\Omega),\, \mathring{C}^{\infty}(\Omega)\right)$. Man kann diesen Graph im Raum $\mathcal{L}\left(C^{-\infty}(\Omega),\, C^{-\infty}(\Omega)\right)$ stetig fortsetzen, indem man den *Operator der verallgemeinerten Ableitung* D^{α} wie folgt definiert:

$$\langle D^{\alpha}\, f,\, u\rangle = (-1)^{|\alpha|}\ \langle f,\, D^{\alpha}\, u\rangle \qquad \forall\, u \in \mathring{C}^{\infty}(\Omega)\ .$$

Eine derartige Fortsetzung erfolgt korrekt für alle Gebiete, in denen die Regel der partiellen Integration unendlich glatter Funktionen in Kraft bleibt. Im Weiteren

[1] Statt $\mathring{C}^{\infty}(\Omega)$ und $C^{-\infty}(\Omega)$ schreibt man oft $\mathcal{D}(\Omega)$ und $\mathcal{D}'(\Omega)$.

wird der Operator D^α als verallgemeinerte Ableitung der $|\alpha|$-ten Ordnung betrachtet. Offenbar ist nun der Graph von D^α in $\mathcal{L}\,(L_2(\Omega),\, L_2(\Omega))$ nicht mehr beschränkt. Wir definieren

$$\mathcal{D}(D^\alpha) = \{u \in L_2(\Omega) :\; D^\alpha\, u \in L_2(\Omega)\}\ ,$$

als Definitionsbereich des linearen Operators und vermerken [60, Kap 7, §6], daß D^α mit einem solchen Definitionsbereich jetzt einen abgeschlossenen Graph hat. Diese Tatsache erlaubt es, ein System von linearen Teilräumen[2]

$$W^m(\Omega) = \bigcap_{0<|\alpha|\le m} \mathcal{D}(D^\alpha)\ , \qquad m \in \mathbf{N}\ , \tag{3.3}$$

in $L_2(\Omega)$ einzuführen, welche mit der Graphnorm [25], [59], [83], [94]

$$\|u\|_{W^m(\Omega)} = \left(\|u\|^2_\Omega + \sum_{0<|\alpha|\le m} \|D^\alpha\, u\|^2_\Omega \right)^{1/2}\ , \tag{3.4}$$

zu selbständigen Hilberträumen werden.[3] Die Räume $W^m(\Omega)$ heißen *Sobolevräume m-ter Ordnung in* Ω [59], [100]. Es gelten offenbar folgende strenge Einbettungen [98, S. 149 (3)]:

$$\mathring{C}^\infty(\Omega) \subset \cdots \subset W^m(\Omega) \subset \cdots \subset W^0(\Omega) = L_2(\Omega)\ , \tag{3.5}$$

welche mengenmäßig und topologisch zu verstehen sind. Die erhaltene Klassifikation der Distributionenräume (3.5) kann verfeinert werden, wenn man die vorangegangenen Überlegungen von Anfang an mit den weiter unten definierten Räumen $W^{0,\lambda}(\Omega)$ statt $W^0(\Omega)$ wiederholt. Der Hilbertraum $W^{0,\lambda}(\Omega)$, $\lambda \in [0,\, 1)$ besteht nämlich aus meßbaren Funktionen, für welche

$$\|u\|_{W^{0,\lambda}(\Omega)} = \left(\|u\|^2_\Omega + \iint\limits_{\Omega\times\Omega} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2\lambda}}\, dxdy \right)^{1/2} < \infty$$

[2]Da wir im Rahmen dieser Dissertation nur mit Hilberträumen zu tun haben, erlauben wir uns die Abkürzung $W^m(\Omega) = W^m_2(\Omega)$.

[3]Hier und desweiteren ergibt sich das Skalarprodukt für zwei Elemente x_1 und x_2 eines Hilbertraumes $\mathcal{X}$ eindeutig aus der Norm nach der Regel [9, Kap. 1] oder [99, § 9]

$$(x_1,\, x_2)_{\mathcal{X}} = (x_1,\, x_2)_* + i\,(x_1,\, i\,x_2)_*\ ,$$

wobei

$$(x_1,\, x_2)_* = \frac{1}{4}\left(\|x_1+x_2\|^2_{\mathcal{X}} - \|x_1-x_2\|^2_{\mathcal{X}} \right)\ .$$

gilt. Schließlich erhalten wir

$$W^{m,\lambda}(\Omega) = \left\{ u \in W^{0,\lambda}(\Omega) : \; D^\alpha u \in W^{0,\lambda}(\Omega) \;\; \forall\, |\alpha| \le m \right\}, \quad m \in \mathbf{N} \tag{3.6}$$

mit

$$\|u\|_{W^{m,\lambda}(\Omega)} = \left(\|u\|^2_{W^{0,\lambda}(\Omega)} + \sum_{0<|\alpha|\le m} \|D^\alpha u\|^2_{W^{0,\lambda}(\Omega)} \right)^{1/2} . \tag{3.7}$$

Die in (3.6), (3.7) eingeführten Räume ergeben für beliebiges $m \in \mathbf{Z}_+$ die Eigenschaft [87], [100]

$$W^{m+1}(\Omega) \subset \cdots \subset W^{m,\lambda}(\Omega) \subset \cdots \subset W^{m,0}(\Omega) \subset W^m(\Omega)$$

und heißen *Sobolev-Slobodeckij-Räume in* Ω. Man kann nachweisen [100], daß alle bisher betrachteten Inklusionen (außer $\mathring{C}^\infty(\Omega)$ in (3.5) !) die Dichtheit der linearen Teilräume in den größeren Räumen voraussetzen. Jetzt ist es recht offensichtlich, daß mit $m \ge |\alpha|$, $\lambda \in [0,\, 1)$ und $\phi(x) \in C^\infty(\Omega)$

$$\phi(x)\, D^\alpha \in \mathcal{L}(W^m(\Omega),\, W^{m-|\alpha|}(\Omega)) , \quad \phi(x)\, D^\alpha \in \mathcal{L}(W^{m,\lambda}(\Omega),\, W^{m-|\alpha|,\lambda}(\Omega)) .$$

Ähnliche Überlegungen kann man im Falle einer beschränkten, unendlich glatten, meßbaren Mannigfaltigkeit $\mathrm{M} \subset \mathbf{R}^n$ an Stelle von Ω entwickeln [29], [100]. Sei $\Gamma \subseteq \partial\Omega$ eine Mannigfaltigkeit solcher Art (nicht unbedingt zusammenhängend, aber mit offenem Rand, falls $\Gamma \subset \partial\Omega$). Wir bezeichnen die entsprechenden Hilberträume mit $\mathring{C}^\infty(\Gamma)$, $W^m(\Gamma)$, $W^{m,\lambda}(\Gamma)$ usw. Es mögen für diese Räume alle bisher abgeleiteten Eigenschaften gelten. Wir führen jetzt auf den Sobolevräumen in Ω einen neuen linearen Operator für die Distributionen ein: den *Spuroperator* $\gamma_0 \in \mathcal{L}(W^m(\Omega),\, W^{m-1,1/2}(\Gamma))$, $m \in \mathbf{N}$, dessen Graph sich durch die stetige Fortsetzung des klassischen Operators der Beschränkung auf eine Teilmenge $\gamma_0 \in \mathcal{L}(C^\infty(\Omega),\, C^\infty(\Gamma))$ (vgl. die Einführung der verallgemeinerten Ableitung) ergibt. Zur richtigen Motivation einer solchen Fortsetzung verweisen wir wieder auf [29], [100].

Wir führen *die Funktion der äußeren Normale* $\nu(\tilde{x}) = (\nu_1(\tilde{x}), \dots, \nu_n(\tilde{x}))^{\mathrm{T}}$, $\tilde{x} \in \partial\Omega$ für den Rand des Gebietes ein. Im Falle unendlich glatter Gebiete kann man leicht feststellen, daß $\nu_1(\tilde{x}), \dots, \nu_n(\tilde{x}) \in C^\infty(\partial\Omega)$. Als *Spuroperator m-ter Ordnung* γ_m für $m \ge 0$ gilt folgende Komposition der bereits bekannten Operatoren [5, §3-5]:

$$\gamma_m = \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \nu_{i_1}(\tilde{x}) \cdots \nu_{i_m}(\tilde{x})\, \gamma_0\, D_{i_1} \cdots D_{i_m} , \tag{3.8}$$

wobei D_i die erste verallgemeinerte Ableitung in der i-ten Koordinate bezeichnet. Im Fall hinreichend glatter Funktionen stimmt (3.8) mit der m-ten klassischen Ableitung entlang der äußeren Normalen überein. Es ist klar, daß $\gamma_j \in$

$\mathcal{L}(W^m(\Omega),\, W^{m-1-j,1/2}(\Gamma))$, $m \in \mathbf{N}$, $j = 0, .., m-1$. Die Nullmengen des Spuroperators

$$\mathcal{N}(\gamma_j) = \left\{ u \in W^m(\Omega) :\ \gamma_j\, u = 0 \ \text{ in } \ W^{m-1-j,1/2}(\Gamma) \right\}, \ \ j = 0, .., m-1$$

sind lineare Teilräume in $W^m(\Omega)$. Es werden neue Sobolevräume

$$W^m_\Gamma(\Omega) = \bigcap_{j=0}^{m-1} \mathcal{N}(\gamma_j)\,, \qquad m \in \mathbf{N}\,, \tag{3.9}$$

mit derselben Norm (3.4) eingeführt. Falls $\Gamma = \partial\Omega$, schreiben wir in (3.9) traditionsgemäß $\mathring{W}^m(\Omega)$ [24], [31], [57]. Falls $\Gamma \subset \partial\Omega$, mes $\Gamma > 0$, dann gilt

$$\mathring{W}^m(\Omega) \subset W^m_\Gamma(\Omega) \subset W^m(\Omega)\,. \tag{3.10}$$

Die Inklusionen in (3.10) sind keine *dichten Einbettungen* wie die meisten zuvor, sondern sogenannte *abgeschlossene Einbettungen*, weil die Teilräume (3.9) in $W^m(\Omega)$ abgeschlossen sind [29], [31]. Wichtig ist aber, daß die Inklusionskette

$$\mathring{C}^\infty(\Omega) \subset \cdots \subset \mathring{W}^{m+1}(\Omega) \subset \mathring{W}^m(\Omega) \subset \cdots \subset L_2(\Omega) \tag{3.11}$$

nur die dichten strengen Einbettungen enthält [31], [100]. Dies bedeutet gleichzeitig, daß die Sobolevräume $\mathring{W}^m(\Omega)$ mittels Vervollständigung von $\mathring{C}^\infty(\Omega)$ in entsprechenden Normen (3.4) definiert werden können [29], [98].

Wir betrachten nun *duale Räume* zu den Hilberträumen in der Kette (3.11).[4] Entsprechend des Darstellungssatzes von Riesz können wir den Raum $L_2(\Omega)$ als *pivot space* [5] erklären, d. h. er wird mit seinem dualen Raum mengenmäßig und topologisch identifiziert. Die anderen dualen Räume bezeichnen wir traditionsgemäß als Sobolevräume mit negativen Indizes $W^{-m}(\Omega) = \left(\mathring{W}^m(\Omega)\right)'$, $m \in \mathbf{N}$. Nach Corollary 1-2 in [5, Kap 2] gelten folgende dichte Einbettungen

$$L_2(\Omega) \subset \cdots \subset W^{-m}(\Omega) \subset W^{-m-1}(\Omega) \subset \cdots \subset C^{-\infty}(\Omega)\,,$$

d. h. die Sobolevräume mit negativen Indizes werden mit bestimmten Teilräumen der Distributionen in Ω identifiziert. Die letzte Einbettungskette ergibt die gewünschte Klassifikation verallgemeinerter Funktionen "außerhalb" des Raumes $L_2(\Omega)$.

Man kann auch die dualen Räume $(W^m_\Gamma(\Omega))'$, $(W^m(\Omega))'$ und $\left(W^{m,\lambda}(\Omega)\right)'$ mit $m \in \mathbf{N}$, $\lambda \in [0,\, 1)$, $\Gamma \subset \partial\Omega$, mes $\Gamma > 0$ in Betracht ziehen. Sie sind aber im allgemeinen in $C^{-\infty}(\Omega)$ nicht enthalten.

[4] Ein dualer Raum $\mathcal{X}'$ enthält entsprechende lineare Funktionale über dem Hilbertraum $\mathcal{X}$ und ist mit einer *Dualnorm* [31, § 6.3] ausgestattet:

$$\|f\|_{\mathcal{X}'} = \sup_{x_0 \in \mathcal{X}\setminus\{0\}} \frac{\langle\, f,\, x_0 \,\rangle_{\mathcal{X}',\,\mathcal{X}}}{\|\, x_0 \,\|_{\mathcal{X}}} \qquad \forall\, f \in \mathcal{X}'\,.$$

3.1.2 Elliptische Operatoren s-ter Ordnung

Wir beginnen die Diskussion der linearen Operatoren mit einer Betrachtung unterschiedlicher *Einbettungsoperatoren* G und *Operatoren des kanonischen Isomophismus* Λ für die in Abschnitt 3.1.1 eingeführten Ketten von Sobolevräumen.

Sei $G_m \in \mathcal{L}(\mathring{W}^{m+1}(\Omega),\, \mathring{W}^{m}(\Omega))$ ein Einbettungsoperator, der den Raum $\mathring{W}^{m+1}(\Omega)$ in den entsprechenden linearen Teilraum von $\mathring{W}^{m}(\Omega)$ auf natürliche Weise überführt, dann spielt der adjungierte Operator

$$G_m^* \in \mathcal{L}(W^{-m}(\Omega),\, W^{-m-1}(\Omega))$$

die Rolle eines Einbettungsoperators, der den Raum $W^{-m}(\Omega)$ in den ihm entsprechenden linearen Unterraum von $W^{-m-1}(\Omega)$ abbildet. Es gilt dabei: (i) $\mathcal{N}(G_m^*) = \{0\}$, (ii) $\mathcal{R}(G_m^*)$ liegt in $W^{-m-1}(\Omega)$ dicht, d. h. G_m^* verwirklicht eine *strenge dichte Einbettung* [5, Kap 2, § 1-4].

Sei $G \in \mathcal{L}(\mathring{W}^{m}(\Omega),\, W^{m}(\Omega))$ ein Einbettungsoperator, der den Raum $\mathring{W}^{m}(\Omega)$ in den entsprechenden linearen Teilraum von $W^{m}(\Omega)$ auf natürliche Weise (3.10) überführt, dann spielt der adjungierte Operator

$$G^* \in \mathcal{L}((W^{m}(\Omega))',\, W^{-m}(\Omega))$$

die Rolle eines Einbettungsoperators, der einen gewissen Teil des Raumes $(W^{m}(\Omega))'$ in den ihm entsprechenden linearen Unterraum von $W^{-m}(\Omega)$ abbildet. Es gilt: (i) $\mathcal{N}(G^*)$ besteht aus denjenigen Funktionalen, welche zum abgeschlossenen linearen Teilraum $\mathring{W}^{m}(\Omega)$ orthogonal sind, (ii) $\mathcal{R}(G^*)$ ist in $W^{-m}(\Omega)$ abgeschlossen, d. h. G^* verwirklicht eine *nichtinjektive abgeschlossene Einbettung* [5, Kap 2, § 1-7].

Wir geben später eine explizite Form für die Beschreibung der Funktionale aus $\mathcal{N}(G^*)$ über Delta-Schichten unterschiedlicher Ordnung an.

Die im Abschnitt 3.1.1 diskutierten Sobolevräume

$$\mathring{W}^{m}(\Omega) \subset L_2(\Omega) \subset W^{-m}(\Omega)\,, \quad W^{m}(\Omega) \subset L_2(\Omega) \subset (W^{m}(\Omega))' \quad \text{u. a.}$$

mit $m \in \mathbf{N}$ bilden sogenannte *Gelfand-Dreier* [31], [100]. Aus dem Darstellungssatz von Riesz folgt unmittelbar die Existenz eines Operators des kanonischen Isomorphismus $\Lambda_m \in \mathcal{L}(\mathring{W}^{m}(\Omega),\, W^{-m}(\Omega))$ [31] mit den Identitäten

$$\begin{aligned} \|\Lambda_m\, u\|_{W^{-m}(\Omega)} &= \|u\|_{W^{m}(\Omega)} \qquad \forall\, u \in \mathring{W}^{m}(\Omega)\,, \\ \|\Lambda_m^{-1}\, f\|_{W^{m}(\Omega)} &= \|f\|_{W^{-m}(\Omega)} \qquad \forall\, f \in W^{-m}(\Omega)\,. \end{aligned}$$

Ähnliche Operatoren existieren ebenfalls für beliebige Gelfand-Dreier der bereits diskutierten Hilberträume, auch dann, wenn eine glatte meßbare Mannigfaltigkeit M an Stelle des Gebiets Ω betrachtet wird.

Beispiel: Durch eine partielle Integration oder Umgruppierung der Integralle kann man leicht explizite Formen der Λ-Operatoren für die Sobolevräume $\mathring{W}^{1}(-1,\,1)$, $\mathring{W}^{2}(-1,\,1)$ und $W^{0,1/2}(-1,\,1)$ erhalten:

- $\Lambda_1 u(x) = -D_x^2 u(x) + u(x)$;
- $\Lambda_2 u(x) = D_x^4 u(x) - D_x^2 u(x) + u(x)$;
- $\Lambda_{1/2} u(x) = 2 \oint_{-1}^{1} \frac{u(x) - u(\xi)}{|x-\xi|^2} \, d\xi + u(x)$.[5]

Die Wirkung der angegebenen Operatoren läßt sich einfach erklären, wenn man einen dichten Definitionsbereich $\mathcal{D}(\Lambda)$ finden kann, welcher aus hinreichend regulären Funktionen besteht. Dann wählt man eine gegen $u(x)$ konvergierende Folge $u_n(x)$ aus dieser Menge und untersucht das Verhalten der Funktionen $\Lambda\, u_n(x)$, welche immer gegen eine entsprechende Distribution $\Lambda\, u(x)$ konvergieren.

□

Es gelingt mit Hilfe der Operatoren des kanonischen Isomorphismus die Interpolation der Sobolevräume mit ganzen Indizes auf die zwischen einem primalen und einem dualen Raum liegenden neuen Hilberträume durchzuführen [59]. Die Interpolationstechnik stützt sich auf die Existenz nichtganzer Potenzen der entsprechenden Λ-Operatoren. Wir identifizieren die Räume $\mathring{H}^m(\Omega)$, $H^m(\Omega)$ und $H^m_\Gamma(\Omega)$ traditionell mit $\mathring{W}^m(\Omega)$, $W^m(\Omega)$ bzw. $W^m_\Gamma(\Omega)$, solange m ganzzahlig ist. Die Einbettung $H^{m_1}(\Omega) \hookrightarrow H^{m_2}(\Omega)$ mit $m_1 > m_2$ ist *kompakt*, wenn Ω beschränkt ist [59]. Es kann dabei eine geringere Glattheit des Randes [31, Satz 6.4.8] oder nur die Beschränktheit des Gebietsvolumens [70, Satz 8.2] vorausgesetzt werden. Somit ist der eingeengte inverse Operator $\Lambda_m^{-1} \in \mathcal{L}(L_2(\Omega))$ immer *vollstetig* und *selbstadjungiert*. Er besitzt folglich ein diskretes Spektrum [60, Kap 6, § 4, Satz 5]. Es gibt ein abzählbares System von Eigenwerten und Eigenfunktionen $\{(\lambda_k,\, \phi_k(x))\}_{k=1}^{\infty}$, mit der Eigenschaft

$$\Lambda_m \, \phi_k(x) = \lambda_k \, \phi_k(x) \, , \qquad k = 1, 2, .. \, .$$

Wir nehmen an, daß jede Eigenfunktion in $L_2(\Omega)$ normiert ist, und definieren die *s-te Potenz des Operators* Λ_m $(-1 \le s \le 1)$ als unendliche Summe

$$\Lambda_m^s \, u \;\stackrel{\text{def}}{=}\; \sum_{k=1}^{\infty} (\lambda_k)^s \, (\phi_k, u)_\Omega \, \phi_k \qquad \forall \, u \in \mathring{H}^m(\Omega) \, .$$

Die Zwischenräume $\mathring{H}^s(\Omega) (0 < s \le m)$ werden als Definitionsbereich des Operators $\Lambda_m^{s/2m}$ erklärt und mit der Graphnorm

$$\|u\|_{\mathring{H}^s(\Omega)} = \left(\|u\|_\Omega^2 + \|\Lambda_m^{s/2m} \, u\|_\Omega^2 \right)^{1/2} \qquad \forall \, u \in \mathring{H}^s(\Omega) \tag{3.12}$$

ausgestattet. Falls s der Menge $\left\{ \frac{1}{2}, \frac{3}{2}, .. \, (m - \frac{1}{2}) \right\}$ angehört, schreiben wir traditionsgemäß $H^s_{00}(\Omega)$ [59, Kap 1, § 11.5]. Der Wert m muß jedesmal entweder im Kontext angegeben oder aus der Beziehung $s \le m < s + 1$ [1, § 7.66] entnommen werden.

[5] Cauchy-Hauptwert eines singulären Integrals ist z.B. in [32, Kap. 7] definiert.

Analog lassen sich die Zwischenräume $H^s(\Omega)$ und $H^s_\Gamma(\Omega)(0 < s \leq m)$ definieren. Man muß dann offenbar andere Spektralprobleme für entsprechende Operatoren des kanonischen Isomorphismus lösen. Die zugehörigen Räume der negativen Indizes $\mathring{H}^{-s}(\Omega)$, $H^{-s}(\Omega)$ und $H^{-s}_\Gamma(\Omega)(0 < s \leq m)$ werden als duale Hilberträume verstanden.

Die in diesem Abschnitt eingeführten Funktionenräume heißen *Sobolev-Lions-Räume in* Ω [57], [59]. Sie bilden lokale Skalen mit dichten kompakten Einbettungen [59, Kap 1, § 16]. Beispielweise bei $m > s_1 > s_2 > 0$ gilt:

$$\mathring{H}^m(\Omega) \hookrightarrow \mathring{H}^{s_1}(\Omega) \hookrightarrow \mathring{H}^{s_2}(\Omega) \hookrightarrow L_2(\Omega) \hookrightarrow \mathring{H}^{-s_2}(\Omega) \hookrightarrow \mathring{H}^{-s_1}(\Omega) \hookrightarrow \mathring{H}^{-m}(\Omega) \; .$$

Die Sobolev-Lions-Räume auf einer glatten meßbaren Mannigfaltigkeit M lassen sich manchmal mit einer ähnlichen Interpolation der Sobolevräume $W^m(\mathrm{M})$ (bzw. $\mathring{W}^m(\mathrm{M})$) ganzzahliger Ordnung m einführen. (Weitere Hinweise auf die Techniken nichtganzzahliger Potenzen des *Laplace-Beltrami-Operators* $-\Delta_\Gamma$ findet man in [57], [59]). Dennoch können die genannten Mannigfaltigkeiten selbst in einem beschränkten Bereich des $\mathbf{R}^n$ eine recht komplizierte Geometrie aufzeigen [29], [69], [70], so daß die für die Existenz des diskreten Spektrums notwendige Kompaktheit der Originaleinbettungen schwer nachzuweisen ist. Trotzdem gelingt es die Idee der Fourier-Darstellung auszunutzen, wenn man die Technik der lokalen Karten verwendet [59, Kap 1, § 7].

Die moderne Funktionalanalysis zeichnet noch eine dritte Art von Sobolevräumen aus. Es sind *Sobolev-Hörmander-Räume* $H_{(s)}(\Omega)$ [36] und die mit ihnen verbundene Theorie der Pseudodifferentialoperatoren [90], [101]. Im Rahmen dieser Dissertation brauchen wir weder die Pseudodifferentialoperatoren noch die Sobolev-Hörmander-Räume und werden sie deshalb hier nicht eingehend beschreiben. Wenn das Gebiet Ω den im Abschnitt 3.1.1 angenommenen Glattheitsvoraussetzungen entspricht, fallen alle drei Arten der Sobolevräume mengenmäßig und topologisch zusammen [59], [100]. Insbesondere gilt $W^{k,\lambda}(\Omega) = H^s(\Omega)$, falls $s = k + \lambda$, $\lambda \in (0,\, 1)$, $m \geq k + 1$. **Für Gebiete allgemeiner Struktur sind aber die genannten drei Arten von Sobolevräumen absolut unterschiedlich [29, S. 16].**

Beispiel: Gemäß der Spurtheorie gehört die Funktion $\ln\left(-\ln\frac{|x|}{2}\right)$ dem Raum $H^{1/2}(-1,\, 1)$ an, obwohl sie unstetig ist. Andererseits liegt die Stufen-Funktion $\mathrm{sign}(x)$ nicht in demselben Raum. Tatsächlich, da $W^{0,1/2}(-1,\, 1) = H^{1/2}(-1,\, 1)$, reicht es zu zeigen, daß $\left(\Lambda_{1/2}\ \mathrm{sign}(x)\right)$ außerhalb $H^{-1/2}(-1,\, 1)$ liegt. Einfache Berechnungen ergeben

$$\Lambda_{1/2}\ \mathrm{sign}(x) = \frac{\mathrm{sign}(x)}{x} - \frac{(1 + \mathrm{sign}(x))^2}{1 + x} - \frac{(1 - \mathrm{sign}(x))^2}{1 - x} \; . \tag{3.13}$$

Obwohl die letzten Summanden von (3.13) zu $H^{-1/2}(-1,\, 1)$ gehören, stellt der erste Summand ein unstetiges Funktional über $H^{1/2}(-1,\, 1)$ dar. Es reicht aus, eine konstante Funktion aus $H^{1/2}(-1,\, 1)$ zu betrachten.

□

Mit Hilfe der Sobolev-Lions-Räume läßt sich die Nullmenge des bereits eingeführten Einbettungsoperators G^* genau beschreiben [5]. Die Funktionale aus $\mathcal{N}(G^*)$ stellen eine lineare Kombination von sogenannten *Delta-Schichten k-ter Ordnung* $g_k\,\delta_\Gamma^k$, $k = 0, .., m-1$ dar.[6] Der Index Γ verweist darauf, daß der Träger der Delta-Schichten mit dem Rand des Gebiets Ω übereinstimmt, falls $\Gamma = \partial\Omega$. Die Funktionale werden gemäß ihrer Wirkung auf $u \in H^m(\Omega)$ definiert:

$$\left\langle\, g_k\,\delta_\Gamma^k,\, u\,\right\rangle_{H^{-m}(\Omega),\,H^m(\Omega)} = (-1)^k\,\langle g_k,\,\gamma_k\,u\rangle_{H^{-m+k+1/2}(\Gamma),\,H^{m-k-1/2}(\Gamma)}\ .$$

Mit g_k wird also ein beliebiges Element aus $H^{-m+k+1/2}(\Gamma)$ bezeichnet. Die oben stehende Definition veranschaulicht gut, warum die Funktionale aus $\mathcal{N}(G^*)$ zu allen Funktionen $u \in \mathring{H}^m(\Omega)$ orthogonal sind.

Sei $\mathcal{V}$ irgendein Sobolevraum $\mathring{H}^{s/2}(\mathrm{M})$, $H^{s/2}(\mathrm{M})$, $\mathring{H}^{s/2}(\Omega)$, $H_\Gamma^{s/2}(\Omega)$ oder $H^{s/2}(\Omega)$ mit $s \in (-\infty,\,\infty)$ und $m \geq s/2$. Sei $a(\cdot,\,\cdot)$ eine *stetige, $\mathcal{V}$-elliptische bilineare Form über* $\mathcal{V} \times \mathcal{V}$. Das heißt, es gelten folgende zwei Bedingungen:

- $\exists\, C > 0: \quad |a(u,\,v)| \leq C\,\|u\|_{\mathcal{V}}\,\|v\|_{\mathcal{V}} \qquad \forall\, u,\,v \in \mathcal{V}\ ;$
- $\exists\, c > 0: \quad a(u,\,u) \geq c\,\|u\|_{\mathcal{V}}^2 \qquad \forall\, u \in \mathcal{V}\ .$

Wir bezeichnen $A \in \mathcal{L}(\mathcal{V},\,\mathcal{V}')$ als *elliptischen Operator s-ter Ordnung*, wenn eine stetige, $\mathcal{V}$-elliptische bilineare Form $a(\cdot,\,\cdot)$ existiert und

$$a(u,\,v) = \langle\, A\,u,\,v\,\rangle_{\mathcal{V}',\,\mathcal{V}} \qquad \forall\, u,\,v \in \mathcal{V}\ . \tag{3.14}$$

Somit ist A ein elliptischer Operator s-ter Ordnung, wenn er dem entsprechenden Operator des kanonischen Isomorphismus $\Lambda_{s/2}$ spektral äquivalent ist. Der Operator A bringt manchmal eine ganze Familie von Operatoren hervor, welche durch eine Einengung des Graphen bestimmt werden kann. So entsteht die Theorie der abstrakt elliptischen Operatoren und die ihr zugehörige Untersuchung der Fredholmschen Eigenschaften [98]. Eine noch tiefere Analyse kommt aus der Theorie der Pseudodifferentialoperatoren [90], [101]: der Begriff von s-ter Ordnung wird durch den komplexeren Begriff eines *Symbols* ersetzt. Der Definitionsbereich des Pseudodifferentialoperators kann auch als direktes Produkt von mehreren Sobolevräumen unterschiedlicher Ordnungen vorliegen.

Die im laufenden Abschnitt angeführte Definition des elliptischen Operators s-ter Ordnung lässt also verschiedene Erweiterungen zu. Bemerkenswert ist auch, daß für den gegebenen Operator A unterschiedliche bilineare Formen $a(\cdot,\,\cdot)$ existieren können, für welche es eine Greensche Darstellung (3.14) gibt [59, Kap 2, § 9.5]. Umgekehrt gilt: man findet zwar in (3.14) eine passende bilineare Form, sie ist aber nicht koerziv über $\mathcal{V}$ [59, Kap 2, § 9.6]. Um die formale Koerzivität einer solchen bilinearen Form $a(\cdot,\,\cdot)$ zurückzugewinnen, muß man spezielle Hilberträume an Stelle der herkömmlichen Sobolevräume einführen. Bei der Definition neuer Räume nimmt man statt D^α feststehende Linearkombinationen als elementare Operatoren und verfolgt anschließend ähnliche Vorgehensweise.

[6] Das Funktional δ_Γ bezeichne desweiteren eine Abkürzung von δ_Γ^0.

3.1.3 Poincaré-Steklov Operatoren

Die Theorie der Poincaré-Steklov Operatoren, welche durch ein elliptisches System mit einem Operator k-ter Ordnung erzeugt werden, bildet einen bequemen analytischen Apparat zur Verkleinerung der ursprünglichen Ordnung des Operators und zur Herstellung neuer elliptischer Operatoren mit anderen Eigenschaften. Wir beschränken unsere Darlegungen auf die Poincaré-Steklov Operatoren erster Ordnung, da gerade dieser Fall ein besonderes Interesse bei der Untersuchung der numerischen Stabilität inkorrekt gestellter Optimalsteuerprobleme hervorruft. Die Diskussion der Kontrollprobleme beginnt erst in Abschnitt 3.3; wir machen hier lediglich die dazu notwendigen Vorbereitungen. Die allgemeine Theorie der Poincaré-Steklov Operatoren mit verschiedenen Anwendungen findet man z.B. bei Agoschkov, Lebedev [4], [55].

Sei Ω ein offenes, beschränktes und zusammenhängendes Gebiet im $\mathbf{R}^n$ mit unendlich glattem Rand $\partial\Omega$. Der Spuroperator γ bildet den Sobolevraum $W^1(\Omega)$ auf den Raum $H^{1/2}(\partial\Omega)$ ab [59]. Wir wählen eine offene, zusammenhängende und unendlich glatte Teilmenge $\Gamma_+ \subset \partial\Omega$ aus. Es möge dabei gelten: mes $\partial\Omega >$ mes $\Gamma_+ > 0$. Wir bezeichnen mit $\Gamma_- = \partial\Omega \setminus \overline{\Gamma}_+$ und $\mathcal{V} = W^1_{\Gamma_-}(\Omega) + W^1_{\Gamma_+}(\Omega)$ gemäß (3.10).[7] Der Spuroperator γ_+ bildet den Raum $\mathcal{V}$ auf $H^{1/2}_{00}(\Gamma_+)$ ab; dabei gilt für die Nullmenge des Operators $\mathcal{N}(\gamma_+) = W^1_{\Gamma_+}(\Omega)$. Der Spuroperator γ_- bildet den Raum $\mathcal{V}$ auf $H^{1/2}_{00}(\Gamma_-)$ ab ; dabei gilt $\mathcal{N}(\gamma_-) = W^1_{\Gamma_-}(\Omega)$. Die linearen Abbildungen γ, γ_- und γ_+ sind auch unter dem Namen *Dirichlet-Operatoren* bekannt [5], [55].

Sei $a(\cdot,\,\cdot)$ eine $\mathcal{V}$-elliptische bilineare Form. Der lineare Operator A in (3.14) besitzt im allgemeinen eine komplexe Struktur. Er enthält in sich einen *formalen Operator* A_0, welcher auf Grund der bilinearen Form $a(\cdot,\,\cdot)$ erzeugt ist und einen *Neumann-Operator* β. Der formale Operator A_0 läßt sich nach

$$a(u,\,v) = \langle\, A_0\, u,\, v\,\rangle_{\mathcal{V}_0',\,\mathcal{V}_0} \qquad \forall\, u \in \mathcal{V},\, v \in \mathcal{V}_0 \tag{3.15}$$

definieren, wobei $\mathcal{V}_0 = \mathring{W}^1(\Omega)$. Mit A_0 kann ein neuer Raum

$$\mathcal{V}(A_0) \;=\; \{u \in \mathcal{V} : A_0\, u \in L_2(\Omega)\}$$

verbunden werden. Nun gilt folgende *Greensche Formel*:

$$\begin{aligned} a(u,\,v) \;&=\; (\,A_0\, u,\, v\,)_\Omega + \langle\, \beta\, u,\, \gamma v\,\rangle_{H^{-1/2}(\partial\Omega),\,H^{1/2}(\partial\Omega)} \\ &\forall\, u \in \mathcal{V}(A_0),\, v \in \mathcal{V}\,, \end{aligned} \tag{3.16}$$

welche gleichzeitig als Definition des Neumann-Operators betrachtet werden kann [5], [55]. Die Wahl des Raumes $\mathcal{V}$ erlaubt die Zerlegung des letzten Summanden von

[7] Die Definition der *direkten Summe* $\oplus$ zweier linearer Teilräume findet man z.B. in [60]. Hier steht eine einfache Summe, da die Summanden einen gemeinsamen Teilraum $\mathring{W}^1(\Omega)$ besitzen.

(3.16) in zwei andere bilineare Komponenten, so daß gilt

$$\begin{aligned} a(u,\, v) \;=\; (\, A_0\, u,\, v\,)_\Omega \;\;+\;\; & \langle\, \beta_-\, u,\, \gamma_- v\,\rangle_{H_{00}^{-1/2}(\Gamma_-),\, H_{00}^{1/2}(\Gamma_-)} + \\ & \langle\, \beta_+\, u,\, \gamma_+ v\,\rangle_{H_{00}^{-1/2}(\Gamma_+),\, H_{00}^{1/2}(\Gamma_+)} \\ & \forall\, u \in \mathcal{V}(A_0),\, v \in \mathcal{V}\;. \end{aligned} \tag{3.17}$$

Somit sind zwei zusätzliche Neumann-Operatoren β_- und β_+ definiert. Im Hinblick auf (3.17) kann man die Operatorform der elliptischen Randwertaufgabe

$$\begin{aligned} A_0\, w &= 0 && \text{in } L_2(\Omega)\;, \\ \gamma_+\, w &= y && \text{in } H_{00}^{1/2}(\Gamma_+)\;, \\ \gamma_-\, w &= 0 && \text{in } H_{00}^{1/2}(\Gamma_-) \end{aligned} \tag{3.18}$$

mit der entsprechenden Variationsform

$$a(w_0,\, v) \;=\; -\, a(\bar\gamma_+\, y,\, v) \qquad \forall\, v \in \mathcal{V}_0 \tag{3.19}$$

durch $w = w_0 + \bar\gamma_+\, y$ in Verbindung bringen, wobei w_0 in $\mathcal{V}_0$ gesucht wird und $\bar\gamma_+ \in \mathcal{L}(H_{00}^{1/2}(\Gamma_+),\, \mathcal{V})$ einen *Fortsetzungsoperator*, den Rechtsinversen zu γ_+ mit dem Wertebereich $\mathcal{R}(\bar\gamma_+) = W_{\Gamma_-}^1(\Omega)$ darstellt [55]. Eine ähnliche Äquivalenz entsteht zwischen

$$\begin{aligned} A_0\, w &= 0 && \text{in } L_2(\Omega)\;, \\ \beta_+\, w &= g && \text{in } H_{00}^{-1/2}(\Gamma_+)\;, \\ \gamma_-\, w &= 0 && \text{in } H_{00}^{1/2}(\Gamma_-) \end{aligned} \tag{3.20}$$

und

$$a(w,\, v) \;=\; \langle\, g,\, \gamma_+ v\,\rangle_{H_{00}^{-1/2}(\Gamma_+),\, H_{00}^{1/2}(\Gamma_+)} \qquad \forall\, v \in W_{\Gamma_-}^1(\Omega)\;, \tag{3.21}$$

wobei w direkt in $W_{\Gamma_-}^1(\Omega)$ gesucht wird. Die Existenz der Lösung von (3.19) und (3.21) folgt aus dem Lax-Milgram-Lemma bei beliebigen $y \in H_{00}^{1/2}(\Gamma_+)$ bzw. $g \in H_{00}^{-1/2}(\Gamma_+)$. Somit kann man auch auf die Lösbarkeit von (3.18) und (3.20) schließen.

Auf Grund des Systems (3.20) besteht eine lineare Abhängigkeit zwischen $\gamma_+\, w$ und dem Element g. Wir können nun einen *Steklov-Operator* S folgendermaßen definieren:

$$S\, g \;=\; \gamma_+\, w \qquad \text{in } H_{00}^{1/2}(\Gamma_+)\;. \tag{3.22}$$

Ganz analog bringt das System (3.18) einen *Poincaré-Operator* P hervor:

$$P\, y \;=\; \beta_+\, w \qquad \text{in } H_{00}^{-1/2}(\Gamma_+)\;. \tag{3.23}$$

Diese Ausführungen zeigen, daß jedem elliptischen Operator A und der entsprechenden Dekomposition des Randes $\partial\Omega = \overline{\Gamma}_+ \cup \overline{\Gamma}_-$ ein Paar neuer linearer Operatoren P und S zugeordnet werden kann. Sie heißen *Poincaré-Steklov Operatoren*. Es gilt die offensichtliche Beziehung [55]:

$$S \;=\; P^{-1} \qquad\qquad \text{in} \quad \mathcal{L}\left(H_{00}^{-1/2}(\Gamma_+),\, H_{00}^{1/2}(\Gamma_+)\right) .$$

Das Paar P und S wird mit zwei Basissystemen ausgestattet. Es sind *Eigenfunktionen* $\{e_k(\tilde{x})\}_{k=1}^{\infty}$ auf Γ_+ und *fundamentale Funktionen* $\{v_k(x)\}_{k=1}^{\infty}$ im Gebiet Ω. Diese Funktionensysteme definieren sich durch

$$P\, e_k \;=\; \lambda_k\, e_k \qquad\qquad \text{in} \quad H_{00}^{-1/2}(\Gamma_+) \; ; \tag{3.24}$$

$$\begin{aligned} A_0\, v_k \;&=\; 0 && \text{in} \quad L_2(\Omega) \; , \\ \gamma_+\, v_k \;&=\; e_k && \text{in} \quad H_{00}^{1/2}(\Gamma_+) \; , \\ \gamma_-\, v_k \;&=\; 0 && \text{in} \quad H_{00}^{1/2}(\Gamma_-) \; , \end{aligned} \tag{3.25}$$

wobei die Zahlen $\{\lambda_k\}_{k=1}^{\infty}$ Eigenwerte von P darstellen. Wir setzen außerdem die Existenz eines *Biorthogonalsystems* $\{\hat{e}_k(\tilde{x})\}_{k=1}^{\infty}$ [60], [94] voraus. Mit dem Kronecker-Symbol δ_{kl} gelte

$$(\hat{e}_k,\, e_l)_\Omega \;=\; \delta_{kl} \; , \qquad (k,\, l) \in \mathbf{N}^2 \; . \tag{3.26}$$

Sei $\mathcal{V}_P = H_{00}^{1/2}(\Gamma_+)$ mit $m = 1$ nach (3.12). Der Poincaré-Operator P ist dann elliptisch mit der Ordnung 1:

$$a_P(y_1,\, y_2) = \langle\, P\, y_1,\, y_2 \,\rangle_{\mathcal{V}_P',\, \mathcal{V}_P} \qquad \forall\, y_1,\, y_2 \in \mathcal{V}_P \; , \tag{3.27}$$

wobei die entsprechende bilineare Form wie folgt definiert werden kann:

$$a_P(y_1,\, y_2) = \sum_{k=1}^{\infty} \lambda_k\, (y_1,\, \hat{e}_k)_\Omega\, (y_2,\, e_k)_\Omega \; .$$

Die Voraussetzung, daß die bilineare Form $a_P(\cdot,\, \cdot)$ $\mathcal{V}_P$-elliptisch ist, läßt sich in den meisten konkreten Fällen leicht überprüfen.

In (3.18) und (3.20) kann man die Dirichlet-Bedingungen auf Γ_- durch homogene gemischte Randbedingungen ersetzten (s. z.B. Abschnitt 3.3). Dies ergibt offenbar ein anderes Paar von Poincaré-Steklov Operatoren [55].

Wenn man P als unbeschränkten Operator in $L_2(\Gamma_+)$ betrachtet, ist oft das Verhältnis zwischen $\mathcal{D}(P)$ und $\mathring{W}^1(\Gamma_+)$ von Interesse.

3.2 Ein inkorrekt gestelltes Steuerproblem mit einem elliptischen Operator 2-ter Ordnung

3.2.1 Das Fuller–Problem

Ein Beispiel für ein inkorrekt gestelltes Optimalsteuerproblem ist das bekannte Problem von Fuller [26]. Die Optimalsteuerfunktion läßt sich dabei explizit berechnen und hat auf einem endlichen Zeitintervall unendlich viele Schaltpunkte. Solche Lösungen heißen *chattering regimes* und sind bei Steuerproblemen mit gewöhnlichen Differentialgleichungen weit verbreitet [103], [104]. Es ist offensichtlich, daß eine numerische Behandlung der Kontrollprobleme mit Chattering eine sehr komplizierte Aufgabe darstellt. Die Schwierigkeit besteht nicht nur in der wachsenden Anzahl der Schaltungen, sondern auch in der Tatsache, daß die adjungierte Zustandsfunktion im kritischen Steuerbereich sehr schnell verschwindet. Dies ruft eine extreme Instabilität der Lösungen hervor, was diese Aufgabenklasse von Steuerproblemen mit einem einfachen "bang-bang" wesentlich unterscheidet.

In der klassischen Formulierung des Fuller–Problems handelt es sich um die Steuerung mit einer Anfangswertaufgabe 2-ter Ordnung, wobei auch der Endzustand der Trajektorie vorgegeben ist. Unter weiter unten angeführten Bedingungen kann man aber die Zustandsgleichung des Fuller–Problems als Randwertproblem mit einem elliptischen Operator 2-ter Ordnung niederschreiben:

$$
\begin{aligned}
& J_0(u) = \frac{1}{2} \, \| \, \theta \, \|^2_{(0,\,1)} \xrightarrow{\;!\;} \min_{u \,\in\, \mathcal{U}_{ad}} \quad , \\
& \left\{ \begin{array}{rclcl} -\,\theta''(x) & = & u(x) & \text{in} & (0\,,\,1)\,, \\ -\,\theta'(0) & = & g\,, & \theta(1) & = \; 0 \end{array} \right.
\end{aligned}
\tag{3.28}
$$

mit

$$
\mathcal{U}_{ad} = \{ u \in L_2(0\,,\,1)\,, \;\; -1 \le u(x) \le 1 \quad \text{f.ü. in } (0,\,1) \; \} \;\; .
$$

Um das Steuerproblem (3.28) in die Operatorform aus Kapitel 2 umzuformulieren, führen wir die Substitution $\theta(x) = y(x) + g(1-x)$ aus. Man erhält:

$$
\begin{aligned}
& J(y,\,u) = \frac{1}{2} \, \| \, y - k_d \, \|^2_{(0,\,1)} \xrightarrow{\;!\;} \min_{(y,\,u) \,\in\, \mathcal{Y} \times \mathcal{U}_{ad}} \quad , \\
& -\,D^2\, y = u \qquad \text{in} \quad L_2(0,\,1)
\end{aligned}
\tag{3.29}
$$

mit $k_d(x) = g(x-1)$ und $\mathcal{Y} = \{ \, y \in W^2(0,\,1) : \, y'(0) = y(1) = 0 \, \}$. Die weiteren Angaben in (3.29) lassen sich nun sehr einfach klären: $\mathcal{K} = \mathcal{H} = \mathcal{U} = L_2(0,\,1)$ und

$A = -D^2 \in \mathcal{L}(\mathcal{Y},\, \mathcal{H})$, $B = I \in \mathcal{L}(\mathcal{U},\, \mathcal{H})$, $C = I \in \mathcal{L}(\mathcal{Y},\, \mathcal{K})$. Offensichtlich ist der Operator C kompakt.

Das Steuerproblem (3.29) besitzt immer noch eine Struktur, die verschieden ist von der in Problem **P** aus Kapitel 2, weil sich die Inhomogenität k_d unter dem Zielfunktional und nicht auf der rechten Seite der Zustandsgleichung befindet. Solche Situationen wurden in Kapitel 2 vorsätzlich vermieden, da im Falle eines beliebigen Zielelements eine endliche Störung im Zielfunktional wegen der vorkommenden Diskretisierungen entstehen kann. Wie man weiter sieht, ist die Lösung des Fuller–Problems im allgemeinen instabil, falls das Funktional J_0 gestört wird. Dennoch ist die Stabilität bezüglich des Wertes g immer vorhanden. *Somit ist die Lösung des Originalproblems (3.28) instabil bezüglich des Zielfunktionals, doch stabil bezüglich des freien Elements auf der rechten Seite der Zustandsgleichung.* Diese Eigenschaft des Steuerproblems ist auch für andere Aufgaben (s. in Kapitel 4) vorausgesetzt, was zugleich die Behandlung solcher Probleme sinnvoll macht, da die Lösung gleichmäßig von den Daten abhängt. Weil sich das Element k_d durch eine sehr einfache Funktion ergibt, kann man den Diskretisierungsfehler und somit die Störung im Zielfunktional völlig ausklammern.

Wir betrachten nun die Optimalitätsbedingungen 1-ter Ordnung für das Fuller–Problem (3.28); dabei wird die adjungierte Zustandsfunktion wieder mit $p(x)$ bezeichnet:

$$
\begin{cases} -\theta''(x) = u(x) & \text{in } (0\,,\,1)\,, \\ -\theta'(0) = g \qquad \theta(1) = 0\,, \end{cases}
$$

$$
\begin{cases} -p''(x) = \theta(x) & \text{in } (0\,,\,1)\,, \\ -p'(0) = 0 \qquad p(1) = 0\,, \end{cases} \tag{3.30}
$$

$$
(p,\, v-u)_{(0,1)} \geq 0 \qquad \forall\, v \in \mathcal{U}_{ad}\,.
$$

Falls $0 < g < g^o$ (g^o ist in (3.34) angegeben), besitzt das System der Optimalitätsbedingungen (3.30) folgende Lösung [103], [104]:

$$
u^*(x) = \begin{cases} u^o\left((g^o/g)\,x\right) & \text{, falls } x \in (0,\, g/g^o) \\ 0 & \text{, falls } x \in (g/g^o,\, 1) \end{cases} \tag{3.31}
$$

$$
\theta^*(x) = \begin{cases} (g/g^o)^2\,\theta^o\left((g^o/g)\,x\right) & \text{, falls } x \in (0,\, g/g^o) \\ 0 & \text{, falls } x \in (g/g^o,\, 1) \end{cases}
$$

$$
p^*(x) = \begin{cases} (g/g^o)^4\,p^o\left((g^o/g)\,x\right) & \text{, falls } x \in (0,\, g/g^o) \\ 0 & \text{, falls } x \in (g/g^o,\, 1) \end{cases}
$$

mit

$$
\begin{aligned}
u^o(x) &= \begin{cases} -1 & , \text{falls } x \in (0,\, s_1) \\ \vdots & \vdots \\ (-1)^{k-1} & , \text{falls } x \in (s_k,\, s_{k+1}) \\ \vdots & \vdots \end{cases} \\
\theta^o(x) &= \begin{cases} \theta_1(x) & , \text{falls } x \in (0,\, s_1) \\ \vdots & \vdots \\ (-q^2)^{k-1}\theta_2\left(\dfrac{x-s_k}{q^{k-1}}+s_1\right) & , \text{falls } x \in (s_k,\, s_{k+1}) \\ \vdots & \vdots \end{cases} \\
p^o(x) &= \begin{cases} p_1(x) & , \text{falls } x \in (0,\, s_1) \\ \vdots & \vdots \\ (-q^4)^{k-1}p_2\left(\dfrac{x-s_k}{q^{k-1}}+s_1\right) & , \text{falls } x \in (s_k,\, s_{k+1}) \\ \vdots & \vdots \end{cases}
\end{aligned}
\tag{3.32}
$$

wobei sich die Parameter $0 < s_1 < \ldots < s_k < \ldots < 1$ der Schaltpunkte nach der Formel der geometrischen Progression

$$
s_k = 1 - q^{k-1}(1-s_1)\,, \qquad k = 2,\, 3,\, ... \tag{3.33}
$$

berechnen. Die Funktionen $\theta_1(x)$ und $\theta_2(x)$ sind Polynome 2-ter Ordnung, während $p_1(x)$ und $p_2(x)$ Polynome 4-ter Ordnung sind. Die Koeffizienten der Polynome kann man ebenfalls durch die zwei Parameter s_1 und q ausdrücken. Wir lassen hier die expliziten Zusammenhänge weg. Für die obere Grenze des Parameters g gilt die Beziehung:

$$
g^o = \frac{2s_1 - (1-q)}{1+q} \tag{3.34}
$$

und die Berechnungsformeln für die Parameter s_1 und q (gemäß der Ergebnisse in [103] und [104]) lauten:

$$
\begin{aligned}
q &= \frac{1}{4}\left(3 + \sqrt{33} - \sqrt{26 + 6\sqrt{33}}\right)\,, \\
s_1 &= \sqrt[3]{-k_1 + \sqrt{k_1^2 + k_2^3}} + \sqrt[3]{-k_1 - \sqrt{k_1^2 + k_2^3}} + \frac{\alpha_1}{3\,\alpha_3}\,,
\end{aligned}
$$

wobei

$$k_1 = \frac{1}{2}\left(\frac{2\,\alpha_2^3}{27\,\alpha_3^3} - \frac{\alpha_1\,\alpha_2}{3\,\alpha_3^2} + \frac{\alpha_0}{\alpha_3}\right) ,$$

$$k_2 = \frac{1}{3}\left(\frac{3\,\alpha_1\,\alpha_3 - \alpha_2^2}{3\,\alpha_3^2}\right) ,$$

mit

$$\begin{aligned}
\alpha_0 &= -q^6 + 6q^5 - 7q^4 - 8q^3 + 21q^2 - 14q + 3 , \\
\alpha_1 &= 3q^6 - 18q^5 + 21q^4 + 12q^3 - 27q^2 + 6q + 3 , \\
\alpha_2 &= -3q^6 + 18q^5 - 21q^4 - 12q^3 + 3q^2 + 18q - 3 , \\
\alpha_3 &= q^6 - 6q^5 + 7q^4 + 4q^3 - q^2 - 14q - 7 .
\end{aligned}$$

Die Funktion (3.31) ergibt gleichzeitig eine Lösung für das eigentliche Fuller–Problem (3.28), da die notwendigen Bedingungen (3.30) im konvexen Fall auch ein hinreichendes Kriterium eines Extremums ist [37, § 1.1, Satz 5]. Somit ist die Frage der Existenz automatisch geklärt.

Um die quantitativen Eigenschaften der Lösung des Fuller–Problems zu veranschaulichen, geben wir jetzt ein paar präzis berechnete Werte an:

$$s_1 = 0.6960352598187639 , \qquad q = 0.2421213735481567 ,$$

$$g^o = 0.5105715968594087 , \qquad J_0^* = 0.004787279758066147 .$$

Die hinreichenden Bedingungen für die Eindeutigkeit der Lösung eines linear-quadratischen Optimalsteuerproblems untersuchen wir jetzt in der allgemeinen Form des Kapitels 2:

$$\mathbf{P}' : \qquad \begin{aligned} & J = \frac{1}{2}\,\|C\,y - k_d\|_{\mathcal{K}}^2 \stackrel{!}{\longrightarrow} \min_{(y,\,u)\,\in\,\mathcal{Y}\times\mathcal{U}_{ad}} , \\ & A\,y = B\,u + f \quad \text{in } \mathcal{H} . \end{aligned}$$

Satz 9 *Sei $\mathcal{U}_{ad} \subset \mathcal{U}$ eine konvexe abgeschlossene Menge. Der Hilbertraum $\mathcal{Y}$ sei in $\mathcal{K}$ und der Hilbertraum $\mathcal{U}$ sei in $\mathcal{H}$ eingebettet. Wenn der Systemoperator $A \in \mathcal{L}(\mathcal{Y},\,\mathcal{H})$ in $\mathbf{P}'$ stetig invertierbar ist und außerdem für den Operator des Zielfunktionals und den Operator der Kontrolle die Beziehungen $C = I \in \mathcal{L}(\mathcal{Y},\,\mathcal{K})$, $B = I \in \mathcal{L}(\mathcal{U},\,\mathcal{H})$ gelten, so besteht die Lösungsmenge von $\mathbf{P}'$ höchstens aus einem einzigen Element $(y^*,\,u^*)$.*

Beweis: Wir beweisen die stärkere Aussage, daß die Lösung des Systems der Optimalitätsbedingungen erster Ordnung:

$$\begin{aligned} A\,y &= u + f \quad \text{in } \mathcal{H}\,, && (3.35)\\ A^*\,p &= \Lambda_{\mathcal{K}}\,(y - k_d) \quad \text{in } \mathcal{Y}'\,, && (3.36)\\ \langle\, p,\, v - u\,\rangle_{\mathcal{U}',\mathcal{U}} &\geq 0 \qquad \forall\, v \in \mathcal{U}_{ad} && (3.37) \end{aligned}$$

eindeutig ist. Nehmen wir an, daß es zwei unterschiedliche Lösungen $(y_1^*,\, p_1^*,\, u_1^*)$ und $(y_2^*,\, p_2^*,\, u_2^*)$ dieses Systems gibt. Für das Zielfunktional J des Problems $\mathbf{P}'$ gelten offensichtliche Gleichheiten:

$$\begin{aligned} \frac{1}{2}\,\|A^{-1}\,(u_2^* + f) - k_d\|_{\mathcal{K}}^2 &= \frac{1}{2}\,\|A^{-1}\,(u_1^* + f) - k_d\|_{\mathcal{K}}^2 \;+\\ \Big(A^{-1}\,(u_1^* + f) - k_d,\, A^{-1}\,(u_2^* - u_1^*)\Big)_{\mathcal{K}} &+ \frac{1}{2}\,\|A^{-1}\,(u_2^* - u_1^*)\|_{\mathcal{K}}^2\,, \qquad (3.38)\\ \frac{1}{2}\,\|A^{-1}\,(u_1^* + f) - k_d\|_{\mathcal{K}}^2 &= \frac{1}{2}\,\|A^{-1}\,(u_2^* + f) - k_d\|_{\mathcal{K}}^2 \;+\\ \Big(A^{-1}\,(u_2^* + f) - k_d,\, A^{-1}\,(u_1^* - u_2^*)\Big)_{\mathcal{K}} &+ \frac{1}{2}\,\|A^{-1}\,(u_2^* - u_1^*)\|_{\mathcal{K}}^2\,. \qquad (3.39) \end{aligned}$$

Aus Abschnitt 2.2.1 wissen wir, daß die Operatoren $(A^{-1})^*$ und $(A^*)^{-1}$ gleich sind, falls A als stetig invertierbar gilt. Die zweiten Summanden auf der rechten Seite von (3.38) und (3.39) ergeben:

$$\begin{aligned} \langle \Lambda_{\mathcal{K}}\,(y_1^* - k_d),\, A^{-1}\,(u_2^* - u_1^*)\rangle_{\mathcal{K}',\mathcal{K}} &= \langle p_1^*,\, (u_2^* - u_1^*)\rangle_{\mathcal{U}',\mathcal{U}} \geq 0\,,\\ \langle \Lambda_{\mathcal{K}}\,(y_2^* - k_d),\, A^{-1}\,(u_1^* - u_2^*)\rangle_{\mathcal{K}',\mathcal{K}} &= \langle p_2^*,\, (u_1^* - u_2^*)\rangle_{\mathcal{U}',\mathcal{U}} \geq 0\,. \end{aligned}$$

Die Ungleichungen folgen dabei aus der Beziehung (3.37). Andererseits sehen wir, daß$A^{-1}\,(u_2^* - u_1^*) \neq 0$ in $\mathcal{K}$, weil $u_2^* - u_1^* \neq 0$ in $\mathcal{U}$ und die Voraussetzungen über die Einbettung der Hilberträume erfüllt sind. Im entgegengesetzten Fall muß $A^{-1} \in \mathcal{L}(\mathcal{H},\, \mathcal{Y})$ einen nicht nullidentischen Kern besitzen, was der stetigen Invertierbarkeit des Operators widerspricht.

Aus (3.38) und (3.39) folgt $J(y_2^*,\, u_2^*) > J(y_1^*,\, u_1^*)$ und $J(y_1^*,\, u_1^*) > J(y_2^*,\, u_2^*)$ gleichzeitig, was zum Widerspruch führt.

□

Weil das Problem (3.29) von der Form $\mathbf{P}'$ ist und offenbar die Bedingungen von Satz 9 erfüllt, besitzt das Fuller–Problem (3.28) eine einzige Lösung (3.31). Die Steuerfunktion $u^*(x)$ enthält unter den genannten Voraussetzungen im Bereich $(0,\, g/g^o)$ unendlich viele Schaltpunkte, deren Koordinaten durch (3.33) bestimmt werden können.

Im Bereich $(g/g^o,\, 1)$, auf dem die Funktion $u^*(x)$ verschwindet, tritt eine *singuläre Lösung 2-ter Ordnung* auf. Die komplete Klassifikation für solche singulären Bereiche ist nur im Falle gewöhnlicher Differentialgleichungen vorhanden, wenn das Pontrjaginsche Maximumprinzip gilt und die Pontrjaginsche Funktion in die Summe $H(p,\, y,\, u) = H_0(p,\, y) + u\, H_1(p,\, y)$ zerlegt werden kann. Falls auf dem offenen Bereich $H_1 = 0$ gilt, so ist die Trajektorie da singulär. Man leitet schließlich H_1 entlang der Lösung des Systems $2k$-mal ab, bis der nach der Differentiation entstandene Ausdruck von der Steuervariablen u explizit abhängt. Die kleinste natürliche Zahl k ergibt dann die Ordnung des singulären Bereichs [104]. Wenn die Variable u unter dem Zielfunktional des Problems nicht explizit auftritt, so wird das hinreichende Kriterium der Existenz von singulärer Lösung noch einfacher: $p(x) = 0,\ x \in \Omega'$, wobei Ω' eine offene Menge ist.

Der Punkt $x_\infty = g/g^o$ heißt der *singuläre Punkt* der Steuerung. Einen wichtigen Aspekt für die Stabilität der Lösung bei aktiven Steuerrestriktionen ergibt das asymptotische Verhalten der adjungierten Zustandsfunktion $p(x)$ in der Umgebung des singulären Punkts. Mit (3.32) und (3.33) erreicht man die Abschätzung:

$$|p(x)| \;\leq\; O(|\, x - x_\infty\,|^4)\;,$$

welche veranschaulicht, daß der "Multiplikator" im informativen Steuerbereich recht schnell verschwindet; die Beschränkung hat somit sehr unterschiedliche Aktivität an unterschiedlichen Punkten, was eine schlechte Skalierung der Daten bedeutet und Schwierigkeiten in der Numerik hervorruft.

Trotz der Eindeutigkeit der Lösung ist das Fuller–Problem (3.28) inkorrekt gestellt [40, Def. 2.1]. Sei $x_\infty < a < b < 1$ und $\delta = b - a$. Wir bilden eine Variation des Originalproblems, indem wir nur das Zielfunktional durch einen konkaven Term stören:

$$J_\delta(u) \;=\; \frac{1}{2}\,\|\,\theta\,(u)\,\|^2_{(0,\,1)} - \|\,u\,\|_{(a,\,b)}\;.$$

Die Variation ist klein

$$\|J_\delta - J_0\|_{C(\mathcal{U}_{ad})} = \sup_{v \in \mathcal{U}_{ad}} |\, J_\delta(v) - J_0(v)\,| \leq \delta\;.$$

Jedoch kann man zeigen, daß die Lösungsmenge des gestörten Problems für beliebiges $\delta > 0$ leer ist. Nehmen wir das Gegenteil an und bezeichnen mit $u^*_\delta(x)$ die Lösung der variierten Aufgabe. Für den Optimalwert des neuen Zielfunktionals gilt die globale Abschätzung:

$$J_\delta(u^*_\delta) = \frac{1}{2}\,\|\,\theta\,(u^*_\delta)\,\|^2_{(0,\,1)} - \|\,u^*_\delta\,\|_{(a,\,b)} \geq J_0(u^*_\delta) - \delta \geq J^*_0 - \delta\;. \tag{3.40}$$

Man kann jetzt eine minimierende Folge in $\mathcal{U}_{ad}$ konstruieren, welche diese untere Grenze (3.40) erreicht. Bei der Herstellung der Folge benutzen wir die Elemente des bekannten orthogonalen Basissystems von Rademacher [77, § 1.5.2.]:

$$\varrho_k(t) = \mathrm{sign}\left(\cos\left(2^{k+1}\pi t\right)\right) \qquad k = 0, 1, 2, \ldots\;.$$

Man wählt mit $u^*(x)$ nach (3.31)

$$v_k(x) \;=\; u^*(x) + \begin{cases} \varrho_k((x-a)/\delta) & , \quad x \in (a,\, b) \\ \\ 0 & , \quad \text{sonst} \end{cases} \qquad k = 0, 1, 2, .. \; .$$

Es ist nicht schwer zu prüfen, daß $\theta(v_k) \to \theta(u^*)$ in $L_2(0,\, 1)$, falls $k \to \infty$. Jedoch ist die Folge v_k im Steuerraum divergent (man nennt diese Erscheinung "sliding modes" [104]). Es widerspricht der Lösbarkeit der gestörten Extremalaufgabe und beweist damit auch die Inkorrektheit des Fuller–Problems.[8]

Oben wurde gezeigt, daß das Problem (3.28) eine Instabilität bezüglich der Störungen im Zielfunktional besitzt. Trotzdem verhält sich die Lösung dieses Problems stabil, wenn man geringe Variationen des Parameters g zuläßt. Die gleichmäßige Abhängigkeit von Daten ergibt eine gute Motivation dafür, daß die Probleme vom Typ (3.28) sinnvoll sind.

Satz 10 *Sei $0 < g_1 < g_2 < g^o$. Den Parametern g_1 und g_2 mögen die Lösungen $u^*_{g_1}(x)$ und $u^*_{g_2}(x)$ des Problems (3.28) entsprechen. Es gilt*

$$\exists\, \varepsilon > 0 \; : \quad g_2 - g_1 < \varepsilon\, g_2 \quad \Rightarrow$$

$$\left\| u^*_{g_2}(x) - u^*_{g_1}(x) \right\|_{(0,\,1)} \le \sqrt{\frac{4\,q}{(1-q)}} \sqrt{\frac{(g_2 - g_1)\, \ln\left(1 - g_1/g_2\right)}{g^o\, \ln q}} \; . \tag{3.41}$$

Beweis: Wir bezeichnen mit $x_1^{(1)}$, $x_2^{(1)}$, .., $x_k^{(1)}$, .. $x_\infty^{(1)}$ sämtliche Schaltpunkte der Lösung $u^*_{g_1}(x)$ und mit $x_1^{(2)}$, $x_2^{(2)}$, .., $x_k^{(2)}$, .. $x_\infty^{(2)}$ die von $u^*_{g_2}(x)$. Die Bedingung, daß die Punkte $x_k^{(2)}$ und $x_{k+1}^{(1)}$ zusammenfallen, läßt sich mit (3.33) folgendermaßen schreiben:

$$g_2 - g_1 = g_2 \left(1 - \frac{1 - q^{k-1}(1 - s_1)}{1 - q^k (1 - s_1)} \right) = g_2\, \frac{(1 - s_1)(1 - q) q^{k-1}}{1 - q^k (1 - s_1)} \; . \tag{3.42}$$

Wir wählen $\varepsilon = q^2$, schätzen in (3.42) nach unten ab und logarithmieren:

$$\ln\left(1 - g_1/g_2\right) \ge \ln\left(\frac{(1 - s_1)(1 - q)}{q^2} \right) + (k+1)\, \ln q \; .$$

Weil der erste Summand positiv ist, kann man die untere Grenze k_0 der Anzahl derjenigen Schaltpunkte definieren, für welche

$$x_l^{(1)} < x_l^{(2)} < x_{l+1}^{(1)} \; , \qquad l = 1, .. \, k_0 \; .$$

Es gilt

$$k_0 \;=\; \left[\frac{\ln\left(1 - g_1/g_2\right)}{\ln q} \right] - 1 \; .$$

[8] vgl. mit Beispiel 7 in [97, Kap 1, §3].

Die Operation $[\cdot]$ bezeichnet den ganzen Betrag von einem reelen positiven Wert. Mit dem gewählten ε gilt immer $k_0 \geq 1$. Man kann nun die gegenseitige Auslenkung der Steuerfunktionen auswerten.

$$\int\limits_0^{x_{k_0}^{(2)}} \left|u_{g_2}^*(x) - u_{g_1}^*(x)\right|^2 dx + \int\limits_{x_{k_0}^{(2)}}^{x_\infty^{(2)}} \left|u_{g_2}^*(x) - u_{g_1}^*(x)\right|^2 dx + 0 \quad \leq$$

$$\sum_{l=1}^{k_0} 4\,(x_l^{(2)} - x_l^{(1)}) + 4\,(x_\infty^{(2)} - x_{k_0}^{(2)}) = \frac{4}{g^o}\left(\sum_{l=1}^{k_0} s_l\,(g_2 - g_1) + g_2\,(1 - s_{k_0})\right) \quad \leq$$

$$\frac{4}{g^o}\left((g_2 - g_1)\,k_0 + \frac{(g_2 - g_1)}{1 - q}\right) \leq \frac{4\,q}{g^o(1-q)}\left((g_2 - g_1)\,\frac{\ln\,(1 - g_1/g_2)}{\ln q}\right) \quad .$$

Der letzte Ausdruck ergibt direkt (3.41).

□

Man sagt im gegebenen Fall: es besteht eine Stabilität vom Hölder-Typ mit der Ordnung $1/2 - 0$. **Die Stabilität vom Lipschitz-Typ [3], [65] kann man für das Problem P′ nicht mehr erwarten, weil die hinreichenden Optimalitätsbedingungen zweiter Ordnung gegebenenfalls fehlen.**

3.2.2 Kovergenz innerer Approximationen im Rahmen des regularisierten Strafverfahrens

Wir fahren mit der Untersuchung des MSR-Verfahrens aus Kapitel 2 fort. Das Fuller–Problem (3.28) ist ein gutes Beispiel, an dem man weitere theoretische und numerische Aussagen für das prox-regularisierte Strafverfahren erarbeiten kann. Das zu betrachtende Steuerproblem stellt eine eindimensionale Vereinfachung angewandter Probleme mit partiellen Ableitungen dar, welche wir später in Kapitel 4 behandeln. Es besteht Hoffnung, daß die meisten für die eindimensionale Reduktion aus Abschnitt 3.2.1 bekannten Eigenschaften auch im Falle zweidimensionaler Gebiete ihre Gültigkeit behalten. Die Schwierigkeiten, auf die man bei der Behandlung des Fuller–Problems stößt, sind offenbar in komplexeren Aufgaben wieder vertreten. Die vorhandene analytische Form der Lösung (3.31) spielt bei praktischen Konvergenzuntersuchungen eine wichtige Rolle, da man den Abstand zwischen einer approximativen und exakten Lösung direkt im vorgegebenen Hilbertraum messen und dadurch die Qualität der Parametereinstellungen des MSR-Verfahrens bewerten kann. Die dabei erworbenen Erfahrungen sind dann für komplexere Fälle von Interesse. Das theoretische Hauptergebnis (Satz 11) läßt sich auf Optimalsteuerprobleme mit über einem Quadrat verteilter Steuerung leicht verallgemeinern, wenn man als Systemoperator einen Laplace-Operator wählt und entsprechende Bisplines im Diskretisierungsschema anwendet.

Um ein möglichst allgemeines Modell in Betracht zu ziehen, erlauben wir in (3.29) die Steuerung nur auf dem ersten Teil des gesamten Intervalls, d. h. nur auf $(0,\ g/g^o)$, indem wir einen nichttrivialen Operator der Kontrolle ansetzen: $B = \mu\, I$. Die Funktion $\mu(x)$ ist überall gleich eins außer auf dem Teilintervall $(g/g^o,\ 1)$, auf dem sie verschwindet. Im Falle eines zweidimensionalen Gebiets Ω aus Kapitel 4 kann eine kompliziertere projizierende Funktion $\mu(x)$, $x \in \Omega$ eingeführt werden, welche auf dasjenige Teilgebiet $\Omega' \subset \Omega$ hinweist, auf dem man die verteilte Steuerung durchzuführen hat. Nach der Berechnung muß die numerische Lösung auf $\Omega \setminus \Omega'$ mit Null ersetzt werden.

Man muß folgendes Optimalsteuerproblem behandeln:

$$\begin{aligned} \mathbf{P}: \qquad & J(y,\ u) = \frac{1}{2}\ \| y - k_d \|^2_{(0,1)} \xrightarrow{\ !\ } \min_{(y,\,u)\ \in\ \mathcal{Y}\times\mathcal{U}_{ad}}\ , \\ & -D^2\, y = \mu\, u \qquad \text{in} \quad L_2(0,\ 1)\ . \end{aligned} \tag{3.43}$$

Die Menge der optimalen Steuerungen $\mathcal{U}^*$ von $\mathbf{P}$ besteht aus allen Funktionen $\mathcal{U}_{ad}$, welche mit der Funktion $u^*(x)$ in (3.31) auf dem Intervall $(0,\ g/g^o)$ übereinstimmen. Die Menge $\mathcal{U}^*$ bildet eine unendlichdimensionale konvexe und abgeschlossene Mannigfaltigkeit im Raum $L_2(0,\ 1)$. Das Problem (3.43) ist nun im klassischen Sinne der Definition 2.1 [40] inkorrekt gestellt, weil keine eindeutige Lösbarkeit vorliegt. Die Lösungsmenge $\mathcal{Z}^* = \mathcal{Y}^* \times \mathcal{U}^*$ ist durch $\mathcal{U}^*$ aufgrund der Zustandsgleichung eindeutig bestimmt.

Das in Kapitel 2 diskutierte MSR-Verfahren garantiert sowohl eine Konvergenz bezüglich des Zielfunktionals, als auch eine schwache Konvergenz der Näherungen gegen ein bestimmtes Element aus der Menge $\mathcal{U}^*$. Das ist mehr, als die nur bezüglich des Zielfunktionalwerts konvergierenden Verfahren [97, Kap 1, § 3, Satz 6] liefern. Die Verfahren ohne Regularisierung gewährleisten lediglich eine schwache Konvergenz der Näherungen gegen die unendlichdimensionale Menge $\mathcal{U}^*$ und sind deshalb numerisch nicht akzeptabel.

Man stellt wieder die Folge $(s = 0, 1, .. s(i),\ i = 0, 1, ..)$ der Ersatzprobleme des MSR-Verfahrens auf:

$$\bar{\mathbf{P}}_{i,s}: \qquad \bar{\Psi}_{i,s}(y,\ u) \xrightarrow{\ !\ } \min_{(y,\,u)\ \in\ \mathcal{Y}\times\mathcal{U}_{ad}}\ ,$$

wobei

$$\begin{aligned} \bar{\Psi}_{i,s}(y,\ u) \ &=\ J(y,\ u) + \frac{1}{2\, r_i}\ \left\| D^2\, y + \mu\, u \right\|^2_{(0,1)} + \\ &+ \frac{\chi_i}{2}\ \left\| u - u^{i,s-1} \right\|^2_{(0,1)}\ . \end{aligned} \tag{3.44}$$

Wir betrachten jetzt eine unendliche Folge von eingeschachtelten gleichmäßigen Gittern auf dem Einheitsintervall. Aufgrund der Bisektion ist die Anzahl der

elementaren Zellen des i-ten Gitters $J = 2^i$. Als Diskretisierungsparameter gilt die Gitterschrittweite $h = 2^{-i}$. Jeder Gitterzelle wird eine konstante Stufenfunktion $\psi_j,\ j = 1, 2, .. J$ zugeordnet. Den Gitterknoten ordnen wir Hermitische Splines $\zeta_k^{(0)}$ und $\zeta_l^{(1)},\ k = 0, 1, .. J-1,\ l = 1, 2, .. J$ [66], [89] zu. Auf jeden inneren Knoten fallen dabei zwei Basisfunktionen; dem ersten Knoten am Ursprungspunkt entspricht nur die Funktion $\zeta_0^{(0)}$ und dem letzten entspricht allein $\zeta_J^{(1)}$. Jetzt können wir die inneren Approximationen im Prozeßraum $\mathcal{Z} = \mathcal{Y} \times \mathcal{U} = \mathcal{Y} \times L_2(0,\, 1)$ einführen. Wir definieren die endlichdimensionalen Räume von Funktionen

$$\begin{aligned} \mathcal{Y}^h &= \left\{ y_h \in \mathcal{Y} :\ y_h(x) = \sum_{k=0}^{J-1} y_k^{(0)}\, \zeta_k^{(0)}(x) + \right. \\ &\qquad \left. \sum_{l=1}^{J} y_l^{(1)}\, \zeta_l^{(1)}(x), \quad y_k^{(0)} \in \mathbf{R},\, y_l^{(1)} \in \mathbf{R} \right\}\ , \\ \mathcal{U}^h &= \left\{ u_h \in L_2(\Omega) :\ u_h(x) = \sum_{j=1}^{J} u_j\, \psi_j(x),\ u_j \in \mathbf{R} \right\}\ . \end{aligned}$$

Die Menge $\mathcal{U}_{ad}$ diskretisieren wir folgendermaßen:

$$\mathcal{U}_{ad}^h = \{ u_h \in \mathcal{U}^h :\ -1 \le u_h(x) \le 1 \ \text{ f.ü. in } (0,\, 1) \}\ .$$

Nun wird jedem Ersatzproblem $\bar{\mathbf{P}}_{i,s}$ ein diskretisiertes Problem

$$\mathbf{P}_{i,s}^h : \qquad \bar{\Psi}_{i,s}(y_h,\, u_h) \xrightarrow{!} \min_{(y_h,\, u_h) \in \mathcal{Y}^h \times \mathcal{U}_{ad}^h} \tag{3.45}$$

zugeordnet. Bei der weiteren theoretischen Analyse setzen wir voraus, daß jedes endlichdimensionale Problem $\mathbf{P}_{i,s}^h$ exakt gelöst wird. (Man denkt z.B. an direkte Verfahren der Optimierung). Diese exakten Lösungen werden als Näherungen im MSR-Verfahren erklärt und wie in Kapitel 2 mit $(y^{i,s},\, u^{i,s})$ bezeichnet. Die Elemente $(\bar{y}^{i,s},\, \bar{u}^{i,s})$ stellen nach wie vor die exakten Lösungen von $\bar{\mathbf{P}}_{i,s}$ dar. Die Formulierung des MSR-Verfahrens im Kapitel 2 bleibt erhalten; man muß nun lediglich die Abschätzung

$$\left\| y^{i,s} - \bar{y}^{i,s} \right\|_{W^2(0,1)}^2 + \left\| u^{i,s} - \bar{u}^{i,s} \right\|_{(0,1)}^2 \ \le\ \varepsilon^2(\chi_i,\, r_i,\, h_i) \tag{3.46}$$

für (3.45) erhalten und anschließend die Ergebnisse von Satz 8 anwenden. Wir fixieren nun die Iterationsnummern i und s des MSR-Verfahrens und werden desweiteren die beiden Indizes möglichst weglassen.

Da der Systemoperator in (3.43) analytisch invertierbar ist, erlauben wir uns folgende Substitution für die Zustandsfunktion: $v := -D^2\, y$ bzw $y := K\, v$, wobei

$$K\, v(x) \ =\ \int_x^1 \left(\int_0^{\xi_1} v(\xi_2)\, d\xi_2 \right) d\xi_1\ , \tag{3.47}$$

auszunutzen. Außerdem definieren wir noch einen Raum der inneren Approximationen

$$\mathcal{V}^h = \left\{ v_h \in \mathcal{U} : \; v_h(x) = \sum_{k=0}^{J-1} v_k^{(0)} \, D_x^2 \, \zeta_k^{(0)}(x) + \sum_{l=1}^{J} v_l^{(1)} \, D_x^2 \, \zeta_l^{(1)}(x), \quad v_k^{(0)} \in \mathbf{R}, \, v_l^{(1)} \in \mathbf{R} \right\} .$$

Es ist nicht schwer zu prüfen, daß $\mathcal{U}_{ad}^h \subset \mathcal{U}^h \subset \mathcal{V}^h$ gilt. Offensichtlich gilt: $\mathcal{V}^h \subset \mathcal{U}$ und $\mathcal{Y}^h \subset \mathcal{Y}$, was genau die Bedingung von inneren Approximationen [27] befriedigt. In diesem Fall läßt sich folgendes Ergebnis für Problem $\mathbf{P}'$ auf der Seite 56 beweisen. Man kann sogar die Zustandsrestriktionen $\mathcal{Y}_{ad} \subset \mathcal{Y}$ berücksichtigen.

Lemma 9 *Man betrachtet die Hilberträume $\mathcal{U}$, $\mathcal{Y}$, $\mathcal{H}$, $\mathcal{K}$, $\mathcal{T}$, $\mathcal{S}$, $\mathcal{Z}$ und die linearen Operatoren A, B, C, W, N mit gleichen Eigenschaften wie in Kapitel 2. Die Existenz und Berechenbarkeit der primären Konstanten aus dem Anhang A seien vorausgesetzt. Seien k_d, f und t_d beliebige Elemente aus $\mathcal{K}$, $\mathcal{H}$ und $\mathcal{T}$. Für den approximierenden Prozeßraum gelte: $\mathcal{Z}^h \subset \mathcal{Z}$. Die Mengen $\mathcal{Z}_{ad}^h \subset \mathcal{Z}^h$ und $\mathcal{Z}_{ad} \subset \mathcal{Z}$ seien konvex und abgeschlossen in $\mathcal{Z}$. Gegeben seien nun zwei Elemente*

$$\bar{z} = (\bar{y}, \, \bar{u}) = \operatorname*{argmin}_{(y, \, u) \in \mathcal{Z}_{ad}} \Phi(y, \, u) \qquad \textit{und} \qquad \bar{z}_h = (\bar{y}_h, \, \bar{u}_h) = \operatorname*{argmin}_{(y_h, \, u_h) \in \mathcal{Z}_{ad}^h} \Phi(y_h, \, u_h) \; ,$$

wobei

$$\Phi(y, \, u) = \frac{1}{2} \, \| C \, y - k_d \|_{\mathcal{K}}^2 + \frac{1}{2 \, r_i} \, \| W (A \, y - B \, u - f) \|_{\mathcal{S}}^2 + \frac{\chi_i}{2} \, \| N \, u - t_d \|_{\mathcal{T}}^2 \; . \tag{3.48}$$

Für diese Elemente mit (2.51) gilt folgende Abschätzung:

$$\| \bar{z}_h - \bar{z} \|_{\mathcal{Z}}^2 \leq \frac{\mathrm{E}^2(y_h, \, u_h) + 2 \sqrt{2 \, \Phi(\bar{y}, \, \bar{u})} \, \mathrm{E}(y_h, \, u_h)}{2 \, c_\kappa(\chi_i, \, r_i)} \qquad \forall \, (y_h, \, u_h) \in \mathcal{Z}_{ad}^h \; , \tag{3.49}$$

wobei

$$\mathrm{E}(y_h, \, u_h) = \left(M_C^2 \, \| y_h - \bar{y} \|_{\mathcal{Y}}^2 + \frac{2 \, M_W^2}{r_i} \, \| A (y_h - \bar{y}) \|_{\mathcal{H}}^2 + \left(\chi_i M_N^2 + \frac{2 \, M_W^2 \, M_B^2}{r_i} \right) \, \| u_h - \bar{u} \|_{\mathcal{U}}^2 \right)^{1/2} . \tag{3.50}$$

Beweis: Wir benutzen ferner die Beweistechnik aus Satz 4.1 von Kapitel 1 in [27]. Statt einer elliptischen bilinearen Form wird die zweite Gâteaux-Ableitung $\nabla^2_{zz}\Phi(z_0)(z)(z)$ (an beliebiger Stelle z_0) genommen. Man schreibt die Optimalitätsbedingungen für $\bar{z}$ und $\bar{z}_h$ nieder, addiert sie miteinander und gruppiert anschließend die Glieder um. Gemäß [27] ergibt sich

$$\frac{1}{2}\,\nabla^2_{zz}\Phi(z_0)(\bar{z}_h-\bar{z})(\bar{z}_h-\bar{z}) \leq \frac{1}{2}\,\nabla^2_{zz}\Phi(z_0)(z_h-\bar{z})(z_h-\bar{z}) +$$

$$\nabla^2_{zz}\Phi(z_0)(\bar{z})(z_h-\bar{z}) - (k_d,\, C(y_h-\bar{y}))_{\mathcal{K}} - \frac{1}{r_i}\,(W\,f,\, WA(y_h-\bar{y}))_{\mathcal{S}} +$$

$$\frac{1}{r_i}\,(W\,f,\, WB(u_h-\bar{u}))_{\mathcal{S}} - \chi_i\,(t_d,\, N(u_h-\bar{u}))_{\mathcal{T}} \leq$$

$$\frac{1}{2}\,\nabla^2_{zz}\Phi(z_0)(z_h-\bar{z})(z_h-\bar{z}) + ((C\,\bar{y}-k_d),\, C(y_h-\bar{y}))_{\mathcal{K}} +$$

$$\left(\frac{1}{\sqrt{r_i}}\,W(A\,\bar{y}-B\,\bar{u}-f),\, \frac{1}{\sqrt{r_i}}\,W(A(y_h-\bar{y})-B(u_h-\bar{u}))\right)_{\mathcal{S}} +$$

$$(\sqrt{\chi_i}\,(N\,\bar{u}-t_d),\, \sqrt{\chi_i}\,N(u_h-\bar{u}))_{\mathcal{T}} \leq \frac{1}{2}\,\nabla^2_{zz}\Phi(z_0)(z_h-\bar{z})(z_h-\bar{z}) +$$

$$\sqrt{2\,\Phi(\bar{y},\,\bar{u})}\,\left(M_C^2\,\|y_h-\bar{y}\|^2_{\mathcal{K}} + \frac{M_W^2}{r_i}\,\left(\|A(y_h-\bar{y})\|_{\mathcal{H}} + M_B\,\|u_h-\bar{u}\|_{\mathcal{U}}\right)^2 +\right.$$

$$\left.\chi_i\,M_N^2\|u_h-\bar{u}\|^2_{\mathcal{U}}\right)^{1/2} \leq \frac{1}{2}\,\mathrm{E}^2(y_h,\,u_h) + \sqrt{2\,\Phi(\bar{y},\,\bar{u})}\,\mathrm{E}(y_h,\,u_h)$$

$$\forall\,(y_h,\,u_h) \in \mathcal{Z}^h_{ad}\,.$$

Der vorletzte Übergang in der Kette gilt aufgrund der Cauchy–Schwarz Ungleichung, die auf die drei letzten Summanden des vorstehenden Ausdrucks angesetzt werden kann. Die weitere Abschätzung von $\nabla^2_{zz}\Phi(z_0)(\bar{z}_h-\bar{z})(\bar{z}_h-\bar{z})$ nach unten erfolgt ganz analog wie in Abschnitt 2.2.2.

□

Nach der Substitution (3.47) für die Zustandsfunktion $y(x)$ nimmt das zu minimierende Funktional (3.44) folgende Form an:

$$\begin{aligned}\bar{\Psi}_{i,s}(v,\,u) &= \frac{1}{2}\,\|\,K\,v-k_d\,\|^2_{(0,1)} + \frac{1}{2\,r_i}\,\|v-\mu\,u\|^2_{(0,1)} + \\ &\quad + \frac{\chi_i}{2}\,\left\|u-u^{i,s-1}\right\|^2_{(0,1)}\,. \end{aligned} \tag{3.51}$$

Die entsprechenden Lösungen von $\bar{\mathbf{P}}_{i,s}$ (bzgl. neuer und alter Variablen) sind miteinander verbunden: $\bar{v} = -D^2\bar{y}$ und $\bar{y} = K\,\bar{v}$. Wir untersuchen nun die Optimalitätsbedingungen 1-ter Ordnung für das Problem $\bar{\mathbf{P}}_{i,s}$ in neuen Variablen. Weil der Operator $K \in \mathcal{L}(L_2(0,\,1))$ selbstadjungiert ist, kann man für $\bar{v}(x)$ schreiben:

$$\begin{aligned} K(K\,\bar{v} - k_d) \;+\; \frac{1}{r_i}\,(\bar{v} - \mu\bar{u}) &= 0\;, \\ \left(I + r_i\,K^2\right)\,\bar{v} &= \mu\bar{u} + r_i K\,k_d\;, \qquad (3.52) \\ \frac{\mu\bar{u} - \bar{v}}{r_i} &= \bar{w} \qquad (3.53) \end{aligned}$$

mit

$$\bar{w} \;=\; K(K\,\bar{v} - k_d)\;. \qquad (3.54)$$

Weil der Operator $\mu \in \mathcal{L}(L_2(0,\,1))$ auch selbstadjungiert ist, schreibt man für $\bar{u}(x)$ mit (3.53) und (3.54) folgendes:

$$\begin{aligned} -\frac{1}{r_i}\,(\bar{v} - \mu\bar{u},\,\mu(\tilde{u} - \bar{u}))_{(0,1)} + \chi_i\,(\bar{u} - u^{i,s-1},\,\tilde{u} - \bar{u})_{(0,1)} &\geq 0\;, \\ (\chi_i\,(\bar{u} - u^{i,s-1}) + \frac{\mu(\mu\bar{u} - \bar{v})}{r_i},\,(\tilde{u} - \bar{u}))_{(0,1)} &\geq 0\;, \\ (\chi_i\,(\bar{u} - u^{i,s-1}) + \mu\bar{w},\,(\tilde{u} - \bar{u}))_{(0,1)} &\geq 0\;. \qquad (3.55) \end{aligned}$$

Die letzten Ungleichungen gelten für beliebiges $\tilde{u} \in \mathcal{U}_{ad}$. Falls $x \in (0,\,1):\ \mu(x) = 0$, so gilt $\bar{u}(x) = u^{i,s-1}$. Andernfalls gilt

$$\bar{u}(x) = \begin{cases} -1 & \text{falls} \quad \bar{w}(x) \geq \chi_i\left(1 + u^{i,s-1}(x)\right)\;, \\ 1 & \text{falls} \quad \bar{w}(x) \leq -\chi_i\left(1 - u^{i,s-1}(x)\right)\;, \\ u^{i,s-1}(x) - \dfrac{\bar{w}(x)}{\chi_i} & \text{sonst}\;. \end{cases} \qquad (3.56)$$

Das Einheitsintervall wird vom i-ten Gitter in gleiche Teilintervalle (in gleiche offene Zellen) $\Omega_j,\ j = 1, 2, .. J$ eingeteilt. Es kann maximal eine Zelle Ω_{j_μ} die Punkte x aus den beiden Bereichen, $\mu(x) = 1$ und $\mu(x) = 0$, enthalten, wobei j_μ den entsprechenden Index eines solchen Teilintervalls bezeichnet. Dem Hauptsatz geht nun folgendes Lemma voraus.

Lemma 10 *Sei $j \neq j_\mu$ und $r_i < 1$. Dann gilt*

$$|\bar{u}'(x)| \leq \frac{c_g(r_i)}{\chi_i} \qquad \textit{f.ü. in } \Omega_j \qquad (3.57)$$

mit

$$c_g(r_i) = \left(\frac{g}{2} + \frac{\sqrt{3} + g\, r_i}{\sqrt{3}\,(1 - r_i)} \right) \ . \tag{3.58}$$

Beweis: Weil die Funktion $u^{i,s-1}(x)$ aufgrund der gewählten inneren Approximation $\mathcal{U}^h$ stückweise konstant ist, verschwindet ihre erste Ableitung in jedem Punkt $x \in \Omega_j$. Falls $\mu(x) = 0$, so gilt $\bar{u}'(x) = 0$. Andernfalls gilt

$$\bar{u}'(x) = \begin{cases} 0 & \text{falls} \quad \bar{w}(x) > w_-(x) \ , \\ 0 & \text{falls} \quad \bar{w}(x) < w_+(x) \ , \\ -\dfrac{\bar{w}'(x)}{\chi_i} & \quad w_-(x) < \bar{w}(x) < w_+(x) \end{cases} \tag{3.59}$$

mit

$$w_-(x) = \chi_i \left(1 + u^{i,s-1}(x)\right) \quad \text{und} \quad w_+(x) = -\chi_i \left(1 - u^{i,s-1}(x)\right) \ .$$

Die letzte Darstellung von $\bar{u}'(x)$ liefert folgende Abschätzung:

$$|\, \bar{u}'(x)\,| \le \frac{1}{\chi_i} \max_{\xi \in (0,1)} |\, \bar{w}'(\xi)\,| \qquad \text{f.ü. in } \Omega_j \ , \tag{3.60}$$

wobei $j \neq j_\mu$. Offensichtlich ist die Norm der Operatoren $K \in \mathcal{L}(L_2(0,\,1))$ und $\mu \in \mathcal{L}(L_2(0,\,1))$ durch die Eins beschränkt. Wir nutzen ferner Abschätzungen für die Norm des inversen schwach gestörten Identitätsoperators [94, S. 140] aus. Weil $\|K\| \le 1$, gilt entsprechend (3.52) mit $r_i < 1$

$$\|\bar{v}\|_{(0,1)} \le \left\|(I + r_i K^2)^{-1}\right\| \left(\|\mu\| \|\bar{u}\|_{(0,1)} + r_i \, \|K\, k_d\|_{(0,1)} \right) \ \le$$

$$\frac{\|\bar{u}\|_{(0,1)} + r_i \, \|K\, k_d\|_{(0,1)}}{1 - r_i \, \|K\|^2} \le \frac{\|\bar{u}\|_{(0,1)} + r_i \, \|K\, k_d\|_{(0,1)}}{1 - r_i} \quad . \tag{3.61}$$

Es ist nicht schwer nachzuprüfen, daß $\|\bar{u}\|_{(0,1)} \le 1$ und $\|K\, k_d\|_{(0,1)} \le g/\sqrt{3}$ gilt. Außerdem sind folgende Abschätzungen leicht nachzuweisen:

$$\left|(K\, k_d(x))'\right| \ \le \ \frac{g}{2} \qquad \text{f.ü. in } \Omega_j \ , \tag{3.62}$$

$$\left|\left(K^2\, \bar{v}(x)\right)'\right| \ \le \ \|\bar{v}\|_{(0,1)} \qquad \text{f.ü. in } \Omega_j \ . \tag{3.63}$$

Wenn man (3.61), (3.62) und (3.63) in (3.54) setzt, so erhält man mit (3.60) die Behauptung des Lemmas.

□

Der nächste Sazt ist praktisch eine Folgerung von Satz 8. Er zeigt: die Folge von Lösungen $z^{i,s}$ zu den Ersatzproblemen $\mathbf{P}^h_{i,s}$ konvergiert schwach gegen irgendein Element der Lösungsmenge $\mathcal{Z}^*$ des Originalproblems $\mathbf{P}$. Die im Satz angeführten Bedingungen für die Einstellung der Verfahrensparameter sind für die Konvergenz hinreichend, vorausgesetzt man hält $h_i = 2^{-i}$ fest.

Satz 11 *Sei $z^{i,s}$ die Folge der Näherungen im MSR-Verfahren, welche sich als Lösungen von $\mathbf{P}^h_{i,s}$ ergeben. Man berechnet die Größen ε_i folgendermaßen:*

$$\varepsilon_i \;=\; \sqrt{\frac{r_i\chi_i + 2}{2\,\chi_i}\,\left(\mathrm{E}_i^2 + 2\sqrt{2\Phi_i}\,\mathrm{E}_i\right)} \tag{3.64}$$

mit

$$\Phi_i \;=\; \min\left\{\; \max_{s=\overline{1,\,s(i)}} \bar{\Psi}_{i,s}(z^{i,s-1})\,,\; \|k_d\|^2_{(0,\,1)} + d_2^2/2 \;\right\}\,, \tag{3.65}$$

$$\begin{aligned}\mathrm{E}_i \;=\; &\left(\left(1+\frac{2}{r_i}\right)\left(h_i^2 c_g^2(r_i)\,\left(\frac{1}{\chi_i}+r_i\right)^2 + 4h_i\right) + \right.\\ &\left.\left(\chi_i+\frac{2}{r_i}\right)\left(\frac{h_i^2}{\chi_i^2}\,c_g^2(r_i) + 4h_i\right)\right)^{1/2}\end{aligned} \tag{3.66}$$

und $c_g(r_i)$ nach (3.58). Die Folge $z^{i,s}$ hat immer endlich viele Glieder in s , d. h.

$$s(i) \le (d_2 + \varepsilon_i)/\vartheta_i \qquad i = 1, 2, 3, .. \tag{3.67}$$

und konvergiert schwach in $W^2(0,\,1) \times L_2(0,\,1)$ **gegen ein Element** *aus der Optimalmenge $\mathcal{Z}^*$, wenn*

$$0 \le \underline{\chi} < \chi_i \le 1\,, \qquad 0 < r_i < 1\,, \qquad 0 < \vartheta_i\,,$$

$$\delta_i \ge \varepsilon_i + \sqrt{6\,(d_2+\varepsilon_i)\,(\varepsilon_i+\vartheta_i) + \frac{4d_1\,r_i}{\chi_i}}\;\;, \tag{3.68}$$

$$\sum_{i=1}^{\infty}\sqrt{\frac{r_i}{\chi_i}} < \infty \qquad \textit{und} \qquad \sum_{i=1}^{\infty}\sqrt[4]{\frac{h_i}{\chi_i^2\,r_i}} < \infty\,. \tag{3.69}$$

Dabei wird die Existenz und a priori Berechenbarkeit der Konstanten d_1 und d_2 gemäß Kapitel 2 vorausgesetzt.

Beweis: Um die Abschätzung (3.46) zu erhalten, benutzen wir die Ergebnisse von Lemma 9. Man muß dabei lediglich die Lipschitz-Konstanten in (3.50) für das Problem $\bar{\mathbf{P}}_{i,s}$ ausrechnen und günstige Elemente $y_h \in \mathcal{Y}^h$ und $u_h \in \mathcal{U}^h_{ad}$ finden, die

den Elementen $\bar{y}$ und $\bar{u}$ "nah genug" liegen. Gemäß [66, Kapitel 2, § 1, Satz 1] kann man

$$u_h(x) = \sum_{j=1}^{J} u_j \psi_j(x) \qquad \text{mit} \qquad u_j = \frac{1}{h_i} \int\limits_{\Omega_j} \bar{u}(\xi)\, d\xi$$

wählen. Selbstverständlich liegt ein solches u_h in $\mathcal{U}_{ad}^h$. Nun gilt

$$\|u_h - \bar{u}\|_{(0,1)}^2 \le h_i^2 \sum_{j \neq j_\mu} \int\limits_{\Omega_j} |\bar{u}'(\xi)|^2\, d\xi + \int\limits_{\Omega_{j_\mu}} |u_h(\xi) - \bar{u}(\xi)|^2\, d\xi \quad \le$$

$$h_i^2 \frac{c_g^2(r_i)}{\chi_i^2} \sum_{j \neq j_\mu} h_i + 4h_i \le h_i^2 \frac{c_g^2(r_i)}{\chi_i^2} + 4h_i \quad . \tag{3.70}$$

Man hat dabei die Ergebnisse von Lemma 10 benutzt.

Weil $\|y_h - \bar{y}\|_{(0,1)} = \|K(-D^2)(y_h - \bar{y})\|_{(0,1)} \le \|(-D^2)(y_h - \bar{y})\|_{(0,1)}$, so reicht es in (3.50) nur das Glied mit $\|A(y_h - \bar{y})\|_{\mathcal{H}}$ abzuschätzen. Dazu benötigt man die Auswertung von $|\bar{v}'(x)|$. Aus (3.53) folgt: falls $x \in (0,\, 1) : \ \mu(x) = 0$, so gilt $\bar{v}'(x) = -r_i \bar{w}'(x)$; sonst gilt $\bar{v}'(x) = \bar{u}'(x) - r_i \bar{w}'(x)$. Endgültig erhält man für $j \neq j_\mu$

$$|\, \bar{v}'(x)\,| \le |\, \bar{u}'(x)\,| + r_i\, |\, \bar{w}'(x)\,| \qquad \text{f.ü. in } \Omega_j\ . \tag{3.71}$$

Gemäß [66, Kapitel 2, § 1, Satz 1] kann man wieder eine Approximation

$$v_h(x) = \sum_{j=1}^{J} v_j \psi_j(x) \qquad \text{mit} \qquad v_j = \frac{1}{h_i} \int\limits_{\Omega_j} \bar{v}(\xi)\, d\xi$$

wählen. Selbstverständlich gilt $v_h \in \mathcal{U}^h$. Aufgrund $\mathcal{U}^h \subset \mathcal{V}^h$ existieren solche Gewichte $v_k^{(0)}$ und $v_l^{(1)}$, daß die andere Darstellung für $v_h(x)$ vorliegt:

$$v_h(x) = \sum_{k=0}^{J-1} v_k^{(0)} D_x^2\, \zeta_k^{(0)}(x) + \sum_{l=1}^{J} v_l^{(1)} D_x^2\, \zeta_l^{(1)}(x)\ .$$

Als y_h nimmt man dann

$$y_h(x) = \ -\sum_{k=0}^{J-1} v_k^{(0)}\, \zeta_k^{(0)}(x) - \sum_{l=1}^{J} v_l^{(1)}\, \zeta_l^{(1)}(x)\ .$$

Weiter gilt:

$$\|v_h - \bar{v}\|_{(0,1)}^2 \le h_i^2 \sum_{j \neq j_\mu} \int\limits_{\Omega_j} |\bar{v}'(\xi)|^2\, d\xi + \int\limits_{\Omega_{j_\mu}} |v_h(\xi) - \bar{v}(\xi)|^2\, d\xi \quad \le$$

$$h_i^2 \left(\frac{c_g(r_i)}{\chi_i} + r_i\, c_g(r_i) \right)^2 + 4h_i \quad . \tag{3.72}$$

Hierbei benutzt man wieder die Ergebnisse von Lemma 10. Die Abschätzungen (3.70), (3.72) und (3.49) zeigen, daß die zweite Bedingung in (3.69) die gewünschte Beziehung

$$\sum_{i=1}^{\infty} \|(y^{i,s},\, u^{i,s}) - (\bar{y}^{i,s},\, \bar{u}^{i,s})\|_{\mathcal{Z}} < \infty \tag{3.73}$$

mit $\mathcal{Z} = W^2(0,\,1) \times L_2(0,\,1)$ impliziert. Die Anwendung von Satz 8 schließt nun den geführten Beweis ab.

Man benötigt noch eine zusätzliche Erklärung für (3.65). Aus Lemma 8 folgt: $\bar{\Psi}_{i,s}(\bar{z}^{i,s}) \le \bar{\Psi}_{i,s}(z^{i,s-1})$. Analog wie im Beweis von Satz 4 erhält man eine absolute obere Grenze für $\bar{\Psi}_{i,s}(\bar{z}^{i,s})$, welche sich auf die Beschränktheit des Steuerbereiches stützt und im Falle des Fuller-Problems explizit berechnet werden kann. Die erste Ungleichung ist für praktische Numerik von Vorteil; die zweite Abschätzung brauchen wir trotzdem, um das richtige asymptotische Ergebnis (3.69) zu sichern.

□

Man könnte genausogut einen nichttrivialen Operator des Zielfunktionals C einführen, der eine Projektion auf ein Teilinterval bewirkt. Man hätte dann denselben Effekt erreicht: das Steuerproblem $\mathbf{P}$ würde nichteindeutig lösbar. Die Nichtinjektivität von C beeinträchtigt auf keinen Fall das sich auf die Konvergenz der inneren Approximationen beziehende Hauptergebnis.

Wenn $\mu(x) = 1,\ \forall\, x \in (0,\,1)$ oder der Schaltpunkt von $\mu(x)$ mit einem der Gitterknoten zusammenfällt, so kann die letzte Bedingung in (3.69) gemildert werden:

$$\sum_{i=1}^{\infty} \sqrt{\frac{h_i}{\chi_i^2 \sqrt{r_i}}} < \infty\ . \tag{3.74}$$

3.2.3 Numerische Tests

Wenn man die Räume $\mathcal{Y}^h$ und $\mathcal{U}^h$ betrachtet, so kann man sehen, daß die Lösung des diskretisierten Hilfsproblems (3.45) in der entsprechenden Auswertung der Koeffiziente $\{y_k^{(0)}\}_{k=0}^{J-1}$, $\{y_l^{(1)}\}_{l=1}^{J}$ und $\{u_j\}_{k=1}^{J}$ besteht. Wir ordnen diese Werte in dem Prozessvektor $\mathbf{z} = (\mathbf{y}_1,\, \mathbf{y}_2,\, \mathbf{u})^{\mathrm{T}}$ an, der somit drei Teilvektoren enthalten muß. Nun kann $\mathbf{P}_{i,s}^h$ in folgendes quadratische Optimierungsproblem (QP) umgewandelt werden:

$$\mathbf{QP}(\chi_i,\, r_i): \qquad \hat{\Psi}(\mathbf{z}) = \frac{1}{2}\, \mathbf{z}^{\mathrm{T}}\, H\, \mathbf{z} - \mathbf{f}^{\mathrm{T}}\, \mathbf{z} \;\to\; \min_{\mathbf{z} \in \hat{\mathcal{Z}}_{ad}} \tag{3.75}$$

mit

$$H = \begin{bmatrix} A_1 & A_q & A_{r1} \\ A_q^{\mathrm{T}} & A_2 & A_{r2} \\ A_{r1}^{\mathrm{T}} & A_{r2}^{\mathrm{T}} & A_u \end{bmatrix}, \qquad \mathbf{f} = (\mathbf{f}_1,\, \mathbf{f}_2,\, \mathbf{f}_u)^{\mathrm{T}}\ .$$

Der Inhalt der Matrizen und Vektoren läßt sich wie folgt beschreiben:

$$
\begin{aligned}
(A_1)_{k'k} &= \left(\zeta_{k'}^{(0)},\, \zeta_k^{(0)}\right)_{(0,1)} + \frac{1}{r_i}\,\left(D^2\,\zeta_{k'}^{(0)},\, D^2\,\zeta_k^{(0)}\right)_{(0,1)}\ ,\\
(A_q)_{k'l} &= \left(\zeta_{k'}^{(0)},\, \zeta_l^{(1)}\right)_{(0,1)} + \frac{1}{r_i}\,\left(D^2\,\zeta_{k'}^{(0)},\, D^2\,\zeta_l^{(1)}\right)_{(0,1)}\ ,\\
(A_2)_{l'l} &= \left(\zeta_{l'}^{(1)},\, \zeta_l^{(1)}\right)_{(0,1)} + \frac{1}{r_i}\,\left(D^2\,\zeta_{l'}^{(1)},\, D^2\,\zeta_l^{(1)}\right)_{(0,1)}\ ,\\
(A_{r1})_{k'j} &= \frac{1}{r_i}\,\left(D^2\,\zeta_{k'}^{(0)},\, \mu\,\psi_j\right)_{(0,1)}\ ,\\
(A_{r2})_{l'j} &= \frac{1}{r_i}\,\left(D^2\,\zeta_{l'}^{(1)},\, \mu\,\psi_j\right)_{(0,1)}\ ,\\
(A_u)_{j'j} &= \frac{1}{r_i}\,\left(\mu\,\psi_{j'},\, \mu\,\psi_j\right)_{(0,1)} + \chi_i\,\left(\psi_{j'},\, \psi_j\right)_{(0,1)}\ ;\\
(\mathbf{f}_1)_k &= \left(\zeta_k^{(0)},\, k_d\right)_{(0,1)}\ ,\\
(\mathbf{f}_2)_l &= \left(\zeta_l^{(1)},\, k_d\right)_{(0,1)}\ ,\\
(\mathbf{f}_u)_j &= \chi_i\,\left(\psi_j,\, u^{i,s-1}\right)_{(0,1)}\ .
\end{aligned}
$$

Die Menge $\hat{\mathcal{Z}}_{ad}$ in (3.75) besteht aus dem Produkt $\mathbf{R}^J \times \mathbf{R}^J \times \hat{\mathcal{U}}_{ad}$, wobei

$$
\hat{\mathcal{U}}_{ad} = \left\{\mathbf{u} = (u_1,\, u_2,\, ..,\, u_J) \in \mathbf{R}^J :\ |u_j| \le 1\,,\ j = 1,..\,J\right\}\ .
$$

Jedes **QP**-Problem hängt von zwei Parametern χ_i und r_i ab. Die auf dasselbe Diskretisierungsniveau bezogenen Probleme sind außerdem über den Prox-Term miteinander gekoppelt. Ein weiterer Zusammenhang zwischen den **QP**'s entsteht dadurch, daß die Lösung des vorangegangenen Problems für den Startpunkt des nächsten Problems verwendet wird. Diese Art des Zusammenhangs bleibt offenbar erhalten, wenn man statt des MSR-Verfahrens das T-Verfahren gemäß Kapitel 2 anwendet. In diesem Fall nehmen wir $\mathbf{f}_u = 0$ an.

Außer der Stufenfunktionen $\psi_j(x)$ kann man stückweise linearen Basisfunktionen $\phi_j(x)$ zur Diskretisierung der Steuerung verwenden. Die Bedingung $\mathcal{U}^h \subset \mathcal{V}^h$ bleibt dabei erhalten und man kann alle Überlegungen im Abschnitt 3.2.2 ohne große Schwierigkeiten wiederholen. Dennoch ergeben die "Dach"-Funktionen bei praktischen Berechnungen bessere Approximationen der Lösung, ohne die Kondition der Matrix A_u wesentlich zu beeinträchtigen. In den unten stehenden Tests wird nur der

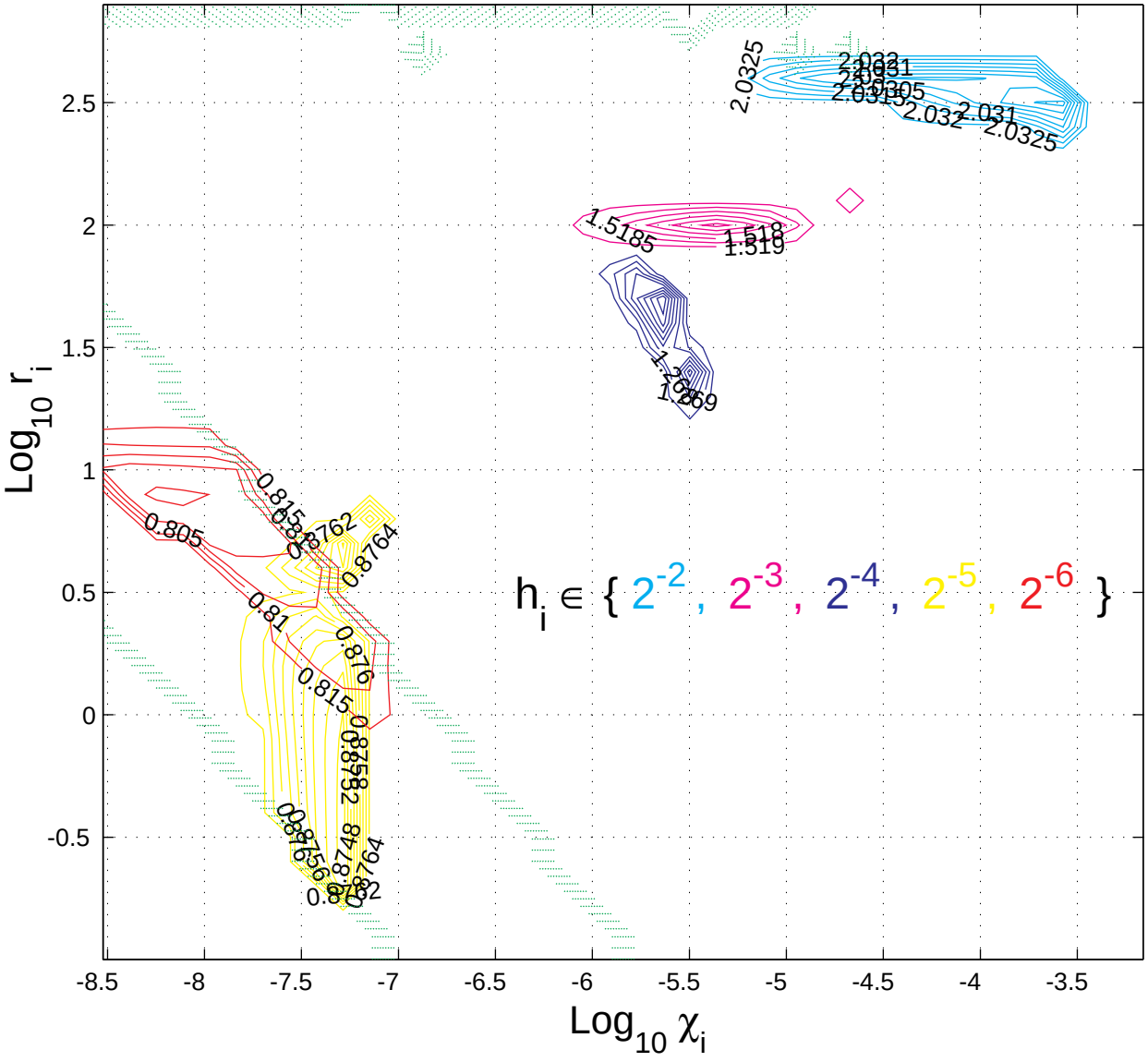


Abbildung 3.1: Tichonov-Regularisierung, $f_i(\chi_i, r_i) = \|z^i - z^*\|_Z$ Abstand bis zur exakten Lösung

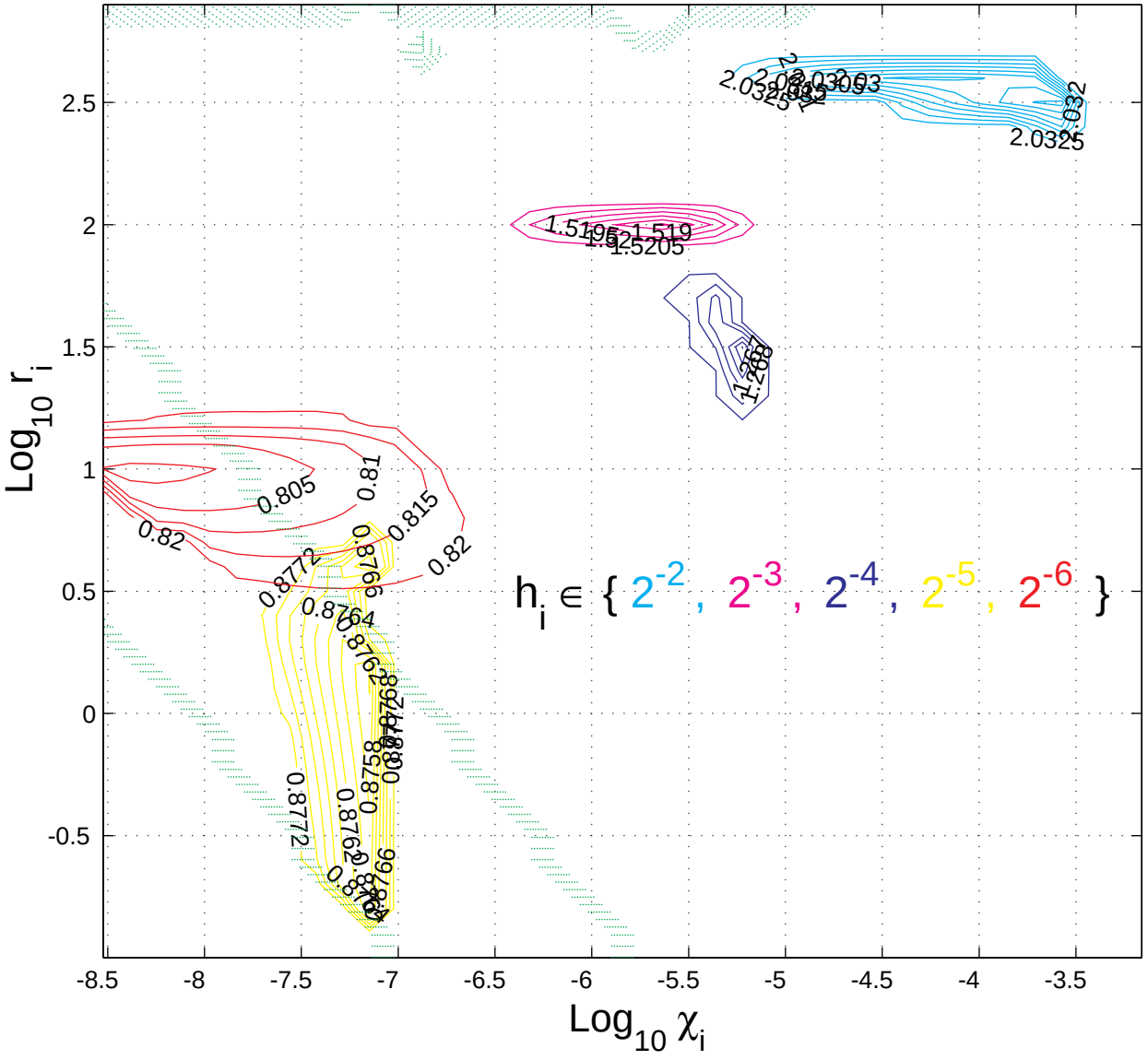


Abbildung 3.2: Prox-Regularisierung (OSR), $f_i(\chi_i, r_i) = \|z^i - z^*\|_Z$ Abstand bis zur exakten Lösung

Fall mit den stückweise linearen Funktionen betrachtet. Es wird außerdem immer $\mu(x) = 1$ vorausgesetzt.

Der Übergang von Gitter zu Gitter wird mit einer exakten Projizierung für Steuer- und Zustandskomponenten vollgezogen. Es werden dabei die im Konvergenzbeweis verwendeten Bedingungen $u^{i+1,0} = u^{i,s(i)}$ und $y^{i+1,0} = y^{i,s(i)}$ gesichert. Nach der Projizierung erhält man die neue Hesse-Matrix H und die neuen Vektoren $\mathbf{z}$ und $\mathbf{f}$ der größeren Dimension.

Test 1: Die Steuerung des OSR- und T-Verfahrens erfolgt auf dieselbe Weise durch Einstellung der Parameter χ_i und r_i zu jedem **QP**. Man startet den Rechenprozess auf dem Gitter $h_2 = 2^{-2}$ und läßt ihn bis zum Gitter $h_7 = 2^{-7}$ laufen. Anhand der analytischen Form der exakten Lösung aus Abschnitt 3.2.1 kann man den Regularisierungs- und Strafparameter auf dem laufenden Gitter "optimal" wählen, indem man durch die Parametereinstellung den Abstand $f_i(\chi_i, r_i)$ zwischen der Näherung und der exakten Lösung minimiert. Der Abstand muß in der Norm des vorgegebenen Steuerprozessraumes $\mathcal{Z}$ gemessen werden. Die Tatsache, daß solche Einstellungsoptima existieren, wurde mit einem numerischen Test belegt, vgl. die Abbildungen 3.1 und 3.2. Hier werden sowohl für die Tichonov- als auch die Prox-Regularisierung einige signifikante Niveaulinien von $f_i(\chi_i, r_i)$ bei optimaler Wahl von χ_i bzw. r_i geplottet. Um die Optimalität nachzuweisen mußten wir auf jedem Diskretisierungsniveau 40 mal 40 **QP**-Probleme mit unterschiedlich eingestellten Parametern lösen. Die Zahlen an den Niveaulinien zeigen, daß der Wert $f_i(\chi_i^{opt}, r_i^{opt})$ auf gröberen Gittern größer ist, als auf feineren. Natürlich existieren weitere Niveaulinien zu höheren Niveaus (für jedes h_i), aber diese bilden keine alternativen lokalen Minima.

Zur Lösung der **QP**-Probleme (3.75) benutzt man die Routine E04NAF aus der NAG-Bibliothek [74]. Das in der Routine verwendete Abstiegsverfahren beruht auf einer aktiven Mengenstrategie und erreicht das Minimum in endlich vielen Schritten. Falls die Kondition der auf die aktive Restriktionsmenge projizierten Hesse-Matrix sehr schlecht wird, erstattet der QP-Solver die Meldung, daß eine interne eigenständige Regularisierung eingeschaltet ist. Dies passiert vornehmlich, wenn der Rechenprozess zu feineren Gittern gelangt und dabei die Größen χ_i und r_i zu klein gewählt werden. Somit wird die ganze Parameterebene in zwei Teile zerlegt: mit zulässiger und unzulässiger Einstellung von χ_i und r_i. Diese Zerlegung ist auf den Abbildungen 3.1 und 3.2 für das 5-te und 6-te Gitter mit zwei schrafierten grünen Trennlinien dargestellt. Betrachtet man z.B. Abbildung 3.2: bei einer Wahl von $r_6 = 10^2$ und $\chi_6 = 10^{-8}$ arbeitet der Solver noch ohne eigenständige Regularisierung, wohingegen bei $r_6 = 1$ und $\chi_6 = 10^{-7}$ der Solver (trotz äußerer Regularisierung) das Problem als inkorrekt auffaßt und intern regularisiert. Bemerkenswert ist, daß der E04NAF-Solver das Fuller-Problem bis zum 4. Diskretisierungsniveau (d.h. $h_i = 2^{-4}$) ohne Regularisierung stabil löst.

Auf derselben Abbildung sieht man die jedem Gitter zugehörigen Minima des

Abstandes zwischen der mit E04NAF berechneten Lösung und der exakten Lösung des Originalproblems. Die Minima des Abstandes sind umso besser ausgeprägt, je größer der Diskretisierungsparameter h_i ist. Die optimalen Einstellungen der Verfahrensparameter, sowie die Größen $\|z^i - z^*\|_Z$ werden in folgenden Tabellen angeführt:

Tichonov-Regularisierung:

h_i	χ_i	r_i	NAG(i)	$\Psi_i(z^i)\ 10^3$	$\|\|z^i - z^*\|\|_Z$
2^{-2}	$2.7\ 10^{-4}$	316	3	0.7205	0.6763
2^{-3}	$4.4\ 10^{-6}$	100	4	0.8932	0.5054
2^{-4}	$2.3\ 10^{-6}$	50	2	0.9627	0.4207
2^{-5}	$5.2\ 10^{-8}$	0.32	6	1.0434	0.2916
2^{-6}	$2.0\ 10^{-8}$	10	6	1.0250	0.2731
2^{-7}	$2.5\ 10^{-7}$	10	5	1.0248	0.3396

Prox-Regularisierung (OSR):

h_i	χ_i	r_i	NAG(i)	$\Psi_i(z^i)\ 10^3$	$\|\|z^i - z^*\|\|_Z$
2^{-2}	$2.7\ 10^{-4}$	316	3	0.7205	0.6762
2^{-3}	$2.3\ 10^{-6}$	100	3	0.8926	0.5060
2^{-4}	$6.0\ 10^{-6}$	32	2	0.9910	0.4213
2^{-5}	$7.0\ 10^{-8}$	1.59	6	1.0409	0.2919
2^{-6}	$2.0\ 10^{-8}$	10	5	1.0250	0.2673
2^{-7}	$1.6\ 10^{-6}$	1.66	4	1.0406	0.2416

Der Haupteffekt in diesem Test besteht darin, daß das T-Verfahren auf feineren Gittern in der Praxis divergiert. Ab einem bestimmten Diskretisierungsniveau kann man nämlich die Regularisierungs- und Strafparameter nicht mehr verkleinern, ohne die den zulässigen Einstellungsbereich absteckende Grenze zu überschreiten. Dieselbe Schwierigkeit entsteht offensichtlich mit dem OSR-Verfahren. Dennoch verhält sich das proximal-regularisierte Verfahren wesentlich flexibler auf feineren Gittern mit wieder größer werdenden χ_i und r_i. Dort liegt weiterhin Konvergenz gegen den Optimalprozess vor. Bemerkenswert ist, daß die Werte des Hilfsfunktionals $\Psi_i(\cdot)$ gegen den Optimalwert J_0^* von unten konvergieren. Es ist scheinbar ein typisches Verhalten für die von uns betrachteten Strafverfahren. In der vierten Spalte beider Tabellen wird die Anzahl der auf dem gegebenen Diskretisierungsniveau vom NAG-Solver gewechselten Restiktionsmengen angeführt. Die Schreibweise $\Psi_i(z^i)\ 10^3$ in der 5. Spalte bedeutet, daß die zugehörigen Werte mit 10^{-3} zu multiplizieren sind. Der kleine Optimalwert des Zielfunktionals ist die Ursache für die relativ kleinen Werte des Regularisierungsparameters χ_i bzw. der großen Strafparameterwerte r_i.

□

Test 2: Anhand der optimalen Einstellungen aus der letzten Tabelle und der exakten Lösung des Fuller–Problems können wir die wichtigste Frage dieses Abschnitts untersuchen: **was kann man mit der Mehr-Schritt-Regularisierung erreichen?** Wir starten den Rechenprozess auf jedem Gitter genau wie beim OSR-Verfahren und erlauben einen zusätzlichen Proximal-Schritt auf demselben Gitter. Zugleich messen wir den Abstand $\|z^{i,2} - z^*\|_Z$ und vergleichen ihn mit $\|z^{i,1} - z^*\|_Z$. Unerwarteterweise ist der Abstand nach dem zweiten Proximal-Schritt größer, als nach dem ersten und das gilt für alle groben Gitter, die man testet. Das heißt, das MSR-Verfahren konvergiert auf groben Diskretisierungsniveaus "zu tief"; die Näherung entfernt sich mit jeder s-Iteration immer weiter von der richtigen Lösung, was dann die spätere Konvergenz auf feineren Gittern negativ beeinflußt.

Ein positiver Effekt kann jedoch entstehen, wenn der Rechenprozess zu feineren Gittern gelangt. Auf dem 6-ten Gitter erbringen die ersten sechs Proximal-Schritte eine gewünschte Verringerung des Fehlers. Auf dem 7-ten Gitter sind das schon 84 Proximal-Schritte, nach denen sich das MSR-Verfahren wieder divergent verhält.

□

Die Konvergenz von T-, OSR- und MSR-Verfahren auf Folgen von Gittern ist in Abbildung 3.3 dargestellt. Die exakte Lösung $u^*(x)$ ist auf derselben Graphik mit grün gezeichnet. Es ist bemerkenswert, daß sich die Näherungen auf den groben Gittern sehr wenig unterscheiden. Das Endergebnis aus dem 7-ten Gitter ist in Abbidung 3.4 angeführt. Man kann sehen, daß die beste Qualität der Reprodunktion einer instabilen Lösung $u^*(x)$ im Falle des MSR-Verfahrens vorliegt.

Wenn man die Strafverfahren ohne Regularisierung startet, so kann man keine Konvergenz beobachten. Unterschiedliche Varianten einer derartigen Divergenz auf dem feinsten Diskretisierungsniveau sieht man auf der Seite 78. Es werden dort unterschiedliche Startpunkte und Strafparameterwerte betrachtet.

3.3 Ein inkorrekt gestelltes Steuerproblem mit einem elliptischen Operator 1-ter Ordnung

Im Abschnitt 3.2.3 wurde gezeigt, daß die inkorrekt gestellten Optimalsteuerprobleme mit elliptischer Zustandsgleichung 2-ter Ordnung wesentliche Schwierigkeiten bei ihrer numerischen Behandlung hervorrufen. Besonders wenn die Berechnung auf feineren Gittern erfolgt. Falls der Diskretisierungsparameter h *einen kritischen Wert* $h < h^*$ unterschreitet, ist zu erwarten, daß man ohne Regularisierung nicht rechnen kann; nur im Falle eines positiven Regularisierungsparameters $\chi > 0$ ist ein stabiler Rechenprozess möglich. Das Ziel des Kapitels 3 besteht darin, die prinzipielle Frage zu beleuchten, wie hängt der charakteristische Wert h^* mit den Eigenschaften des Diskretisierungsschemas bzw. des Originalproblems zusammen?

Die in den diskretisierten Ersatzproblemen entstehenden Massen- und Steifigkeitsmatrizen sind von der Wahl der Basisfunktionen der Finite-Element-Technik

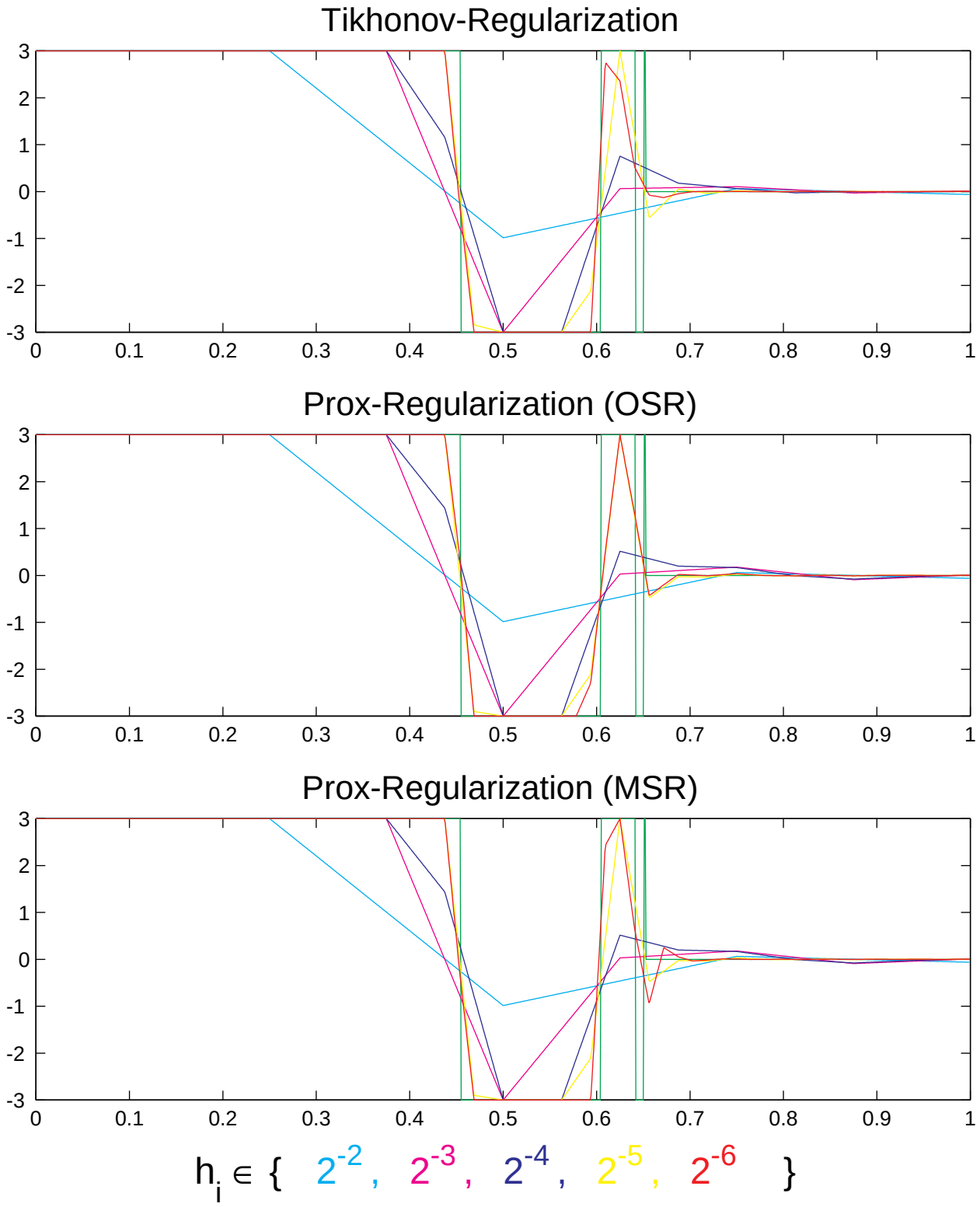


Abbildung 3.3: Konvergenz auf Folgen von Gittern

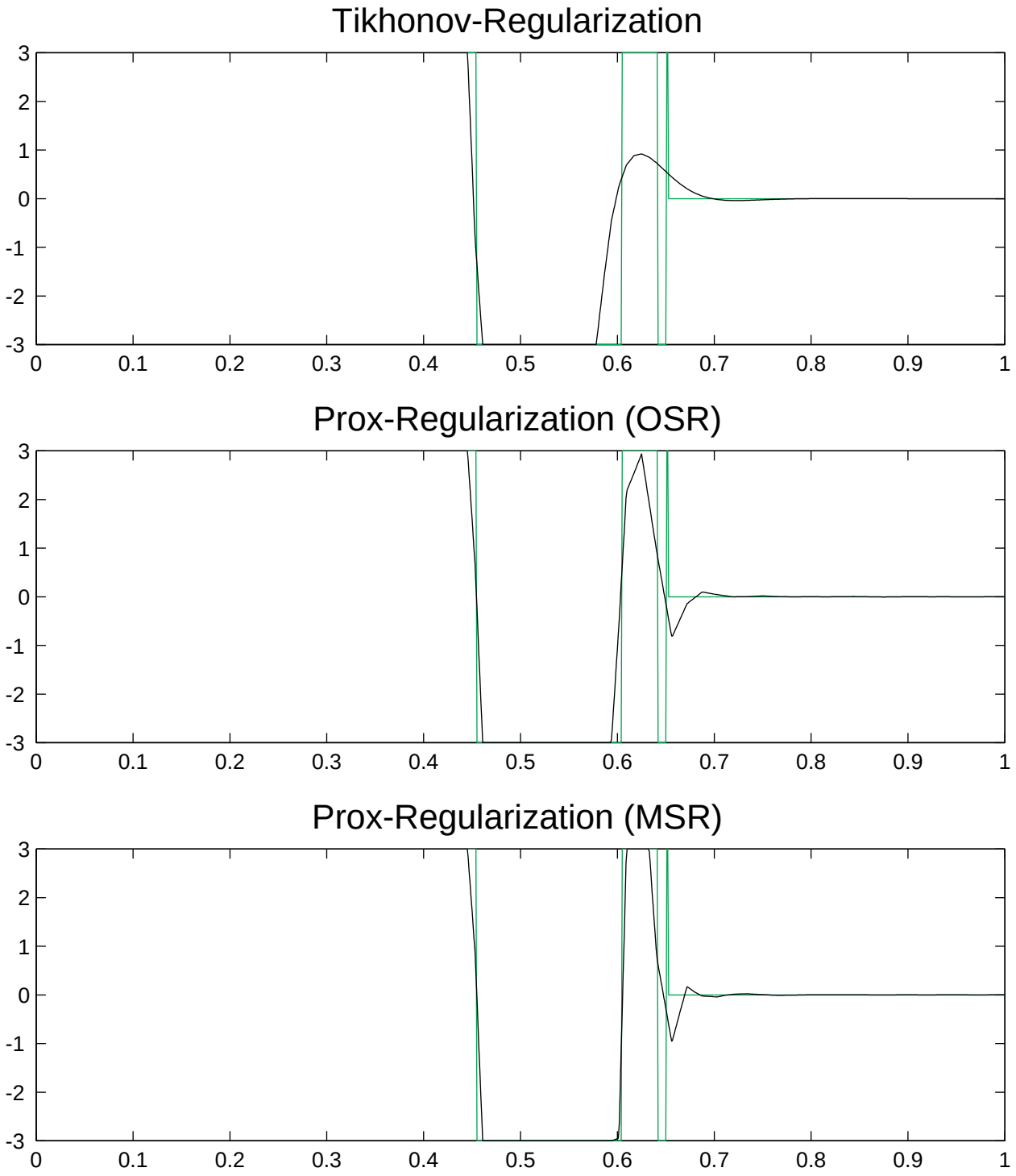


Abbildung 3.4: Vergleich der Regularisierungsverfahren, $h = 2^{-7}$

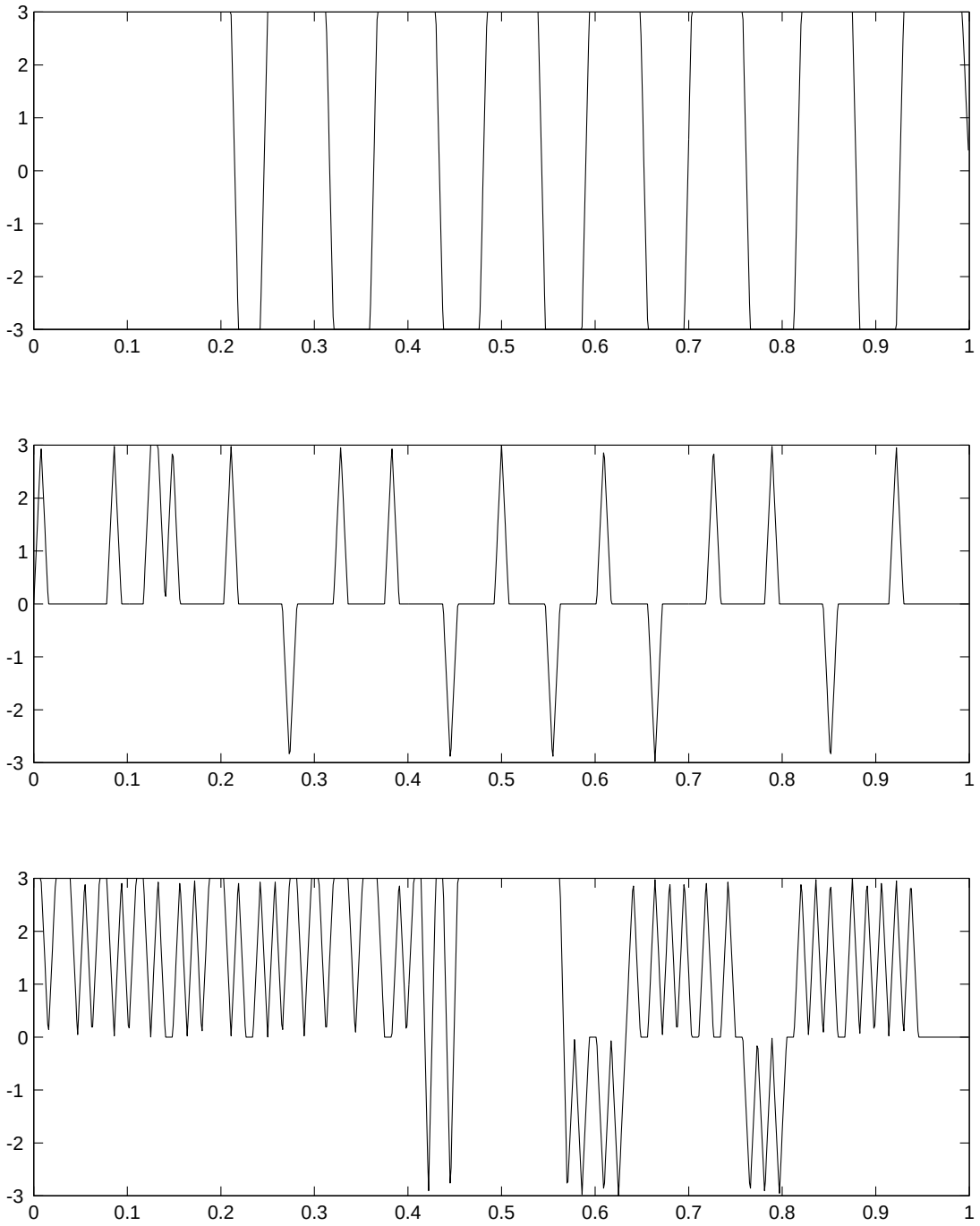


Abbildung 3.5: Divergenz des Strafverfahrens ohne die Regularisierung

abhängig. Wie diese Wahl den kritischen Wert h^* im Falle inkorrekt gestellter Steuerprobleme beeinflußt, bleibt im allgemeinen eine offene Frage. Es ist lediglich bekannt, daß Basiselemente, die durch Polynome höherer Ordnung erzeugt werden, zu Matrizen mit schlechteren Spektraleigenschaften führen. Diesem Gedanken folgend, muß die Wahl linearer Elemente immer wünschenswerter sein, als die der Hermetischen Splines. Diese Überlegung hat uns gezwungen, in Kapitel 4 auf die Straftechniken zu verzichten, um mit günstigeren Basisfunktionen operieren zu können. Es sei nochmal daran erinnert, daß im Falle der Straftechnik die approximative Lösung in $W^2 \times L_2$ gesucht werden muß, während das zu minimierende Hilfsfunktional in Kapitel 4 einen größeren Definitionsbereich besitzen kann.

Die prinzipielle Abhängigkeit zwischen h^* und den Eigenschaften der Operatoren A, B, C des Originalproblems (**P** oder **P**′) läßt sich durch folgende Hypothese ausdrücken: **bei fixierten Operatoren B und C muß der kritische Wert h^* des Diskretisierungsparameters um so kleiner werden, je kleiner die Ordnung des Operators A ist.** Außerdem ist denkbar, daß eine solche Auswahl der Operatoren A und C existieren kann, bei der das entsprechende Steuerproblem seine Korrektheit wiedererhält, obwohl keine quadratischen Steuerkosten vorhanden sind [57, Kap. 2, § 3]. Es kann auch vorkommen, daß der dem Originalproblem entsprechende kritische Wert h^* so gering ist, daß die erfolgreiche Berechnung ohne Regularisierung bis zu einem sehr feinen Gitter möglich ist. Es sei hierbei auf die numerischen Experimente in [68] hingewiesen. In solchen Fällen werden wir von Steuerproblemen mit *schwach ausgeprägter Inkorrektheit* sprechen.

3.3.1 Beobachtung und Steuerung auf gleichem Randteil

Um ein Optimalsteuerproblem mit schwach ausgeprägter Inkorrektheit zu konstruieren, müssen wir entsprechend unserer Hypothese die Ordnung des elliptischen Systemoperators A verkleinern. Der Operator des Zielfunktionals C sei wieder von der Ordnung Null. Wir lösen die Aufgabe der Ordnungsverkleinerung durch die Technik der Poincaré-Steklov Operatoren aus Abschnitt 3.1.3. Zunächst betrachten wir ein angewandtes Kontrollproblem auf einem Quadratgebiet (Abb. 3.6).

Man betrachte eine Meeresküste mit der Strömung des Wassers entlang des Strandes, dessen Länge als Eins angenommen wird. Die Geschwindigkeit der Strömung v sei von links nach rechts gerichtet. Auf dem linken Teil des Strandes befinde sich eine Verschmutzungsquelle, deren Länge gleich $1/2$ sei und die eine gleichmäßige Verschmutzung mit der Intensität g erzeugt. Der Schmutz diffundiert in das Wasser und wird zugleich durch die Strömung weggetragen. Wir führen ein Koordinatensystem ein, wie auf der Abbildung 3.6 dargestellt, und bezeichnen mit $\sigma(x)$ die Konzentration der Verschmutzung im Wasser. Da die Verbreitung des Schmutzes einen quasistationären Prozess im Meer bildet, erfüllt die Funktion σ die stationäre Diffusionsgleichung mit einem Konvektionsglied, das für die Strömung verantwort-

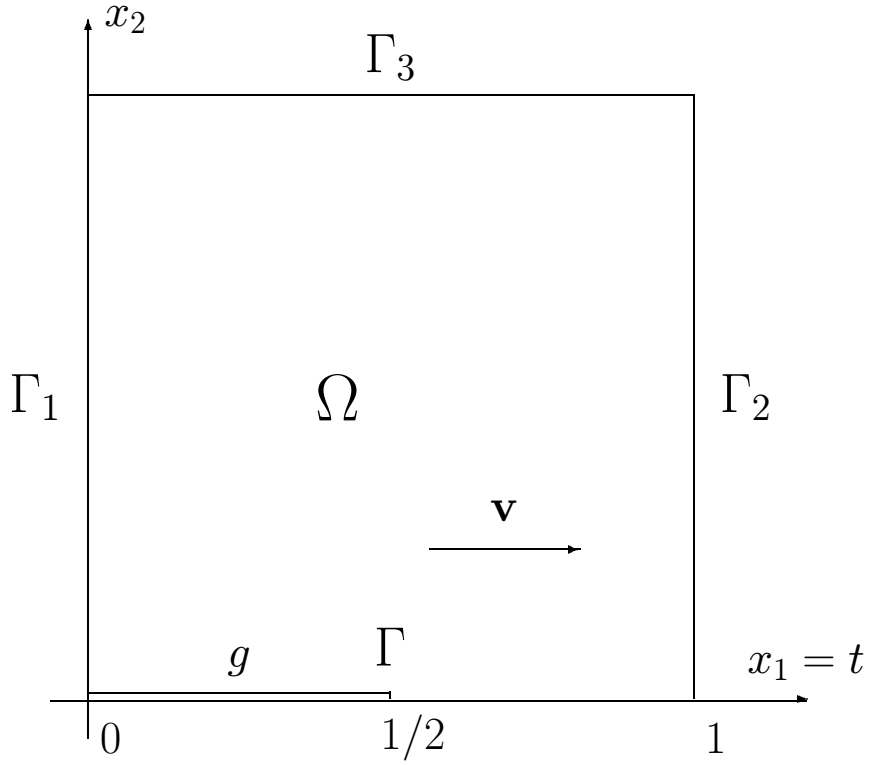


Abbildung 3.6: Die Verschmutzungsquelle g auf dem Rand des Gebiets

lich ist:

$$\begin{cases} -\Delta\,\sigma(x) + 2v\,\dfrac{\partial\sigma}{\partial x_1}(x) = 0 & \text{in } \Omega\,, \\ \sigma(x) = 0 & \text{auf } \Gamma_1 \cap \Gamma_2\,, \\ \dfrac{\partial\sigma}{\partial x_2}(x) = 0 & \text{auf } \Gamma_3\,, \\ -\dfrac{\partial\sigma}{\partial x_2}(x) = u(x) + g(x) & \text{auf } \Gamma\,. \end{cases} \tag{3.76}$$

Hierbei bezeichnet Δ einen Laplace-Operator ($\Delta = \partial^2/\partial x_1^2 + \partial^2/\partial x_2^2$). Außer der Funktion g, welche die Verschmutzungsquelle charakterisiert, findet man im System (3.76) die Steuerung u. Diese Funktion beschreibe die Reinigungsmaßnahmen, welche man am restlichen Teil der Küste durchzuführen hat. Der Träger von u sei also $[1/2, 1]$. Durch die Reinigungsmaßnahmen braucht man nur die Konzentration der Verschmutzung σ an der Küste, aber nicht im ganzen Meer auf das "Null-Niveau" zu bringen; das heißt, man muß das Meer nicht stärker als notwendig reinigen. Außerdem sind die Reinigungsquellen in ihrer Leistung begrenzt. Dies ergibt punktweise Beschränkungen auf die Amplitude der Steuerfunktion u.

Um das erklärte Steuerproblem in der Form des Problems **P** aus Kapitel 2 zu formulieren, muß man eine neue Art der Zustandsgleichung aufbauen. Als System-

operator nimmt man den Poincaré-Operator P an, welcher die lineare Abhängigkeit $P : \gamma\sigma \to \beta\sigma$, zwischen der Spur und der äußeren normalen Ableitung der Funktion $\sigma(x)$ auf dem Randteil Γ bestimmt. Gemäß dem Abschnitt 3.1.3 heißen γ und $\beta = -\gamma D_{x_2}$ Dirichlet- und Neumann-Operatoren. Man kann die Graphen dieser linearen Operatoren so fortsetzen, daß die Bedingung

$$\gamma \in \mathcal{L}(H^1(\Omega),\, H_{00}^{1/2}(\Gamma))\,, \qquad \delta \in \mathcal{L}(H^1(\Omega),\, H_{00}^{-1/2}(\Gamma))$$

erfüllt ist. Wir bezeichnen mit $y(t)$, $t \in (0,\, 1)$ die Spur von $\sigma(x)$ auf Γ und behandeln diese Funktion einer Variablen als unsere neue Zustandsfunktion. Das zu lösende Problem läßt sich wie folgt formulieren:

$$\begin{aligned} J(y,\, u) = \frac{1}{2}\,\|y\|^2_{(0,\,1)} &\xrightarrow{\;!\;} \min_{(y,\,u)\,\in\,\mathcal{Y}\times\mathcal{U}_{ad}} \\ P\,y = u + g \qquad &\text{in}\;\; H_{00}^{-1/2}(\Gamma)\,. \end{aligned} \tag{3.77}$$

In (3.77) gilt $P \in \mathcal{L}(H_{00}^{1/2}(\Gamma),\, H_{00}^{-1/2}(\Gamma))$. Nun müssen wir kurz vorführen, warum P ein elliptischer Operator erster Ordnung ist und wie bequem die Sobolev-Lions-Räume sein können.

Wir beginnen mit der Diskussion der fundamentalen Funktionen, Eigenfunktionen und Eigenwerte von P entsprechend Abschnitt 3.1.3. Man kann leicht prüfen, daß die Daten

$$v_k(x_1,\, x_2) = \sqrt{2}\, e^{vx_1} \sin \pi k x_1\, \frac{\cosh \mu_k (1 - x_2)}{\cosh \mu_k}\;; \tag{3.78}$$

$$e_k(t) = \sqrt{2}\, e^{vt} \sin \pi k t\,, \quad \hat{e}_k(t) = \sqrt{2}\, e^{-vt} \sin \pi k t\;; \quad \lambda_k = \mu_k\, \tanh \mu_k \tag{3.79}$$

mit

$$\mu_k \;\; = \;\; \sqrt{v^2 + (\pi k)^2} \tag{3.80}$$

sämtliche Bedingungen in (3.24) und (3.25) für das zu betrachtende Problem (3.76) erfüllen. In Abschnitt 3.1.2 findet man auch ein Beispiel für den Operator des kanonischen Isomorphismus im Raum $\mathring{W}^1(\Gamma)$:

$$\Lambda_1\, y(t) = \, - D_t^2\, y(t) + y(t)\,.$$

Durch diesen Operator wird die Norm des Raumes $H_{00}^{1/2}(\Gamma)$ erzeugt. Desweiteren benötigen wir noch folgende Beziehungen:

$$p \geq \pi \;\Rightarrow\; 2p > 1 + \sqrt{p^2 + 1}\;, \quad p\, \tanh p < 1 + \sqrt{p^2 + 1}\;. \tag{3.81}$$

Wir betrachten den einfachsten Fall $v = 0$, d.h. $\hat{e}_k(t) = e_k(t) = \sqrt{2}\,\sin \pi k t$. Die Eigenfunktionen von P und Λ_1 fallen dann zusammen. Entsprechend Abschnitt 3.1.2 und (3.81) für $y \in H_{00}^{1/2}(\Gamma)$ gilt

$$a_P(y,\, y) = \sum_{k=1}^{\infty} \pi k\, \tanh \pi k\, (e_k,\, y)_\Gamma^2 \;\; > \;\;$$

$$\sum_{k=1}^{\infty} \frac{\tanh \pi}{2}\, \left(1 + \sqrt{(\pi k)^2 + 1}\right)\, (e_k,\, y)_\Gamma^2 \;\; = \;\;$$

$$\frac{\tanh \pi}{2} \int\limits_0^1 \left(\sum_{k=1}^{\infty} \left((\pi k)^2 + 1 \right)^{1/4}\, (e_k,\, y)_\Gamma\, e_k(t) \right)^2 dt \;\; + \;\;$$

$$\frac{\tanh \pi}{2} \int\limits_0^1 \left(\sum_{k=1}^{\infty} (e_k,\, y)_\Gamma\, e_k(t) \right)^2 dt = \frac{\tanh \pi}{2}\, \left(\left\| \Lambda_1^{1/4} y \right\|_\Gamma^2 + \|y\|_\Gamma^2 \right) \quad .$$

Aus der Cauchy-Schwarz Ungleichung und der letzten Ungleichung in (3.81) folgt

$$a_P(y_1,\, y_2) \le \sqrt{a_P(y_1,\, y_1)\, a_P(y_2,\, y_2)} \;\; = \;\;$$

$$\left(\sum_{k=1}^{\infty} \pi k\, \tanh \pi k\, (e_k,\, y_1)_\Gamma^2 \right)^{1/2} \left(\sum_{k=1}^{\infty} \pi k\, \tanh \pi k\, (e_k,\, y_2)_\Gamma^2 \right)^{1/2} \;\; < \;\;$$

$$\left(\sum_{k=1}^{\infty} \left(1 + \sqrt{(\pi k)^2 + 1}\right)\, (e_k,\, y_1)_\Gamma^2 \right)^{1/2} \;\; \times \;\;$$

$$\left(\sum_{k=1}^{\infty} \left(1 + \sqrt{(\pi k)^2 + 1}\right)\, (e_k,\, y_2)_\Gamma^2 \right)^{1/2} \;\; = \;\;$$

$$\left(\left\| \Lambda_1^{1/4} y_1 \right\|_\Gamma^2 + \|y_1\|_\Gamma^2 \right)^{1/2} \left(\left\| \Lambda_1^{1/4} y_2 \right\|_\Gamma^2 + \|y_2\|_\Gamma^2 \right)^{1/2} \quad .$$

Somit ist der Operator P elliptisch mit der Ordnung $s = 1$.

Gemäß dem allgemeinen Schema aus Kapitel 2 muß man in (3.77) folgende Daten wählen: $\mathcal{Y} = \mathring{W}^1(\Gamma)$, $\mathcal{U} = L_2(\Gamma)$, $\mathcal{K} = L_2(\Gamma)$ und

$$\mathcal{U}_{ad} = \{u \in \mathcal{U} :\; u(x) = 0 \;\text{ f.ü. in }\; (0,\, 1/2)\;,\;\; |u(x)| \le 1 \;\text{ f.ü. in }\; (1/2,\, 1)\}\;;$$

der Operator des Zielfunktionals C stellt dabei wieder einen Identitätsoperator dar.

Weil $g \in L_2(\Gamma)$ gilt, kann man die in den vorigen Abschnitten eingeführten und untersuchten Strafverfahren erfolgreich anwenden. Als Kopplungsraum $\mathcal{H}$ in (3.77) muß dann L_2 statt $H_{00}^{-1/2}$ betrachtet werden. Außerdem muß noch die

Existenz beider Lipschitz-Konstanten für den eingeschränkten Poincaré-Operator $P \in \mathcal{L}(\mathring{W}^1(\Gamma),\, L_2(\Gamma))$ bewiesen werden. Dazu muß man die Struktur dieses Operators erforschen.

Satz 12 *Für beliebiges $y \in \mathring{W}^1(\Gamma)$ ergibt sich die explizite Form des Poincaré-Operators P folgendermaßen*

$$Py\,(t) = \int\limits_0^1 \Big(v^2\, K_0(t,\tau)\, y(\tau) + 2v\, K_2(t,\tau)\, D_\tau\, y(\tau) - K_2(t,\tau)\, D_\tau^2\, y(\tau) \Big)\; d\tau \;, \tag{3.82}$$

wobei

$$K_0(t,\tau) \;=\; e^{v\,(t-\tau)} \sum_{k=1}^{\infty} \frac{2\,\tanh \mu_k}{\mu_k} \cdot \sin \pi k t\;\, \sin \pi k \tau \;, \tag{3.83}$$

$$K_2(t,\tau) \;=\; e^{v\,(t-\tau)} \sum_{k=1}^{\infty} \frac{2(\pi k)^2\,\tanh \mu_k}{\mu_k^3} \cdot \sin \pi k t\;\, \sin \pi k \tau \tag{3.84}$$

und D_t die verallgemeinerte Ableitung bezeichnet.

Beweis: Die Summen in (3.83) und (3.84) konvergieren für $(t,\,\tau) \in (0,\,1) \times (0,\,1) \setminus \{(t,\,\tau) :\;\; t = \tau\,\}$. Die Kerne $K_0(t,\tau)$ und $K_2(t,\tau)$ in den zu betrachtenden Integraloperatoren sind schwach singulär [32] und besitzen logarithmische Asymptotik, falls $|t-\tau| \to 0$. Tatsächlich, wenn man beispielsweise die Parametrisierung der Variablen $\tau(\varepsilon) = (\sqrt{2} - \varepsilon)/2,\;\; t(\varepsilon) = (\sqrt{2} + \varepsilon)/2$ ausführt, so erhält man unter den unendlichen Summen die Ansätze

$$\frac{\cos \pi k \varepsilon}{\mu_k} \qquad \text{bzw.} \qquad \frac{(\pi k)^2 \cos \pi k \varepsilon}{\mu_k^3} \;.$$

Andererseits [47, § 55.15] hat man für $0 < \varepsilon < 2\pi$ die Beziehung

$$\sum_{k=1}^{\infty} \frac{\cos \pi k \varepsilon}{k} \;=\; -\ln\left(2 \sin \frac{\varepsilon}{2}\right) \;,$$

welche die oben erwähnte Asymptotik festlegt. Um die explizite Darstellung (3.82) nachzuweisen, untersuchen wir die Wirkung der dort stehenden Operatoren auf eine beliebige Eigenfunktion $e_l(x)$ des Poincaré-Operators. Weil für hinreichend glatte Funktionen verallgemeinerte und klassische Ableitungen zusammenfallen, gilt folgendes

$$D_\tau\; (e^{v\tau} \sin \pi l \tau) \;=\; v\, e^{v\tau} \sin \pi l \tau + \pi l\, e^{v\tau} \cos \pi l \tau \;,$$

$$D_\tau^2\; (e^{v\tau} \sin \pi l \tau) \;=\; (v^2 - (\pi l)^2)\, e^{v\tau} \sin \pi l \tau + 2v \pi l\, e^{v\tau} \cos \pi l \tau \;.$$

Aufgrund der Orthogonalität der trigonometrischen Polynome kann man die Kosinus-Glieder eliminieren. Entsprechend des Operators auf der rechten Seite von (3.82) erhält man

$$v^2\mu_l^2 + 2v^2(\pi l)^2 - (\pi l)^2(v^2 - (\pi l)^2) = \mu_l^4 \ ,$$

was zuletzt die Definition von P durch seine Spektralzerlegung ergibt.

□

Aus Satz 12 kann man sehen, daß der Poincaré-Operator P in $\mathcal{L}(L_2(0,\,1))$ unsymmetrisch ist, solange $v > 0$. **Aber die in Kapitel 2 diskutierten Strafverfahren sind auf Symmetrieeingenschaften des Systemoperators nicht angewiesen.** Sie benötigen lediglich die Existenz beider Lipschitz-Konstanten für diesen Operator in den vorgegebenen Sobolevräumen.

Satz 13 *Für den Operator P und beliebiges $y \in \mathring{W}^1(\Gamma)$ gilt*

$$m_P \, \|y\|_{W^1(\Gamma)} \leq \| \, Py \, \|_\Gamma \leq M_P \, \|y\|_{W^1(\Gamma)}$$

mit den Lipschitz-Konstanten

$$m_P = \frac{e^{-v}}{\sqrt{2}} \tanh \sqrt{\pi^2 + v^2} \ , \qquad M_P = e^v \ . \tag{3.85}$$

Beweis: Wir betrachten zunächst eine Folge unendlich glatter Funktionen $y_n(t)$, $n = 1, 2, ..,$ welche gegen die zu untersuchende Funktion $y(t)$ in $W^1(\Gamma)$ konvergiert. Die Existenz einer solchen Folge läßt sich aus der Dichtheit des Raumes $C^\infty(\overline{\Gamma})$ in $W^1(\Gamma)$ ersehen [1]. Aus der Gleichheit

$$e^{-vt} \sin \pi k t = \ - \left(\frac{e^{-vt}}{\mu_k^2} \left(\pi k \cos \pi k t + v \sin \pi k t \right) \right)'$$

folgt nach der partiellen Integration

$$\int_0^1 e^{-2vt} (P\, y_n)^2(t) \, dt \ =$$

$$\int_0^1 \left(\sum_{k=1}^{\infty} \frac{2\pi k \tanh \mu_k}{\mu_k} \sin \pi k t \int_0^1 y_n'(\tau_1) \, e^{-v\tau_1} \left(\cos \pi k \tau_1 + \frac{v}{\pi k} \sin \pi k \tau_1 \right) d\tau_1 \right) \ \times$$

$$\left(\sum_{l=1}^{\infty} \frac{2\pi l \tanh \mu_l}{\mu_l} \sin \pi l t \int_0^1 y_n'(\tau_2) \, e^{-v\tau_2} \left(\cos \pi l \tau_2 + \frac{v}{\pi l} \sin \pi l \tau_2 \right) d\tau_2 \right) \, dt \ =$$

$$\int_0^1 e^{-2vt} y_n'(t) \, (Q\, y_n')(t) \, dt$$

mit dem neuen Operator

$$Qu\,(t) = \sum_{k=1}^{\infty} \frac{2(\pi k)^2 \tanh^2 \mu_k}{\mu_k^2}\, e^{vt} \left(\cos \pi k t + \frac{v}{\pi k} \sin \pi k t\right) \quad \times$$

$$\int_0^1 u(\tau)\, e^{-v\tau} \left(\cos \pi k \tau + \frac{v}{\pi k} \sin \pi k \tau\right) d\tau \quad .$$

Wir führen eine $\mathcal{W}$-Norm im Raum $L_2(\Gamma)$ ein:

$$\|u\|_{\mathcal{W}} = \left(\int_0^1 e^{-2vt} u^2(t)\, dt \right)^{1/2}$$

Man kann vermerken, daß Q in $\mathcal{L}(\mathcal{W})$ symmetrisch ist und ein $\mathcal{W}$-orthogonales System der Eigenfunktionen $\{h_k(t)\}_{k=1}^{\infty}$, sowie die Eigenwerte $\{\varsigma_k\}_{k=1}^{\infty}$ besitzt. Es gilt

$$h_k(t) = e^{vt} \left(\cos \pi k t + \frac{v}{\pi k} \sin \pi k t\right) , \qquad \varsigma_k = \tanh^2 \mu_k \; .$$

Gemäß [60, Kap. 7, § 4, Satz 3] ergeben die Spektralgrenzen die untere und obere Abschätzung der zu betrachtenden quadratischen Form:

$$\varsigma_1 \, \|y_n'\|_{\mathcal{W}}^2 \le \int_0^1 e^{-2vt} (P\, y_n)^2(t)\, dt \le \|y_n'\|_{\mathcal{W}}^2 \; . \tag{3.86}$$

Zum Abschluß muß man vermerken, daß $\|y_n\|_\Gamma \le \|y_n'\|_\Gamma$ gilt und daß die Exponentialfunktion auf Γ positiv ist. Somit folgt (3.85) bei nahe direkt aus (3.86).

□

Die explizite Form (3.82) veranschaulicht nochmal, warum der Poincaré-Operator P von erster Ordnung ist. Im letzten Term aus (3.82) wird eine zweite Ableitung genommen, aber dann erfolgt ein rückwärtiges Glätten der Distribution, weil der logarithmische Kern $K_2(t,\, \tau)$ eine sehr schwache Singularität besitzt. Satz 13 zeigt deutlich, daß beide Lipschitz-Konstanten sehr schnell auseinander laufen, falls $v \to \infty$. Es ist ein typisches Verhalten des singulär gestörten Operators: er wird in den vorgegebenen Räumen nicht beschränkt.

3.3.2 Numerische Tests

Die Ergebnisse im Abschnitt 3.3.1 erlauben nun eine Anwendung der regularisierten Strafverfahren aus Kapitel 2 zur Lösung des Problems (3.77). Wir beginnen die Untersuchung mit dem T-Verfahren ($t_d = 0$) und zeigen durch einen numerischen

Test, daß man in diesem Verfahren den Regularisierungsterm vollständig ausschalten kann ($\chi_i = 0,\ \forall\, i$), ohne daß die Ersatzprobleme ihre Stabilität auf feineren Gittern verlieren. Zur Diskretisierung der Ersatzprobleme scheint es vernünftig, die Eigenfunktionen des Poincaré-Operators (3.79) zu verwenden. Somit betrachtet man die Folge von Problemen:

$$\mathbf{P}_i^h: \qquad \frac{1}{2}\,\|y_h\|^2_{(0,\,1)} + \frac{1}{2r_i}\,\|P\,y_h - u_h - g\|^2_{(0,\,1)} \stackrel{!}{\longrightarrow} \min_{(y_h,\,u_h)\,\in\,\mathcal{Y}^h\times\mathcal{U}^h_{ad}}\,, \tag{3.87}$$

wobei die inneren Approximationen wie folgt definiert sind:

$$\mathcal{Y}^h = \left\{ y_h \in \mathring{W}^1(\Omega):\ y_h(x) = \sum_{l=1}^{L} y_l\, e_l(x),\ y_l \in \mathbf{R} \right\}\,,$$

$$\mathcal{U}^h = \left\{ u_h \in L_2(\Omega):\ u_h(x) = \sum_{j=1}^{J} u_j\, \psi_j(x),\ u_j \in \mathbf{R} \right\}\,.$$

Die Menge $\mathcal{U}_{ad}$ diskretisiert man folgendermaßen:

$$\begin{aligned} \mathcal{U}^h_{ad} = \{u_h \in \mathcal{U}^h:\ \ & u_h(x) = \ \ 0 \ \text{ f.ü. in } (0,\,1/2), \\ & -1 \le \ u_h(x) \le \ \ 1 \ \text{ f.ü. in } (1/2,\,1)\}\,. \end{aligned}$$

Bei solcher Art der Diskretisierung begegnet man einer prinzipiell neuen Situation. Im Gegensatz zu Abschnitt 3.2.3 gilt nun mit einer konventionellen Bezeichnung $\mathcal{U}^h \not\subseteq P\,\mathcal{Y}^h$. Die große Lücke zwischen den Räumen kann bei einem kleinen Strafparameterwert r_i die sensible Approximation der Steuerung zerstören. Um diese Schwierigkeit zu vermeiden, muß man die Anzahl der zur Diskretisierung verwendeten Eigenfunktionen L im Vergleich mit J wesentlich größer wählen. (Im konkret berechneten Fall wurden auf dem letzten Diskretisierungsniveau $L = 2048$ und $J = 256$ gewählt). Schließlich erhält man folgendes Problem der quadratischen Programmierung:

$$\mathbf{QP}(r_i): \qquad \hat{\Psi}(\mathbf{z}) = \frac{1}{2}\,\mathbf{z}^{\mathrm{T}}\,H\,\mathbf{z} - \mathbf{f}^{\mathrm{T}}\,\mathbf{z} \ \stackrel{!}{\longrightarrow} \min_{\mathbf{z}\,\in\,\hat{\mathcal{Z}}_{ad}} \tag{3.88}$$

mit

$$H = \begin{bmatrix} A_y & A_r \\ A_r^{\mathrm{T}} & A_u \end{bmatrix}\,, \qquad \begin{aligned} \mathbf{z} &= (\mathbf{y},\,\mathbf{u})^{\mathrm{T}}\,, \\ \mathbf{f} &= (\mathbf{f}_y,\,\mathbf{f}_u)^{\mathrm{T}}\,. \end{aligned}$$

Der Inhalt der Matrizen und Vektoren läßt sich wie folgt beschreiben:

$$(A_y)_{l'l} = \left((\mu_{l'} \tanh \mu_{l'})^{-1} (\mu_l \tanh \mu_l)^{-1} + \frac{1}{r_i} \right) (e_{l'},\, e_l)_{(0,\,1)}\,,$$

$$(A_r)_{l'j} = -\,\frac{1}{r_i}\,(e_{l'},\, \psi_j)_{(0,\,1)}\,,$$

$$(A_u)_{j'j} = \left(\frac{1}{r_i} + \chi_i\right) (\psi_{j'}, \psi_j)_{(0,1)} ;$$

$$(\mathbf{f}_y)_l = \frac{1}{r_i} (e_l, g)_{(0,1)} ,$$

$$(\mathbf{f}_u)_j = \frac{1}{r_i} (\psi_j, g)_{(0,1)} .$$

Die Massenmatrix $\left[(\psi_{j'}, \psi_j)_{(0,1)}\right]$ für die stückweise linearen Funktionen besitzt eine gut bekannte Struktur [89]. Die Elemente der Massenmatrix zu den Eigenfunktionen $e_k(t)$ lassen sich auch explizit berechnen:

$$(e_k, e_l)_{(0,1)} = \frac{v\pi^2 kl \left((-1)^{k-l} e^{2v} - 1\right)}{4\,\mu^2_{(k-l)/2}\,\mu^2_{(k+l)/2}}$$

Eine graphische Darstellung der Elementwerte in der Matrix $\left[(e_l, \psi_j)_{(0,1)}\right]$ ist auf der Seite 89 angeführt. Es sei vermerkt, daß man nun mit dicht besetzten Matrizen arbeiten muß. Jedoch entspricht die Dimension des Zustandsvektors $\mathbf{z}$ der Anzahl der Knoten auf dem Rand Γ, welche wesentlich geringer ist, als die gesammte Anzahl der Knoten bei einer eventuellen Diskretisierung des Gebiets Ω. Man vergesse nicht, daß es sich in (3.77) um ein zweidimensionales Steuerproblem handelt.

Test: Die Folge der gekoppelten **QP**-Probleme (3.88) wird nur anhand des Strafparameters r_i gesteuert, dessen Werte aus folgender Tabelle zu entnehmen sind:

h_i	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}
r_i	20	10	5	1	0.5	0.1

Die Konvergenz des unregularisierten Strafverfahrens auf Folgen von Gittern wird in Abbildung 3.7 angeführt. Man erkennt wieder zwei Zonen der Steuerung. Links befindet sich eine einfache Bang-Steuerung (es ist kein "Chattering" mehr) und rechts liegt eine nicht nullidentische singuläre Steuerung vor, welche nur im singulären Punkt $x = x_\infty$ und im Punkt $x = 1$ verschwindet. Im Punkt $x = 1$ ist außerdem die Zustandsfunktion gleich Null. Die interessante Eigenschaft jedes nicht trivialen Optimalprozesses besteht darin, daß das Maß derjenigen Mengen, auf denen beide Funktionen $u^*(x)$ und $y^*(x)$ verschwinden, immer gleich Null ist [86, Satz 2.1].

Auf der unteren Graphik ist die Wurzel aus der adjungierten Zustandsfunktion $\sqrt{p(x)}$ angeführt. Aufgrund der gut erkennbaren Geraden kann man auf die Asymptotik

$$|p(x)| \leq O(|\, x - x_\infty\,|^2)$$

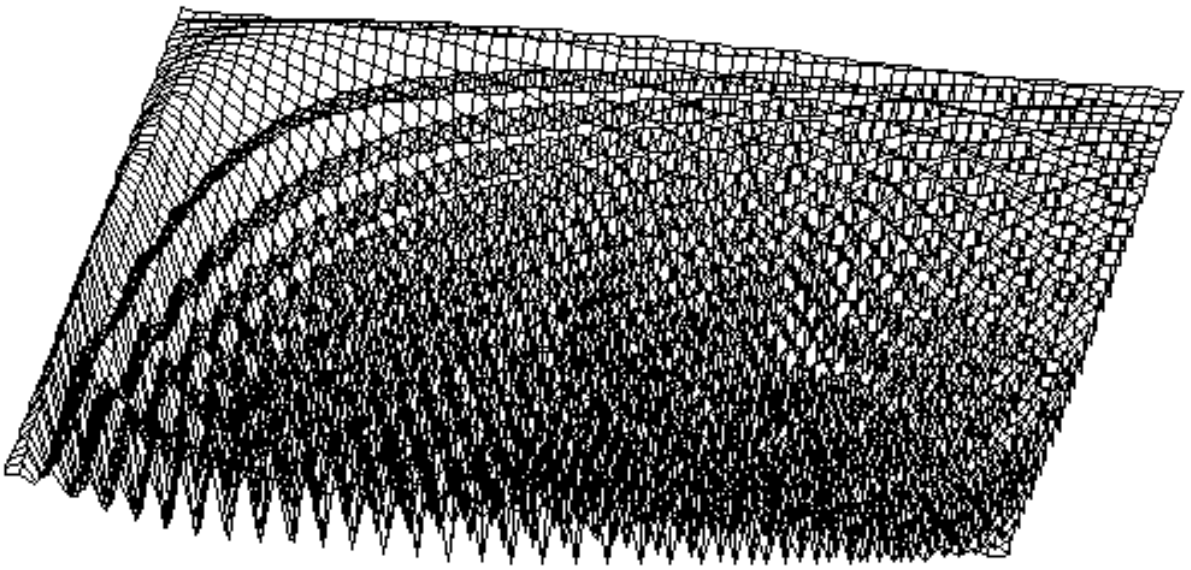
schließen. Somit verschwindet die adjungierte Zustandsfunktion in der Umgebung des singulären Punkts zweimal langsamer als im ähnlichen Fall des Fuller Problems aus Abschnitt 3.2.1.

Zum Abschluß kann man die fundamentalen Funktionen (3.78) für die Fortsetzung der Lösung $y^* \to \sigma^*$ in das ganze Gebiet Ω verwenden. Die Verteilung der Verschmutzung im Wasser ohne und mit den Reinigungsmaßnahmen sieht man auf der Abbildung 3.8. In der unteren Graphik von 3.8 ist im Bereich $x_\infty < x < 1$ ein nichtnullidentischer Zustand zu erkennen, d.h. $y^*(x) \neq k_d(x) = 0$. Somit ist das Bang-Bang-Prinzip nicht erfüllt. Eine plausible Erklärung dazu ergibt sich, wenn man die adjungierte Zustandsgleichung in allgemeiner Form

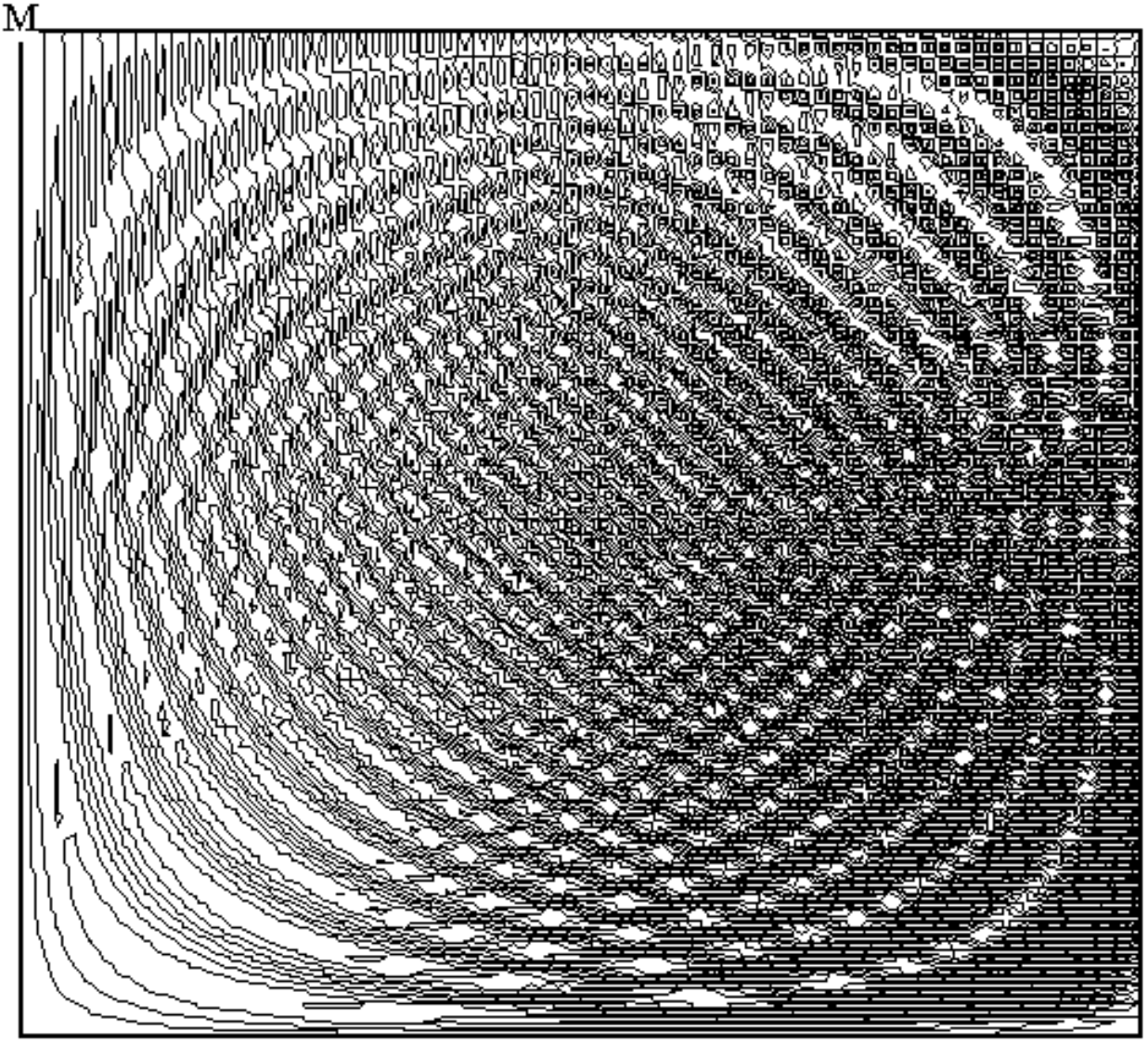
$$A^* p \;\; = \;\; y - k_d$$

(ggf. gilt: $A^* = P^*$) betrachtet. Weil der Operator P^* eine globale Wirkung über dem Bereich (0, 1) hat, folgt daraus nicht, daß dort, wo $p(x)$ verschwindet, auch die Werte $A^* p(x)$ verschwinden müssen. Bemerkenswert an diesem Beispiel ist, daß obwohl der Poincaré-Operator sämtliche Eigenschaften eines elliptischen Operators besitzt, kein verallgemeinertes Bang-Bang-Prinzip [15], [16] gilt. Ähnliche Beispiele könnte man auch für parabolische Operatoren erster Ordnung konstruieren.

□



M



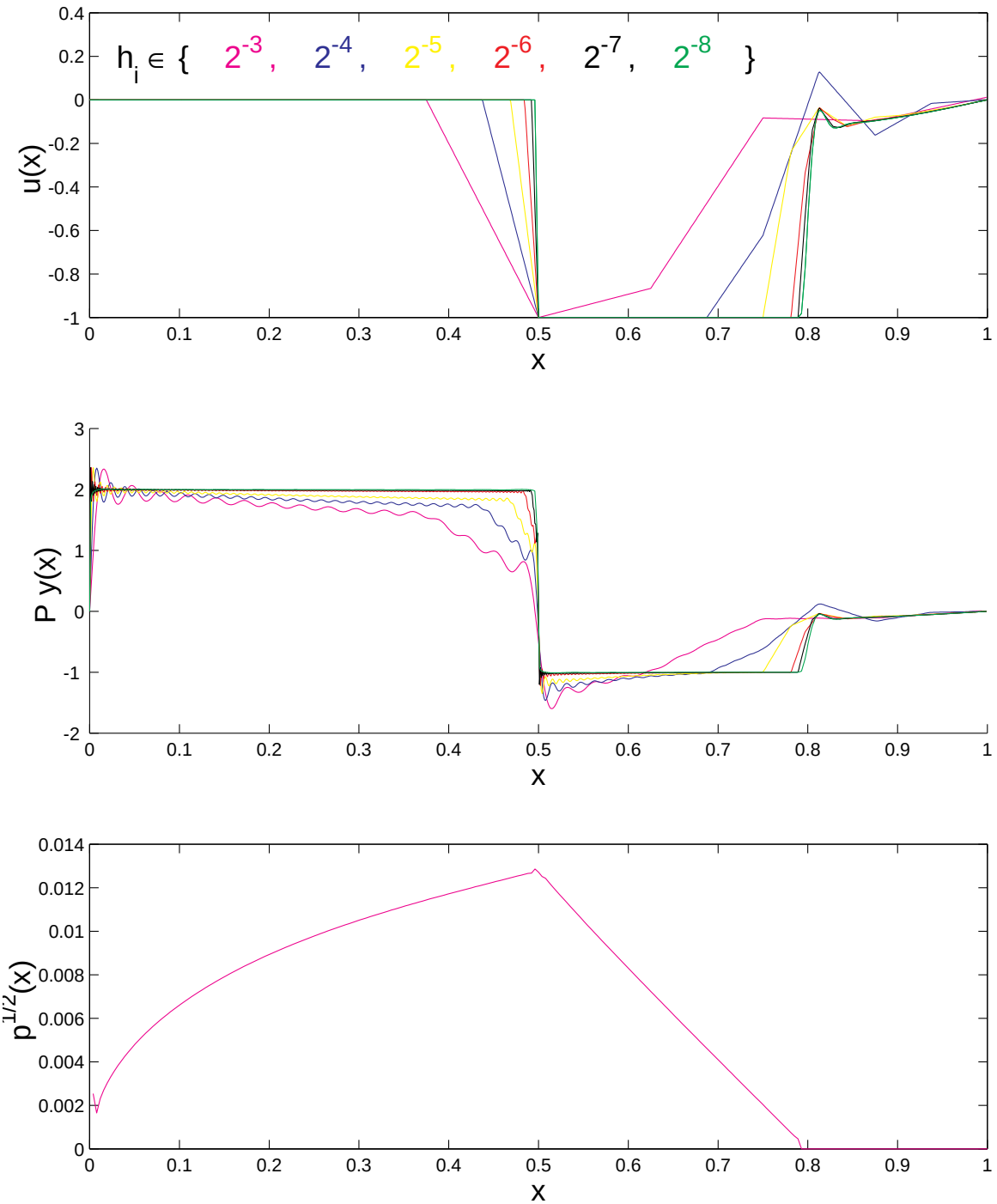


Abbildung 3.7: Konvergenz des Strafverfahrens ohne die Regularisierung

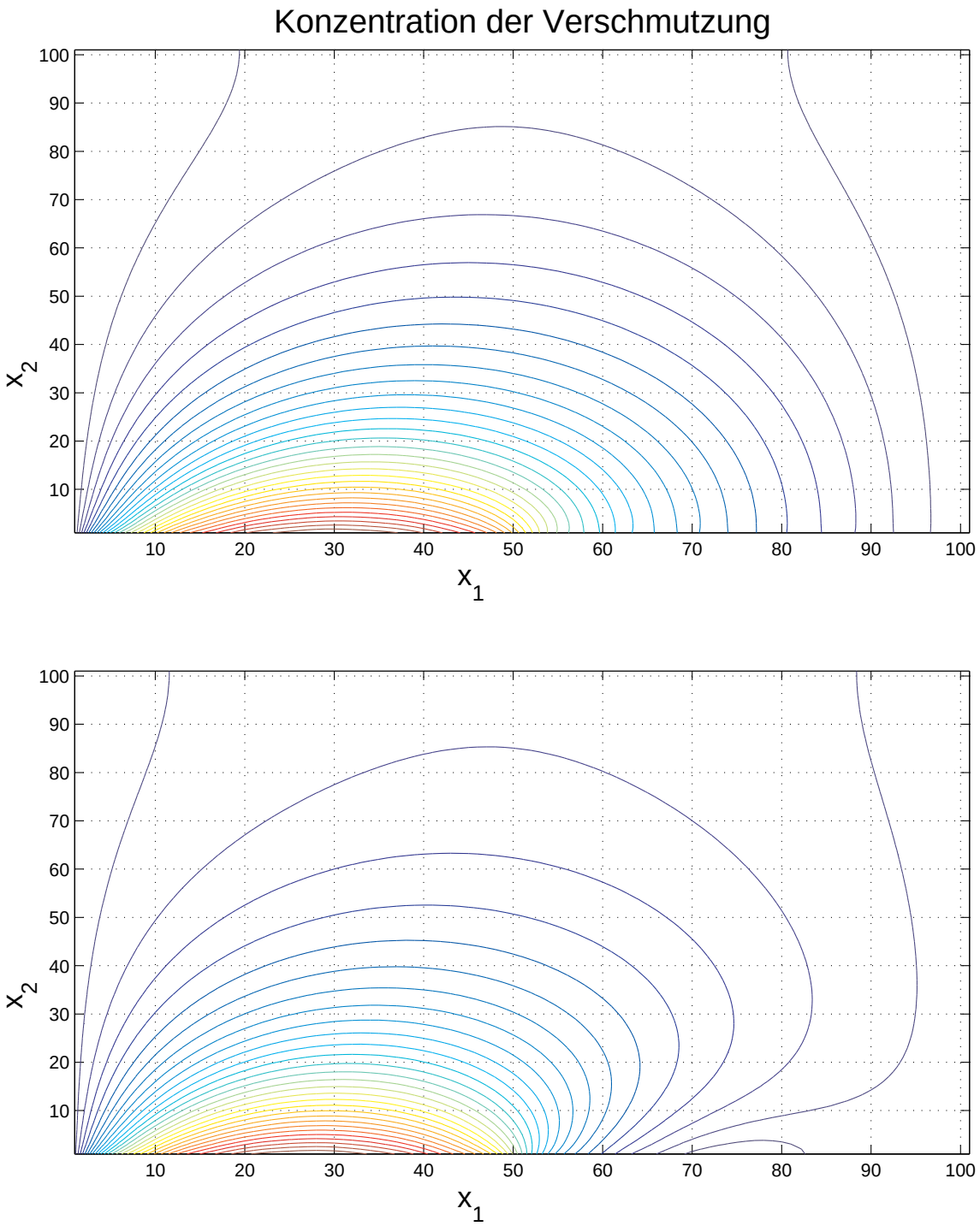


Abbildung 3.8: Zustand des Wassers ohne und mit den Reinigungsmaßnahmen

Kapitel 4

Konvergenz innerer Approximationen im MSR-Verfahren

In diesem Kapitel konzentrieren wir uns auf die Untersuchung schlecht gestellter Kontrollprobleme, in denen der Bereich der zulässigen Steuerungen eine konvexe, abgeschlossene und **beschränkte** Teilmenge des Steuerraumes $\mathcal{U}$ darstellt. Die Fälle mit beschränkten Steuerprozessen sind in der Praxis am häufigsten anzutreffen und wurden stets ausführlich diskutiert [14], [28], [68]. Die Beschränktheit des Bereichs der zulässigen Steuerungen sichert die Koerzivität eines linear-quadratischen Optimalsteuerproblems und somit seine Lösbarkeit. Dennoch kann Existenz einer Lösung auch im Falle der unbeschränkten Steuerprozesse vorliegen (s. Beispiel in [35]). Wenn der Radius R_0 einer solchen Lösung von oben abschätzbar ist, so kann man eine der zwei Strategien verfolgen: entweder fügt man eine zusätzliche Restriktion $\|u\|_{\mathcal{U}} \leq R_0$ hinzu [13, Remark 3.2], welche während der numerischen Realisierung von Algorithmen einen größeren Rechenaufwand verursacht, oder man benutzt eine Regularisierung mit speziell gewählten Parametern [35, Part II], [39], was aus praktischer Hinsicht günstiger sein wird. Im zweiten Fall ist aber eine Straftechnik gefragt, um ein schlecht gestelltes Kontrollproblem auf eine (nicht stark) konvexe Variationsungleichung [40] zurückzuführen.

Weil die meisten uns bekannten Fälle schlechtgestellter Kontrollprobleme mit einem unbeschränkten Bereich der zulässigen Steuerungen rein akademische Beispiele sind, haben wir sie hier nicht betrachtet. Wir können deshalb auf die in den vorigen Kapiteln sorgfältig untersuchten Strafverfahren ohne jegliche Schwierigkeiten verzichten, d. h. die lineare Zustandsgleichung in Ersatzproblemen des MSR-Verfahrens wird nun in keinen Strafterm eingebettet, sondern erzeugt eine selbständige exakte Restriktion. Bei der Diskretisierung der Ersatzprobleme müssen jetzt für eine von dieser Restriktion im Prozessraum erzeugte Menge richtige innere Approximationen konstruiert werden, welche man mit einem Ritz- oder Galerkin-Verfahren [66] realisieren kann. Die erhaltenen diskretisierten Probleme behandelt man anschließend

mit Standardverfahren der klassischen Optimiereung [17], [80].

In Abschnitt 4.1 beginnen wir mit der Analyse einer elliptischen Gleichung 2-ter Ordnung mit gemischten Randbedingungen. Die Regularitätsvoraussetzungen für die Lösung und die notwendigen Abschätzungen für den Diskretisierungsfehler werden diskutiert. Im nachfolgenden Abschnitt formulieren wir ein konkretes linear-quadratisches Optimalsteuerproblem mit der beschriebenen elliptischen Zustandsgleichung, sowie ein MSR-Verfahren zur approximativen Lösung dieses Problems. In Abschnitt 4.3 werden die Konvergenzkriterien für die Näherungen im MSR-Verfahren erhalten. Es gelingt dabei einen expliziten Zusammenhang zwischen dem Abbruchparameter des inneren Prox-Zyklus δ_i und anderen Einstellungsparametern des MSR-Verfahrens herzuleiten. Der Verzicht auf die Straftechniken erlaubt außerdem die bessere Asymptotik (4.53) zu bekommen, welche die Diskretisierungs- und Regularisierungsparameter vereinbart:

$$\sum_{i=1}^{\infty} \frac{h_i}{\chi_i} < \infty \, .$$

Jetzt darf man den Regularisierungsparameter auf einer Folge von Gittern schneller verkleinen, was immer von Vorteil ist, wenn man das Proximal-Punkt-Verfahren auf schlecht konditionierte Aufgaben anwendet. Aus Abschnitt 4.4 kann man sehen, daß die T- und MSR-Verfahren auch praktisch realisierbar sind und erlauben, recht große Optimalsteuerprobeme mit partiellen Ableitungen zu berechnen. Selbst solch komplizierte Phenomena wie "Chattering regimes" und singuläre Steuerbereiche lassen sich mit den von uns entwickelten numerischen Verfahren nachvollziehen.

4.1 Approximation der elliptischen Gleichung 2-ter Ordnung mit gemischten Randbedingungen

Der erste Schritt bei der Untersuchung eines Optimalsteuerproblems ist die Diskussion der Zustandsgleichung. Dabei gilt es nachzuweisen, daß diese Gleichung für jede zulässige Steuerung eine möglichst eindeutige Lösung besitzt, die einem vorgegebenen Raum angehört. Die Theorie elliptischer Differentialgleichungen ist in dieser Beziehung sehr weit ausgearbeitet. An dieser Stelle ist es nicht möglich, auf alle bekannten Feinheiten einzugehen. Der Abschnitt verfolgt nur das Ziel, Eigenschaften zusammenzustellen, die bei der weiteren Untersuchung der inneren Approximationen für das linear-quadratische, inkorrekt gestellte Optimalsteuerproblem notwendig sind und zu zeigen, daß diese Aussagen die Voraussetzungen aus Kapitel 1 erfüllen. Bei der Auswahl der Voraussetzungen geht es dabei vor allem um Bedingungen, die ohne große Schwierigkeiten überprüft werden können.

Die in diesem Abschnitt erhaltenen Resultate könnten sicherlich für den Fall eines n-dimensionalen Gebiets ($n \geq 3$) mit unendlich glattem Rand [56], [81] ausgearbeitet werden. Dabei gilt es die Schwierigkeiten bei der Approximation des

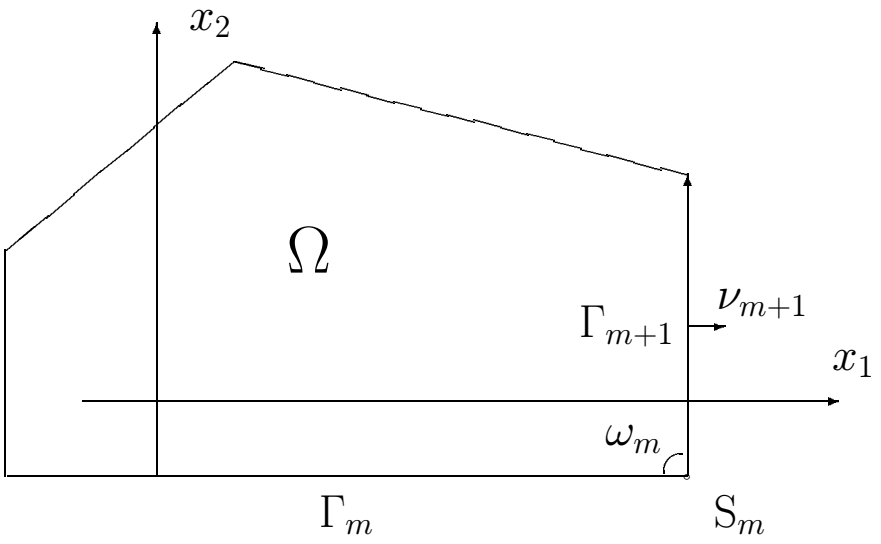


Abbildung 4.1: Das Polygongebiet

Randes zu überwinden. Die Resultate in [21] und [82] zeigen, daß die notwendigen Regularitätsergebnisse auch im Falle scharfer Kanten und Ecken für $n \geq 3$ erreicht werden können. Solche Erweiterungen sind auch möglich, wenn man statt der Poisson-Gleichung eine allgemeine elliptische Gleichung 2-ter Ordnung mit genügend glatten Koeffizienten betrachtet. Für die elasto-mechanischen Zustandsgleichungen kann man entsprechende Resultate in [45] finden.

Wir betrachten nun ein konvexes zweidimensionales Polygongebiet Ω mit Eckpunkten $S_m,\ m \in \mathcal{M} = \{1, \dots, M\}$. Der Rand des Gebiets wird in zwei Teile partitioniert: $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$. Jeder Teil läßt sich wieder als Vereinigung einer bestimmten Anzahl von Seiten des Polygons $\Gamma_m,\ m \in \mathcal{M}$ darstellen:

$$\Gamma_D = \bigcup_{m \in \mathcal{M}_D} \Gamma_m \,, \qquad \Gamma_N = \bigcup_{m \in \mathcal{M}_N} \Gamma_m \,,$$

mit $\mathcal{M}_D \cup \mathcal{M}_N = \mathcal{M}$, $\mathcal{M}_D \neq \emptyset$. Die Seiten und Ecken des Polygons werden gemäß der positiven Orientierung des Randes (gegen den Uhrzeigersinn) numeriert. Dabei sei S_m der Endpunkt von Γ_m . Jede Ecke wird durch ihren Winkelbetrag $\omega_m \in (0,\, \pi],\ m \in \mathcal{M}$ charakterisiert.

Auf einem solchen polygonalen Gebiet untersuchen wir das Randwertproblem

$$\begin{aligned} -\Delta\, y(x) &= f(x) && \text{in } \Omega \,, \\ \frac{\partial y}{\partial \nu_m}(x) &= g_m(x) && \text{auf } \Gamma_m, \quad m \in \mathcal{M}_N \,, \\ y(x) &= 0 && \text{auf } \Gamma_m, \quad m \in \mathcal{M}_D \,, \end{aligned} \tag{4.1}$$

mit $f \in L_2(\Omega),\ g_m \in L_2(\Gamma_m)$.

Sei $s = k + \lambda > 0$, $k \in \mathbf{Z}_+$, $\lambda \in [0,\, 1)$ und $\mathrm{M} \subset \mathbf{R}^n$ meßbar, mes $\mathrm{M} > 0$. Wir führen folgende Normenbezeichnungen ([29], [105] und [22]) ein:

$$\|u\|_{s,\,\mathrm{M}} = \|u\|_{W^k(\mathrm{M})} = \left(\sum_{|\alpha| \le k} \int\limits_{\mathrm{M}} |D^\alpha\, u(x)|^2\, dx \right)^{1/2}, \qquad \text{falls } \lambda = 0\ ,$$

sonst

$$\|u\|_{s,\,\mathrm{M}} = \left(\|u\|^2_{W^k(\mathrm{M})} + \sum_{|\alpha| = k} \iint\limits_{\mathrm{M}\times\mathrm{M}} \frac{|D^\alpha\, u(x) - D^\alpha\, u(y)|^2}{|x-y|^{n+2\lambda}}\, dxdy \right)^{1/2},$$

und

$$\|f\|_{-s,\,\mathrm{M}} = \sup_{\phi \in C^\infty(\overline{\mathrm{M}})\setminus\{0\}} \frac{\langle\, f,\, \phi\,\rangle}{\|\phi\|_{s,\,\mathrm{M}}}\ .$$

Um die schwache Formulierung des Problems (4.1) zu ermöglichen, führen wir folgenden Hilbertraum ein:

$$\mathcal{Y} = \left\{\, y \in W^1(\Omega) :\ \gamma_0^{(m)}\, y = 0 \quad \text{in } H^{1/2}(\Gamma_m)\,,\ \ m \in \mathcal{M}_D \,\right\}, \tag{4.2}$$

der im weiteren als Zustandsraum behandelt wird. Die Spuroperatoren $\gamma_0^{(m)}$ beziehen sich in (4.2) auf diejenigen Seiten des Polygons, welche den homogenen Dirichlet Randbedingungen entsprechen. Jetzt besteht das Variationsproblem darin, daß man nach einem Element $y \in \mathcal{Y}$ sucht, welches für beliebiges $\tilde{y} \in \mathcal{Y}$ folgende Gleichung erfüllt:

$$\sum_{j=1,2} \left(D_{x_j}\, y,\, D_{x_j}\, \tilde{y} \right)_\Omega = (f,\, \tilde{y})_\Omega + \sum_{m \in \mathcal{M}_N} \left(g_m,\, \gamma_0^{(m)} \tilde{y} \right)_{\Gamma_m}\ . \tag{4.3}$$

Da die Greensche Formel für beschränkte Lipschitz-glatte Gebiete gilt ([29], [76]), erfüllt jede "klassische" Lösung von (4.1) die Variationsgleichung (4.3) . Gleichung (4.3) kann man in die äquivalente Operatorform

$$-\,\bar{\Delta}\, y = f + \sum_{m \in \mathcal{M}_N} g_m\, \delta_{\Gamma_m} \qquad \text{in } \mathcal{Y}' \tag{4.4}$$

umschreiben wobei sich der verallgemeinerte Laplace-Operator und die Delta-Schichten auf dieselbe Weise wie im Abschnitt 3.1.2 definieren lassen. Die Operatorform der elliptischen Gleichung (4.4) ist mit dem klassischen Randwertproblem (4.1) nicht identisch. In (4.1) werden alle auftretenden Funktionen als messbar angenommen, wohingegen es sich in (4.4) um allgemeine Distributionen handelt. Die Elemente f und g_m dürfen jetzt außerhalb des Raumes L_2 liegen. Dashalb muß man zusätzliche Regularitätsbedingungen für die Elemente g_m, $m \in \mathcal{M}_N$ finden, welche die Existenz einer "klassischen" Lösung $y \in \mathcal{Y} \cap W^2(\Omega)$ von (4.4) sicherstellen [29, Kap. 4] .

Jedem Eckpunkt S_m auf dem orientierten Rand $\partial\Omega$ ordnen wir eine periodische Vektorfunktion $x_m \in [C^{0,1}(-\infty, \infty)]^2$ zu. Es gilt dabei, daß der Vektor $x_m(s)$ unendlich oft den Rand $\partial\Omega$ umläuft, wärend sich der Parameter s , der als vorzeichenorientierte Distanz vom Punkt S_m auf der Kurve $\partial\Omega$ definiert ist, von $-\infty$ bis ∞ ändert. Wir untersuchen nun jeden Eckpunkt S_m und stellen folgende (bi)quadratischen Funktionale zusammen:

A. Wenn $m \in \mathcal{M}_N$ und $(m+1) \in \mathcal{M}_D$, dann

$$T_m(g_m) = \int_0^{c_m} \frac{|g_m(x_m(-s))|^2}{s}\, ds \;;$$

B. Wenn $m \in \mathcal{M}_D$ und $(m+1) \in \mathcal{M}_N$, dann

$$T_m(g_{m+1}) = \int_0^{c_m} \frac{|g_{m+1}(x_m(s))|^2}{s}\, ds \;;$$

C. Wenn $m \in \mathcal{M}_N$ und $(m+1) \in \mathcal{M}_N$, dann

$$T_m(g_m,\, g_{m+1}) = \int_0^{c_m} \frac{|g_{m+1}(x_m(s)) - g_m(x_m(-s))|^2}{s}\, ds \;,$$

wobei $c_m = \min\{\text{mes}\,\Gamma_m, \text{mes}\,\Gamma_{m+1}\}$. (Der Index $M+1$ sei mit dem Index 1 identifiziert, wobei M die Anzahl der Seiten des Polygons bezeichnet). Wir können jetzt gemäß der oben angeführten Klassifikation der Eckpunkte entsprechende Indexuntermengen $\mathcal{M}_A \subset \mathcal{M}$, $\mathcal{M}_B \subset \mathcal{M}$ und $\mathcal{M}_C \subset \mathcal{M}$ bestimmen.

Satz 14 *Wenn das System der Elemente g_m, $m \in \mathcal{M}_N$ in (4.4) folgende Bedingungen erfüllt:*

- $g_m \in H^{1/2}(\Gamma_m)$, $m \in \mathcal{M}_N$;
- $T_m(g_m) < \infty$, *falls* $\omega_m = \pi/2$ *und* $m \in \mathcal{M}_A$;
- $T_m(g_{m+1}) < \infty$, *falls* $\omega_m = \pi/2$ *und* $m \in \mathcal{M}_B$;
- $T_m(g_m,\, g_{m+1}) < \infty$, $m \in \mathcal{M}_C$

und $\omega_m \le \pi/2$, $m \in \mathcal{M}_A \cup \mathcal{M}_B$, so gehört die Lösung der Zustandsgleichung (4.4) dem Raum $W^2(\Omega)$ an.

Den Beweis des Satzes 14 findet man in [29, Kap. 1 und 4]. Die im Falle von Dirichlet-Neumann-Ecken bei $\omega_m > \pi/2$ existierenden Singularitäten sind in [78] ausführlich beschrieben. Da $y(x)$ nun eine "klassische" Lösung ist, gilt auch der Satz 4.4.1.2 in [29] über die Eindeutigkeit dieser Lösung.

Dennoch benötigen wir weitere Untersuchungen, um die Korrektheit der Aufgabe (4.4) (im Sinne von Hadamard) nachzuweisen, da die Existenz und Eindeutigkeit der Lösung von (4.4) in $\mathcal{Y} \cap W^2(\Omega)$ unter den getroffenen Bedingungen keine Information über das Verhalten der Lösung bei in $\mathcal{Y}'$ endlichen Störungen der rechten Seite in (4.4) ergibt. Zum Zweck einer solchen Stabilitätsuntersuchung formulieren wir die Operatorgleichung (4.4) als Extremalproblem im Hilbertraum $\mathcal{Y}$:

$$F(y) \xrightarrow{!} \min_{y \in \mathcal{Y}} \tag{4.5}$$

mit

$$F(y) = \frac{1}{2} \sum_{j=1,2} \left\| D_{x_j} y \right\|_\Omega^2 - \left\langle \bar{f},\, y \right\rangle_{\mathcal{Y}',\mathcal{Y}} \ .$$

Im Fall des Randwertproblems (4.1) ergibt sich das Funktional $\bar{f} \in \mathcal{Y}'$ auf folgende Weise:

$$\left\langle \bar{f},\, y \right\rangle_{\mathcal{Y}',\mathcal{Y}} = (f,\, y)_\Omega + \sum_{m \in \mathcal{M}_N} \left(g_m,\, \gamma_0^{(m)} y \right)_{\Gamma_m} \ .$$

Um nun die Korrektheit von (4.4) zu zeigen, genügt es die starke Konvexität des Zielfunktionals $F(y)$ in (4.5) nachzuweisen [40]. Dafür benötigen wir den Satz von Friedrichs.

Satz 15 *Für jedes Element $y \in \mathcal{Y}$ mit $\mathcal{M}_D \neq \emptyset$ existiert eine positive von y unabhängige Konstante c_F , so daß folgende Ungleichung gilt:*

$$\|y\|_\Omega \leq c_F \sqrt{\sum_{j=1,2} \left\| D_{x_j} y \right\|_\Omega^2} \ . \tag{4.6}$$

Beweis: Man findet den allgemeinen Beweis der Existenz von c_F in [19, Theorem 1.2.1]. Die richtige Auswertung dieser Konstanten führen wir unter speziellen Annahmen durch. Sei $\omega_1 < \pi$, $\omega_M < \pi$ und $\Gamma_1 = AB \subset \Gamma_D$, dann kann man das Gebiet Ω (s. Abb. 4.2) mit dem Dreieck ABC fortsetzen, so daß ein Kreis mit endlichem Radius d in dieses Dreieck eingebettet werden kann. Nach klassischer Beweistechnik [94] betrachten wir zunächst eine Folge unendlich glatter Funktionen $y_k(x)$, $k = 1, 2, ..$, welche gegen die zu untersuchende Funktion $y(x)$ in $W^1(\Omega)$ konvergiert. Die Existenz einer solchen Folge läßt sich aus der Dichtheit des Raumes $C^\infty(\overline{\Omega})$ in $W^1(\Omega)$ ersehen [29, § 1.4.2]. Da die Dichtheitsverhältnisse mit derselben Norm $\| \cdot \|_{1,\Omega}$ auf dem abgeschlossenen linearen Teilraum $\mathcal{Y}$ in $W^1(\Omega)$ erhalten

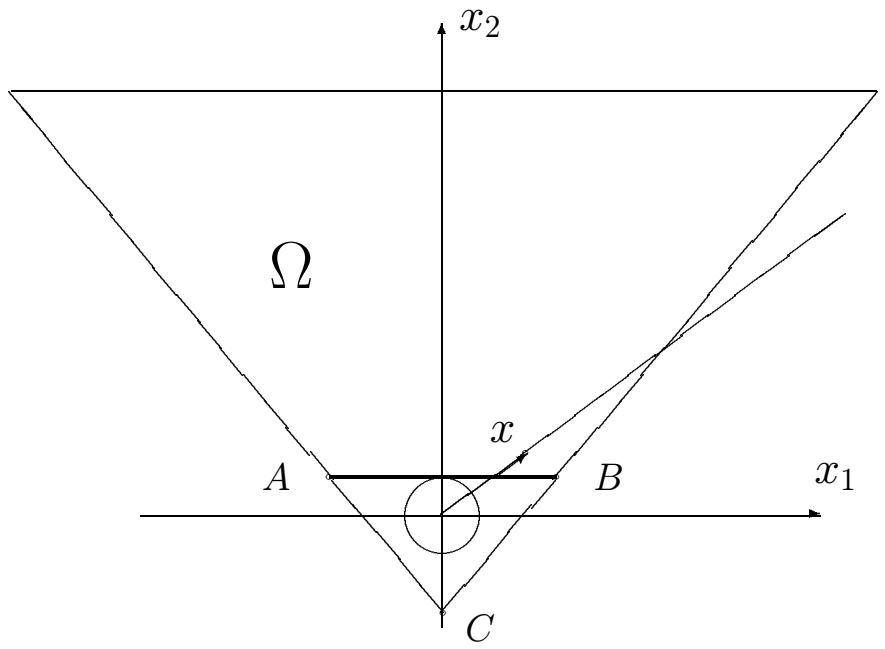


Abbildung 4.2: Friedrichssche Ungleichung

bleiben, kann man die Funktionen $y_k(x)$ so konstruieren, daß ihre Werte auf Γ_1 gleich Null sind. Wir setzen jetzt alle Funktionen $y(x)$, $y_1(x)$, $y_2(x)$, .. mit Null im Dreieck ABC fort, ohne die Bezeichnungen zu ändern und wählen das Koordinatensystem, wie es auf der Abbildung 4.2 gezeigt ist. Seien (ρ, ω) die entsprechenden Polarkoordinaten. Den Funktionswert im Inneren des Gebiets erhählt man dann folgendermaßen

$$y_k(x) = \int_0^r \frac{\partial y_k}{\partial \rho}(\rho, \omega)\, d\rho \,.$$

Die Funktion unter dem Integral ist stückweise stetig in ρ . Nach der Cauchy–Schwarz Ungleichung gilt

$$|y_k(x)|^2 \le D \int_0^r \left|\frac{\partial y_k}{\partial \rho}\right|^2 d\rho \,,$$

wobei D den Diameter von Ω darstellt. Wir integrieren die letzte Ungleichung über dem Gebiet und erhalten

$$\int_\Omega |y_k(x)|^2\, dx \;\le\; D \int_{\omega_{min}}^{\omega_{max}} d\omega \int_{R_{min}(\omega)}^{R_{max}(\omega)} r\, dr \int_0^r d\rho \left|\frac{\partial y_k}{\partial \rho}\right|^2 \le$$

$$\frac{D^3}{2} \int_{\omega_{min}}^{\omega_{max}} d\omega \int_{R_{min}(\omega)}^{R_{max}(\omega)} \frac{\rho}{d}\, d\rho \left|\frac{\partial y_k}{\partial \rho}\right|^2 \le \frac{D^3}{2\,d} \int_\Omega \sum_{j=1,2} \left(D_{x_j}\, y_k \right)^2 dx \,.$$

Bei obigen Abschätzungen nutzt man die Tatsache, daß $\rho/d \geq 1$ gilt. Für $k \to \infty$ erhält man

$$c_F = \sqrt{\frac{D^3}{2\,d}} \ .$$

□

Folgerung 1 *Es gilt* $\forall\, y_0 \in \mathcal{Y} \cap W^2(\Omega),\ \forall\, y \in \mathcal{Y}$ *:*

$$\nabla^2_{yy} F(y_0)(y)(y) \geq \frac{1}{1 + c_F^2}\, \|y\|^2_{1,\Omega} \ .$$

Folgerung 2 *Der Operator* $-\bar{\Delta} \in \mathcal{L}(\mathcal{Y},\, \mathcal{Y}')$ *ist stetig invertierbar und es gilt*

$$m_A\ \|y\|_{\mathcal{Y}} \leq \left\| -\bar{\Delta}\, y \right\|_{\mathcal{Y}'} \leq M_A\ \|y\|_{\mathcal{Y}} \qquad \forall\, y \in \mathcal{Y}\ ,$$

mit den Lipschitz-Konstanten

$$m_A = \frac{1}{1 + c_F^2}\ , \qquad M_A = 1\ . \tag{4.7}$$

Die Folgerung 1 sichert die starke Konvexität des Funktionals in (4.5), sowie die Korrektheit der Zustandsgleichung (4.4).

Für die spätere Analyse im Abschnitt 4.3 benötigen wir noch die Abschätzung der L_2-Norm einer Funktion auf dem Rande durch die W^1-Norm derselben Funktion über dem Gebiet. Die Abschätzung läßt sich verbessern, erfordert aber die Auswertung der Konstanten durch die geometrischen Parameter des Gebiets.

Satz 16 *Für jedes Element* $y \in W^1(\Omega)$ *existiert eine positive von* y *unabhängige Konstante* c_γ *, so daß folgende Ungleichung gilt:*

$$\|\gamma_0\, y\|_{\partial\Omega} \leq c_\gamma\ \|y\|_{1,\Omega} \tag{4.8}$$

Beweis: Wie im Beweis des Satzes 15 wählen wir eine gegen $y(x)$ in $W^1(\Omega)$ konvergierende Folge $y_k(x) \in C^\infty(\Omega)$, $k = 1, 2, ..$ aus. Auf Grund seiner scharfen Winkel erlaubt das Polygon immer die Einbettung eines Kreises der Art, daß der Polarstrahl des auf der Abbildung 4.3 aufgezeichneten Koordinatensystems genau einen Schnittpunkt $\tilde{x}$ mit dem Rand des Gebiets Ω bildet. Seien $(\rho,\, \omega)$ die Polarkoordinaten und der Einheitsvektor $\kappa(\omega) = (\cos\omega,\, \sin\omega)$ bestimme dann eine Polarrichtung. Wir bezeichnen mit $l(\omega)$ die positive Distanz zwischen dem Ursprungspunkt und dem Punkt $\tilde{x}$, d. h. $l(\omega) = |\,\tilde{x}\,|$. Die Bezeichnung $\nu(\omega)$ gilt für den Einheitsvektor der entsprechenden äußeren Normale. Man benötigt desweiteren noch die Gleichheit

$$\frac{\partial}{\partial r}\left(r^{3/2}\, y_k(r,\, \omega)\right) \ = \ \frac{3}{2}\sqrt{r}\ y_k(r,\, \omega) + r^{3/2}\, \frac{\partial y_k(r,\, \omega)}{\partial r}\ . \tag{4.9}$$

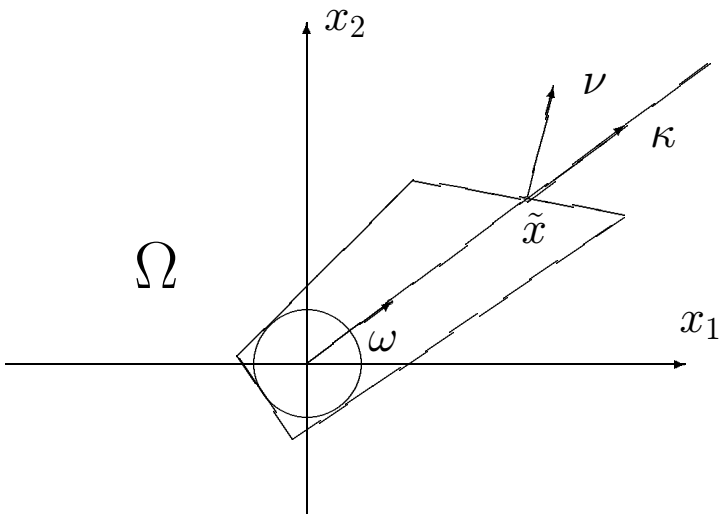


Abbildung 4.3: Integration über dem polygonalen Gebiet

Da die Funktionen $y_k(x)$ stetig sind, können wir ihre Werte auf dem Rand $\partial\Omega$ untersuchen. Es gilt

$$y_k(\tilde{x})\, l^{3/2}(\tilde{x}) \;=\; \int_0^{l(\omega)} \frac{\partial}{\partial r} \left(\rho^{3/2}\, y_k(\rho,\, \omega)\right)\, d\rho\ .$$

Jetzt quadrieren wir diese Gleichung und integrieren über dem Rand des Gebiets. Man entwickelt folgende Kette mit Hilfe der Cauchy-Schwarz Ungleichung, mit (4.9) und mit der leicht zu prüfenden Beziehung $(a+b)^2 \le 2a^2 + 2b^2$:

$$\int_{\partial\Omega} y_k^2(\tilde{x})\, l^3(\tilde{x})\, d\tilde{x} = \int_{\partial\Omega} \left(\int_0^{l(\omega)} \frac{\partial}{\partial r} \left(\rho^{3/2}\, y_k(\rho,\, \omega)\right)\, d\rho \right)^2 d\tilde{x} =$$

$$\int_0^{2\pi} \left(\int_0^{l(\omega)} \frac{\partial}{\partial r} \left(\rho^{3/2}\, y_k(\rho,\, \omega)\right)\, d\rho \right)^2 \frac{l(\omega)}{(\kappa(\omega),\, \nu(\omega))}\, d\omega \le$$

$$\frac{\max\limits_{0<\omega\le 2\pi} l^2(\omega)}{\min\limits_{0<\omega\le 2\pi} (\kappa(\omega),\, \nu(\omega))} \int_0^{2\pi} \int_0^{l(\omega)} \left(\frac{3}{2}\, \sqrt{\rho}\; y_k(\rho,\, \omega) + \rho^{3/2}\, \frac{\partial y_k(\rho,\, \omega)}{\partial \rho} \right)^2 d\rho\, d\omega \le$$

$$\frac{R_{max}^2}{C_{min}} \int_0^{2\pi} \int_0^{l(\omega)} \left(\frac{9}{2}\, \rho\, y_k^2(\rho,\, \omega) + 2\, \rho^3 \left| \frac{\partial y_k(\rho,\, \omega)}{\partial \rho} \right|^2 \right)\, d\rho\, d\omega \le$$

$$\frac{R_{max}^2}{C_{min}} \left(\frac{9}{2} \|y_k\|_\Omega^2 + 2\, R_{max}^2 \left\| \frac{\partial y_k}{\partial \rho} \right\|_\Omega^2 \right) \leq$$

$$\frac{R_{max}^2}{C_{min}} \max\{ \frac{9}{2}, 2\, R_{max}^2 \}\, \|y_k\|_{1,\Omega}^2 \,.$$

Endgültig bekommt man im Limes $k \to \infty$ folgende Auswertung

$$c_\gamma \;=\; \sqrt{\frac{R_{max}^2}{R_{min}^3 C_{min}} \max\{ \frac{9}{2}, 2\, R_{max}^2 \}}$$

mit

$$C_{min} \;=\; \min_{0<\omega\leq 2\pi} (\kappa(\omega), \nu(\omega)) \,,$$

$$R_{min} \;=\; \min_{0<\omega\leq 2\pi} l(\omega) \,,$$

$$R_{max} \;=\; \max_{0<\omega\leq 2\pi} l(\omega) \,.$$

Offensichtlich sind die letzten drei Größen streng größer als Null. Sie lassen sich in konkreten Fällen leicht berechnen.

□

Man kann die Beweise der Sätze 15 und 16 auf den n-dimensionalen Fall übertragen.

Es gelingt sehr selten, eine geschlossene Form für die Lösung von (4.4) zu erhalten. Somit stellen approximative Lösungen manchmal die einzige Möglichkeit dar, Informationen über die exakte Lösung zu bekommen.

Nach der Finite-Element-Technik [66], [72], [89], beginnen wir die Approximation mit einer Dreieckszerlegung (oder Viereckseinteilung) des Polygongebiets. Wir betrachten konkret eine Folge eingeschachtelter Triangulationen. Jedes Gitter erhält eine Nummer $i = 1, 2, \ldots$. Es werden die Mengen aller Knoten $\{\mathrm{K}_j^i,\ j \in \mathcal{J}^i\}$ und aller elementaren Teilgebiete $\{\Omega_j^i,\ j \in \mathcal{I}^i\}$ auf dem i-ten Diskretisierungsniveau bestimmt. Man definiert den *Diskretisierungsparameter* des Gitters

$$h_i = \max_{j\in\mathcal{I}^i} \mathrm{diam}(\Omega_j^i) \,,$$

der die Feinheit der Triangulation charakterisiert. Wir sondern eine Untermenge der Indizes $\mathcal{J}_0^i \subset \mathcal{J}^i$ für diejenigen Knoten aus, welche nicht auf dem Dirichlet-Teil des Randes Γ_D liegen. Wenn es sich um ein fixiertes Diskretisierungsniveau handelt, lassen wir den Index i weg. Jedem K_j, $j \in \mathcal{J}_0$ wird eine bekannte "Dach"-Funktion $\varphi_j(x)$ [27], [31], [66], [89] zugeordnet. Man definiert nun den Raum der *inneren Approximationen* entsprechend [27] folgendermaßen:

$$\mathcal{Y}^h = \left\{ y_h \in \mathcal{Y} :\ y_h(x) = \sum_{j\in\mathcal{J}_0} y_j\, \varphi_j(x),\ y_j \in \mathbf{R} \right\}$$

und betrachtet das diskretisierte Extremalproblem (4.5):

$$F(y_h) \xrightarrow{!} \min_{y_h \in \mathcal{Y}^h} \tag{4.10}$$

mit

$$F(y_h) = \frac{1}{2} \sum_{i=1,2} \|D_{x_i}\, y_h\|_\Omega^2 - (f,\, y_h)_\Omega - \sum_{m \in \mathcal{M}_N} \left(g_m,\, \gamma_0^{(m)} y_h\right)_{\Gamma_m} .$$

Es entsteht nun die Frage: wie gut nähert die Lösung von (4.10) die entsprechende Lösung des exakten Problems (4.5) an? Die bekannten Approximationsergebnisse in [66], [89] betreffen nur die Randbedingungen ersten, zweiten oder dritten Typs, berühren aber keine gemischten Randwertprobleme. Dennoch sind die folgenden zwei bekannten Sätze auch in unserem Fall von Bedeutung.

Satz 17 *Sei $y \in \mathcal{Y}$ die exakte Lösung des Extremalproblems (4.5) und $y_h \in \mathcal{Y}^h$ die exakte Lösung des diskretisierten Problems (4.10) , dann gilt*

$$\sum_{i=1,2} \|D_{x_i}\,(y_h - y)\|_\Omega^2 \le \sum_{i=1,2} \|D_{x_i}\,(\tilde{y}_h - y)\|_\Omega^2 \qquad \forall\, \tilde{y}_h \in \mathcal{Y}^h$$

Satz 18 *Wenn die Funktion $u(x)$ dem Raum $W^2(\Omega)$ angehört, dann gilt*

$$\sum_{|\alpha|=1} \|D^\alpha\,(u - u_I)\|_\Omega^2 \le C^2\, h^2 \sum_{|\alpha|=2} \|D^\alpha\, u\|_\Omega^2$$

mit der Interpolante

$$u_I(x) = \sum_{j \in \mathcal{J}} u(x_j)\, \varphi_j(x) \; . \tag{4.11}$$

Der Beweis des Satzes 17 läßt sich im Fall der gemischten Randbedingungen analog wie in [66] durchführen, wo man Eigenschaften des Ritz-Verfahrens für homogene Dirichlet-Bedingungen untersucht. Der Satz 18 stellt einen Sonderfall des Satzes 3.3 in [89] dar. Gemäß Satz 15 und Satz 17, sowie der Definition der inneren Approximationen, erhält man die Konvergenz

$$\lim_{h \to 0} \|y_h - y\|_{1,\Omega} = 0 \; .$$

Um die Abschätzung der Konvergenzgeschwindigkeit zu bekommen, wenden wir den Satz 18 auf die Lösung $y(x)$ des Problems (4.5) an und berücksichtigen dabei folgende Eigenschaft auf dem Dirichlet-Teil des Randes:

$$\text{Weil } y(x_j) = 0 \quad \forall\, j \in \mathcal{J} \setminus \mathcal{J}_0 \; , \;\; \text{gilt } y_I \in \mathcal{Y}^h \; .$$

Die zweiten Ableitungen der Funktion $y(x)$ sind nach Satz 14 in $L_2(\Omega)$ beschränkt, was zu der folgenden endgültigen Abschätzung führt. Sei $y \in \mathcal{Y}$ die exakte Lösung des Extremalproblems (4.5) und $y_h \in \mathcal{Y}^h$ die exakte Lösung des diskretisierten Problems (4.10) , dann gilt

$$\|y - y_h\|_{1,\Omega} \leq \tilde{C}_1\, h \left(\|f\|_\Omega + \sum_{m \in \mathcal{M}_N} \|g_m\|_{1/2,\Gamma_m} \; + \right.$$

$$\left. \sum_{m \in \mathcal{M}_A} T_m^{1/2}(g_m) + \sum_{m \in \mathcal{M}_B} T_m^{1/2}(g_{m+1}) + \sum_{m \in \mathcal{M}_C} T_m^{1/2}(g_m,\, g_{m+1}) \right) .$$

Da $y - y_h \in \mathcal{Y} \subset L_2(\Omega)$ und die homogenen Neumannschen Bedingungen die Voraussetzungen des Satzes 14 erfüllen, kann man die Abschätzungstechnik von Nitsche [89] anwenden. Es gilt

$$\|y - y_h\|_\Omega \leq \tilde{C}_0\, h^2 \left(\|f\|_\Omega + \sum_{m \in \mathcal{M}_N} \|g_m\|_{1/2,\Gamma_m} \; + \right. \tag{4.12}$$

$$\left. \sum_{m \in \mathcal{M}_A} T_m^{1/2}(g_m) + \sum_{m \in \mathcal{M}_B} T_m^{1/2}(g_{m+1}) + \sum_{m \in \mathcal{M}_C} T_m^{1/2}(g_m,\, g_{m+1}) \right) .$$

4.2 Approximation des linear-quadratischen Optimalsteuerproblems mit elliptischer Zustandsgleichung

Im Kapitel 2 wurde ein abstraktes linear-quadratisches Steuerproblem im Hilbertraum eingeführt und untersucht. Um das Verhalten der Iterierten im MSR-Verfahren zu studieren, betrachten wir nun einen Sonderfall, in dem man die im Kapitel 2 vorbereiteten Operatoren, Räume und konvexen Mengen wiederfindet.

$$\begin{aligned} J(y,\, u) = \frac{1}{2}\, \|y\|^2_{\Omega'} \; &\stackrel{!}{\longrightarrow} \min_{(y,\, u) \in \mathcal{Y} \times \mathcal{U}_{ad}} \\ -\,\bar{\Delta}\, y = u + f + \sum_{m \in \mathcal{M}_N} g_m\, \delta_{\Gamma_m} \qquad &\text{in } \mathcal{Y}' \,, \end{aligned} \tag{4.13}$$

wobei der Zustandsraum $\mathcal{Y}$ nach (4.2) definiert ist. Für den Steuerraum und den Bereich der zulässigen Steuerungen wählen wir

$$\mathcal{U} = L_2(\Omega)\,, \qquad \mathcal{U}_{ad} = \{u \in \mathcal{U} :\ \xi_0(x) \le u(x) \le \xi_1(x)\ \text{ f.ü. in } \Omega\}\ ,$$

mit

$$\xi_0(x) \in W^1(\Omega) \cap L_\infty(\Omega)\,, \qquad \xi_1(x) \in W^1(\Omega) \cap L_\infty(\Omega)\,.$$

Das Gebiet Ω' kann als offenes meßbares Teilgebiet [44] von Ω gewählt werden. Je kleiner das Teilgebiet Ω' ist, umso stärker ist der Beobachtungsraum $\mathcal{K} = L_2(\Omega')$ eingeengt.

Aus Kapitel 3 wissen wir, daß das Fehlen der quadratischen Steuerkosten unter dem Zielfunktional die Inkorrektheit des Steuerproblems (4.13) hervorrufen kann. Die dort beschriebenen numerischen Erfahrungen zeigen, daß die inkorrekt gestellten Kontrollprobleme eine sehr schwierige Klasse restringierter Extremalprobleme bilden und darum die Entwicklung spezieller numerischer Algorithmen erfordern. Im Rahmen dieses Kapitels untersuchen wir die bekannten "Chattering regimes" [103], [104] und setzen uns das Ziel, die gleichen Effekte für die verteilte Steuerung über einem zweidimensionalen Polygongebiet zu entdecken. Solche Aufgaben korrelieren auch direkt mit Problemstellungen in [35], in denen man inhomogene Neumannsche Randbedingungen betrachtet. Da das Verhalten der Lösung ohnehin sehr kompliziert ist (s. Abschnitt 3.2.1), beschränken wir unser Studium auf den Fall, daß es in (4.13) keine Zustandsrestriktionen gibt: $\mathcal{Z}_{ad} = \mathcal{Y} \times \mathcal{U}_{ad}$. Im Abschnitt 2.2.3 ist die Existenz der Lösung für ein Problem in allgemeiner Form bewiesen. Wir betrachten deshalb wieder die Lösungsmenge $\mathcal{Z}^*$ im Prozeßraum $\mathcal{Z} = \mathcal{Y} \times \mathcal{U}$, welcher mit der Norm

$$\|z\|_{\mathcal{Z}} \;=\; \sqrt{\|y\|^2_{1,\Omega} + \|u\|^2_{\Omega}} \tag{4.14}$$

versehen ist. Aufgrund des "kleinen" Beobachtungsraumes kann man keine eindeutige Lösung vom Problem (4.13) erwarten, weil Satz 9 jetzt nicht mehr gültig ist.

Wir korrigieren den MSR-Algorithmus (vgl. S. 11), der zur näherungsweisen Lösung des Steuerproblems (4.13) verwendet wird, indem wir eine (endliche) Abweichung des Näherungselements $z^{i,s}$ von der exakten Lösung $\bar{z}^{i,s}$ des Ersatzproblems

$$\bar{\Psi}_{i,s}(y,\,u) = \frac{1}{2}\ \|y\|^2_{\Omega'} + \frac{\chi_i}{2}\ \left\|u - u^{i,s-1}\right\|^2_{\Omega} \xrightarrow{\ !\ } \min_{(y,\,u)\,\in\,\mathcal{Z}_{ad}}\ ,$$

$\bar{\mathbf{P}}_{i,s}$:

$$-\,\bar{\Delta}\,y = u + f + \sum_{m \in \mathcal{M}_N} g_m\,\delta_{\Gamma_m} \quad \text{in } \mathcal{Y}' \tag{4.15}$$

in der schwächeren Norm (4.14) zulassen. Im übrigen bleibt die Struktur des MSR-Algorithmus für das auf der Seite 107 beschriebene Verfahren erhalten.

Nun erklären wir konkreter, woraus der Fehler beim Lösen des Ersatzproblems $\bar{\mathbf{P}}_{i,s}$ besteht. Insgesamt werden wir zwei Anteile des Fehlers finden. Wir beginnen mit der Untersuchung des Diskretisierungsfehlers und benötigen dazu noch die inneren Approximationen für $\mathcal{U}$ und $\mathcal{U}_{ad}$ analog zur Approximation des Zustandsraumes $\mathcal{Y}$ im Abschnitt 4.1. Jedem elementaren Dreieck Ω_j wird eine stückweise konstante Basisfunktion $\psi_j(x)$, $j \in \mathcal{I}$ [22], [66], [89] zugeordnet. Jetzt können wir auch die innere Approximation definieren:

$$\mathcal{U}^h \;=\; \left\{ u_h \in L_2(\Omega) :\; u_h(x) = \sum_{j \in \mathcal{I}} u_j \, \psi_j(x),\; u_j \in \mathbf{R} \right\} .$$

Um die innere Aproximation der konvexen Menge $\mathcal{U}_{ad}$ gemäß [27] zu bestimmen, führen wir folgende Bezeichnung für den Mittelwert über dem elementaren Dreieck Ω_j einer beliebigen Funktion $\psi \in L_2(\Omega)$ ein:

$$M_j(\psi) = \frac{1}{\text{mes } \Omega_j} \int\limits_{\Omega_j} \psi \, dx \, . \tag{4.16}$$

Nun definieren wir die Menge der inneren Approximationen folgendermaßen:

$$\mathcal{U}^h_{ad} = \{ u_h \in \mathcal{U}^h :\; M_j(\xi_0) \le u_h(x) \le M_j(\xi_1) \;\text{ f.ü. in } \Omega_j \,,\; \forall\, j \in \mathcal{I} \} \, .$$

Es sei bemerkt, daß hier zur Vereinfachung alle Approximationsmengen $\mathcal{Y}^h, \mathcal{U}^h, \mathcal{U}^h_{ad}$ durch den gemeinsamen Diskretisierungsparameter h_i verbunden sind, wohingegen man in [22] mit zwei unterschiedlichen Diskretisierungsparametern — für die Zustands- und Steuerfunktionen getrennt — arbeitet.

Sei $z^{i,s}_h = (y^{i,s}_h,\, u^{i,s}_h)$ die exakte Lösung des diskretisierten Ersatzproblems

$$\mathbf{P}^h_{i,s}: \qquad \bar{\Psi}_{i,s}(y_h,\, u_h) \xrightarrow{\;!\;} \min_{(y_h,\, u_h) \,\in\, \mathcal{Y}^h \times \mathcal{U}^h_{ad}} \, , \tag{4.17}$$

$$\sum_{i=1,2} (D_{x_i}\, y_h,\, D_{x_i}\, \tilde{y}_h)_\Omega \;=\; (u_h + f,\, \tilde{y}_h)_\Omega + \sum_{m \in \mathcal{M}_N} \left(g_m,\, \gamma_0^{(m)} \tilde{y}_h \right)_{\Gamma_m} \quad \forall\, \tilde{y}_h \in \mathcal{Y}^h \, ,$$

mit der in die Variationsform umgeschriebenen Zustandsgleichung. Das Problem (4.17) ist ein endlichdimensionales Problem der quadratischen Programmierung (s. Abschnitt 4.4), welches mit numerischen Verfahren der klassischen Optimierung [17], [80] behandelt werden kann. Diese Verfahren besitzen auch einen Fehler, so daß die exakte Lösung $z^{i,s}_h$ nimmer berechnet werden kann. Somit müssen wir eine inexakte Lösung $(y^{i,s},\, u^{i,s})$ des Ersatzproblems (4.17) betrachten. Das MSR-Verfahren zur

Gegeben $u^{0,0} \in \mathcal{U}\,;\ s(0) := 0\,;\ \ i := 1\,;$

(a) $u^{i,0} := u^{i-1,s(i-1)}\ ;\ \ s := 0\,;$

(b) Gegeben $u^{i,s} \in \mathcal{U}\,;\ s := s+1\,;$

Berechne eine Nährung $z^{i,s} = (y^{i,s},\, u^{i,s}) \in \mathcal{Y}^h \times \mathcal{U}^h_{ad}$

der exakten Lösung $z^{i,s}_h = (y^{i,s}_h,\, u^{i,s}_h) \in \mathcal{Y}^h \times \mathcal{U}^h_{ad}$

des diskretisierten Hilfsproblems $\mathbf{P}^h_{i,s}$, so daß

$$\bar{\Psi}_{i,s}(y^{i,s},\, u^{i,s}) - \bar{\Psi}_{i,s}(y^{i,s}_h,\, u^{i,s}_h) \le \eta_i\ , \tag{4.18}$$

und

$$\sum_{i=1,2} \Big(D_{x_i}\, y^{i,s},\, D_{x_i}\, \tilde{y}_h\Big)_\Omega = \Big(u^{i,s} + f,\, \tilde{y}_h\Big)_\Omega + \sum_{m \in \mathcal{M}_N} \Big(g_m,\, \gamma_0^{(m)} \tilde{y}_h\Big)_{\Gamma_m} \quad \forall\, \tilde{y}_h \in \mathcal{Y}^h\ ;$$

(c) Wenn $\| u^{i,s} - u^{i,s-1} \|_\Omega \ge \delta_i\ ,$

dann gehe nach (b)

sonst $s(i) := s\ ;\ i := i+1\,;$ gehe nach (a)

Abbildung 4.4: Das MSR-Verfahren mit der Näherungsstrategie

Lösung des Originalproblems (4.13) erhält somit seine endgültige Fassung (s. Abbildung 4.4).

Der Verfahrensparameter η_i charakterisiert den Rechnungsfehler. Die Annahme, daß die diskretisierte Zustandsgleichung exakt lösbar ist, ist erfüllbar, da die Zustandsgleichung selbst, wie im Abschnitt 4.1 gezeigt, eine korrekt gestellte Aufgabe ist und sich somit mit großer Genauigkeit lösen läßt.

4.3 Das Konvergenzkriterium der Näherungen im MSR-Verfahren

Das Ziel dieses Abschnitts ist der Beweis des Konvergenzsatzes (Satz 19). Wir beginnen mit dem einfachsten Teil der Untersuchung, mit der Abschätzung des Rechnungsfehlers durch die Verfahrensparameter.

Lemma 11 *Für die im Abschnitt 4.2 eingeführten Näherungen gilt*

$$\|y^{i,s} - y_h^{i,s}\|_{1,\Omega} \leq \sqrt{1 + c_F^2}\, \sqrt{\frac{2\,\eta_i}{\chi_i}}\ , \quad \|u^{i,s} - u_h^{i,s}\|_{\Omega} \leq \sqrt{\frac{2\,\eta_i}{\chi_i}}\ . \tag{4.19}$$

Beweis: Man betrachtet das Funktional $\bar{\Psi}_{i,s}(\cdot\,,\cdot)$ in (4.15), die Differenz (4.18) und gruppiert die entstehenden Glieder um:

$$\frac{1}{2}\ \left\|y^{i,s} - y_h^{i,s}\right\|_{\Omega'}^2 + \frac{\chi_i}{2}\ \left\|u^{i,s} - u_h^{i,s}\right\|_{\Omega}^2 + \tag{4.20}$$

$$\left(y_h^{i,s},\, y^{i,s} - y_h^{i,s}\right)_{\Omega'} + \chi_i\ \left(u_h^{i,s} - u^{i,s-1},\, u^{i,s} - u_h^{i,s}\right)_{\Omega} \leq \eta_i\ .$$

Die Optimalitätsbedingungen 1-ter Ordnung für das Problem $\mathbf{P}_{i,s}^h$ ergeben für ein beliebiges Element $(\tilde{y}_h,\, \tilde{u}_h) \in \mathcal{Y}^h \times \mathcal{U}_{ad}^h$, das die diskretisierte Zustandsgleichung in (4.17) genau erfüllt,

$$\left(y_h^{i,s},\, \tilde{y}_h - y_h^{i,s}\right)_{\Omega'} + \chi_i\ \left(u_h^{i,s} - u^{i,s-1},\, \tilde{u}_h - u_h^{i,s}\right)_{\Omega} \geq 0\ . \tag{4.21}$$

Gemäß den Annahmen im Abschnitt 4.2 erfüllt die inexakte Lösung $(y^{i,s},\, u^{i,s})$ die Zustandsgleichung genau. Somit können wir in (4.21) die Substitution $(\tilde{y}_h,\, \tilde{u}_h) := (y^{i,s},\, u^{i,s})$ ausführen und mit dem erhaltenen Ergebnis die Abschätzung (4.20) nach unten fortsetzen. Man erhält dadurch die zweite Ungleichung in (4.19).

Für die y-Komponenten gilt aufgrund der diskretisierten Zustandsgleichung

$$\sum_{i=1,2} \left(D_{x_i}\ \left(y^{i,s} - y_h^{i,s}\right),\, D_{x_i}\,\tilde{y}_h\right)_{\Omega} = \left(u^{i,s} - u_h^{i,s},\, \tilde{y}_h\right)_{\Omega} \quad \forall\, \tilde{y}_h \in \mathcal{Y}^h\ .$$

Man setzt $\tilde{y}_h := y^{i,s} - y_h^{i,s}$ und erhält durch die Friedrichssche Ungleichung

$$\|y^{i,s} - y_h^{i,s}\|_{1,\Omega} \leq \sqrt{1 + c_F^2}\ \|u^{i,s} - u_h^{i,s}\|_{\Omega}\ .$$

□

Die nächste zu klärende Frage ist, wie gut das diskretisierte Problem $\mathbf{P}^h_{i,s}$ das Problem $\bar{\mathbf{P}}_{i,s}$ approximiert? Wir fixieren die Indizes i und s im MSR-Verfahren und werden sie zur Vereinfachung der Schreibweise bei sich bietender Gelegenheit weglassen. Es sei erwähnt, daß die meisten unten durchgeführten Beweistechniken und Bezeichnungen aus [22] entnommen wurden. Die Zustandsgleichungen in (4.15) und (4.17) ergeben zwei affine Abbildungen $\bar{y}(\bar{u})$ und $y_h(u_h)$, welche nach den Optimalitätsbedingungen 1-ter Ordnung für $\mathbf{P}_{i,s}$ und $\mathbf{P}^h_{i,s}$ folgende Ungleichungen erfüllen:

$$(\bar{y}(\bar{u}),\, \bar{y}(v) - \bar{y}(\bar{u}))_{\Omega'} + \chi_i \left(\bar{u} - u^{i,s-1},\, v - \bar{u}\right)_{\Omega} \geq 0 \qquad \forall\, v \in \mathcal{U}_{ad}\,, \tag{4.22}$$

$$(y_h(u_h),\, y_h(v_h) - y_h(u_h))_{\Omega'} + \chi_i \left(u_h - u^{i,s-1},\, v_h - u_h\right)_{\Omega} \geq 0 \quad \forall\, v_h \in \mathcal{U}^h_{ad}\,. \tag{4.23}$$

Addiert man (4.22) und (4.23) miteinander, so erhält man die Ungleichung

$$a_y(v,\, v_h) + \chi_i\, a_u(v,\, v_h) \geq 0 \qquad \forall\, v \in \mathcal{U}_{ad}\,, \quad \forall\, v_h \in \mathcal{U}^h_{ad} \tag{4.24}$$

mit

$$\begin{aligned} a_u \;=\; & \left(\bar{u} - u^{i,s-1},\, v - \bar{u}\right)_{\Omega} + \left(u_h - u^{i,s-1},\, v_h - u_h\right)_{\Omega} = \\ & - \|\bar{u} - u_h\|^2_{\Omega} + (\bar{u},\, v - u_h)_{\Omega} + (\bar{u},\, v_h - \bar{u})_{\Omega} + \\ & (u_h - \bar{u},\, v_h - \bar{u})_{\Omega} - \left(u^{i,s-1},\, (v - u_h) + (v_h - \bar{u})\right)_{\Omega} \end{aligned} \tag{4.25}$$

und

$$\begin{aligned} a_y \;=\; & (\bar{y}(\bar{u}),\, \bar{y}(v) - \bar{y}(\bar{u}))_{\Omega'} + (y_h(u_h),\, y_h(v_h) - y_h(u_h))_{\Omega'} = \\ & - \|\bar{y}(\bar{u}) - y_h(u_h)\|^2_{\Omega'} + (\bar{y}(\bar{u}),\, (\bar{y}(v) - \bar{y}(u_h)) + (\bar{y}(v_h) - \bar{y}(\bar{u})))_{\Omega'} + \\ & (\bar{y}(\bar{u}),\, (\bar{y}(u_h) - y_h(u_h)) + (y_h(v_h) - \bar{y}(v_h)))_{\Omega'} + \\ & ((y_h(u_h) - \bar{y}(\bar{u})),\, (y_h(v_h) - \bar{y}(v_h)) + (\bar{y}(v_h) - \bar{y}(\bar{u})))_{\Omega'}\,. \end{aligned} \tag{4.26}$$

Somit ergibt sich aus (4.24) die folgende Beziehung

$$\begin{aligned} \chi_i\, \|\bar{u} - u_h\|^2_{\Omega} + \|\bar{y}(\bar{u}) - y_h(u_h)\|^2_{\Omega'} \;\leq\; & \\ \chi_i \left(\left(u^{i,s-1},\, (v - u_h) + (v_h - \bar{u})\right)_{\Omega} \;+\; & \\ (\bar{u},\, v - u_h)_{\Omega} + (\bar{u},\, v_h - \bar{u})_{\Omega} + (u_h - \bar{u},\, v_h - \bar{u})_{\Omega}\right) \;+\; & \\ (\bar{y}(\bar{u}),\, (\bar{y}(v) - \bar{y}(u_h)) + (\bar{y}(v_h) - \bar{y}(\bar{u})))_{\Omega'} \;+\; & \end{aligned} \tag{4.27}$$

$$(\bar{y}(\bar{u}),\, (\bar{y}(u_h) - y_h(u_h)) + (y_h(v_h) - \bar{y}(v_h)))_{\Omega'} \quad +$$

$$((y_h(u_h) - \bar{y}(\bar{u})),\, (y_h(v_h) - \bar{y}(v_h)) + (\bar{y}(v_h) - \bar{y}(\bar{u})))_{\Omega'}$$

Die Ungleichung (4.27) wird in Lemma 17 zur Abschätzung des Diskretisierungsfehlers in der Norm (4.14) zugrundegelegt. Die aus (4.27) stammenden Ansätze erfordern außerdem die Auswertung des Approximationsfehlers der Steuervariablen in der L_2-Norm, was immer nur in dem Fall möglich ist, wenn man eine obere Grenze für die Werte $\|\bar{u}\|_{1,\Omega_j}$, $j \in \mathcal{I}$ berechnen kann [66], [89]. Diese Berechnung erfolgt nach Lemma 12, welches eine Verallgemeinerung zu Lemma 2 in [22] für $\mathbf{P}_{i,s}$ darstellt, und erfordert ihrerseits, daß der vorletzte Prox-Punkt $u^{i,s-1}$ auf demselben oder auf dem vorherigen gröberen Gitter diskretisiert ist. **Somit ist wesentlich, daß die Gitter exakt eingeschachtelt sind.** Das konkrete Beispiel mit einem Fragment von zwei nacheinander folgenden Gittern wird auf der Abbildung 4.5 angeführt.

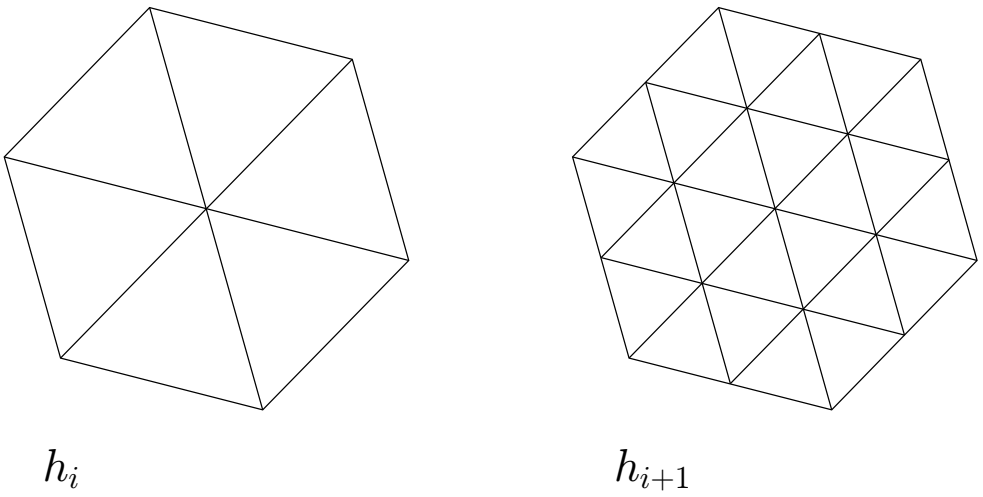


Abbildung 4.5: Die eingeschachtelten Triangulationen

Lemma 12 *Wenn die Funktion $u^{i,s-1}$ dem Raum $\bigoplus_{j\in\mathcal{I}} W^1(\Omega_j)$[1] angehört, dann gilt mit (4.30)*

$$\sum_{j\in\mathcal{I}} \|\bar{u}\|^2_{1,\Omega_j} \le \frac{2\tilde{c}_y^2(1+c_F^2)}{\chi_i^2} + 2\left\|u^{i,s-1}\right\|^2_{\Omega} + \|\xi_0\|^2_{1,\Omega} + \|\xi_1\|^2_{1,\Omega} \ .$$

Beweis: Die Regularität der Lösung $\bar{u}$ des Ersatzproblems $\bar{\mathbf{P}}_{i,s}$ kann mit Standardtechniken von J.-L. Lions [57] gezeigt werden. Das Endergebnis unterscheidet

[1]Die Mengen $W^1(\Omega_1)$, $W^1(\Omega_2)$,.. können als Teilräume von $\mathcal{U}$ gelten, wenn man die zugehörigen Funktionen jeweils mit Null in das ganze Gebiet Ω fortsetzt.

sich von dem in [57] unerheblich:

$$\bar{u}(x) = \begin{cases} u^{i,s-1}(x) - \dfrac{\bar{p}(x)}{\chi_i} & \text{falls} \quad \xi_0(x) \le u^{i,s-1}(x) - \dfrac{\bar{p}(x)}{\chi_i} \le \xi_1(x) \ , \\ \xi_0(x) & \text{falls} \quad u^{i,s-1}(x) - \dfrac{\bar{p}(x)}{\chi_i} < \xi_0(x) \ , \\ \xi_1(x) & \text{falls} \quad u^{i,s-1}(x) - \dfrac{\bar{p}(x)}{\chi_i} > \xi_1(x) \ , \end{cases} \tag{4.28}$$

wobei $\bar{p}(x)$ die Lösung der adjungierten Zustandsgleichung darstellt. Da für Hilberträume $\mathcal{Y}'' = \mathcal{Y}$ gilt und der verallgemeinerte Laplace-Operator selbstadjungiert ist, d. h. $(-\bar{\Delta})^* = -\bar{\Delta}$ in $\mathcal{L}(\mathcal{Y}, \mathcal{Y}')$, kann $\bar{p}(x)$ als klassische Lösung des Poisson-Problems mit gemischten Randbedingungen

$$\begin{aligned} -\Delta\,\bar{p}(x) &= \bar{y}(x) && \text{in } \Omega \ , \\ \frac{\partial \bar{p}}{\partial \nu_m}(x) &= 0 && \text{auf } \Gamma_m, \quad m \in \mathcal{M}_N \ , \\ \bar{p}(x) &= 0 && \text{auf } \Gamma_m, \quad m \in \mathcal{M}_D \end{aligned} \tag{4.29}$$

betrachtet werden. Nach Satz 14 gilt offensichtlich $\bar{p} \in \mathcal{Y} \cap W^2(\Omega)$. Ganz analog wie in [22] kann man das Gebiet Ω gemäß (4.28) in drei Teile zerlegen. Wir führen diese Zerlegung lokal auf jedem Ω_j, $j \in \mathcal{I}$ durch:

$$\|\bar{u}\|^2_{1,\Omega_j} \le \left\| u^{i,s-1} - \frac{\bar{p}}{\chi_i} \right\|^2_{1,\Omega_j} + \|\xi_0\|^2_{1,\Omega_j} + \|\xi_1\|^2_{1,\Omega_j} \ .$$

Weil man mit eingeschachtelten Gittern zu tun hat, existieren die Größen $\|u^{i,s-1}\|_{1,\Omega_j}$, $j \in \mathcal{I}$, und da die verteilte Steuerung mit stückweise konstanten Basisfunktionen diskretisiert ist, gilt außerdem $\|u^{i,s-1}\|_{1,\Omega_j} = \|u^{i,s-1}\|_{\Omega_j}$.

Die Abschätzung für die adjungierte Zustandsfunktion $\bar{p}(x)$ kann man jetzt aus Lemma 13 entnehmen und somit den geführten Beweis abschliessen.

□

Lemma 13 *Die Funktionen $\bar{y}(x)$ und $\bar{p}(x)$ sind beschränkt, dabei gilt*

$$\|\bar{y}\|_\Omega \le \tilde{c}_y \ , \qquad \|\bar{p}\|_{1,\Omega} \le \tilde{c}_y \sqrt{1 + c_F^2} \ ,$$

mit

$$\tilde{c}_y = c_F^2 \left(\|f\|_\Omega + \sup_{v \in \mathcal{U}_{ad}} \|v\|_\Omega + c_\gamma \sqrt{1 + c_F^{-2}} \sum_{m \in \mathcal{M}_N} \|g_m\|_{\Gamma_m} \right) \ . \tag{4.30}$$

Beweis: Die Beschränktheit der exakten Zustandsfunktion $\bar{y}$ in $L_2(\Omega)$ ergibt sich aus der Variationsform der Zustandsgleichung

$$\sum_{j=1,2} \left(D_{x_j}\, \bar{y},\, D_{x_j}\, \tilde{y} \right)_\Omega = (\bar{u} + f,\, \tilde{y})_\Omega + \sum_{m \in \mathcal{M}_N} \left(g_m,\, \gamma_0^{(m)} \tilde{y} \right)_{\Gamma_m} , \tag{4.31}$$

mit $\tilde{y} := \bar{y}$ und aus den Sätzen 15 und 16. Auf das letzte Glied in (4.31) wird außerdem die Cauchy-Schwarz Ungleichung angesetzt. Man erhält

$$\begin{aligned} \frac{1}{c_F}\, \|\bar{y}\|_\Omega \;\; &\leq \;\; \sqrt{\sum_{j=1,2} \left\| D_{x_j}\, \bar{y} \right\|_\Omega^2} \leq \\ & c_F\; (\|\bar{u}\|_\Omega + \|f\|_\Omega) + \sqrt{1 + c_F^2}\; c_\gamma \sum_{m \in \mathcal{M}_N} \|g_m\|_{\Gamma_m} \;, \end{aligned} \tag{4.32}$$

wobei $\bar{u}$ im Raum $\mathcal{U} = L_2(\Omega)$ beschränkt ist (s. auch M_{ad} unter primären Konstanten im Anhang A). Die Berechnung der Konstante $\tilde{c}_y$ beruht unmittelbar auf (4.32).

Die Beschränktheit der adjungierten Zustandsfunktion $\bar{p}$ in $W^1(\Omega)$ ergibt sich analog, auf Grund der Variationsform zu (4.29) aus folgender Ungleichungskette

$$\|\bar{p}\|_{1,\Omega} \leq \sqrt{1 + c_F^2}\; \|\bar{y}\|_\Omega \leq \tilde{c}_y \sqrt{1 + c_F^2}\;.$$

□

Die folgenden drei Lemmata (mit gering geänderten Bezeichnungen) entsprechen Lemmen 3, 4 und 5 in [22] . Der Beweis von Lemma 15 stimmt dabei mit dem von Lemma 4 aus [22] vollständig überein. Diese Ergebnisse stellen ein klassisches Werkzeug zur Analyse der inneren Approximationen für konvexe Mengen im L_2-Raum dar. Da es desweiteren sich nur um die Existenz einiger Konstanten handelt, betrachten wir eine gemeinsame, majorisierende Konstante C .

Lemma 14 *Sei $\bar{u}$ die Lösung von $\bar{\mathbf{P}}_{i,s}$ und v_h gegeben durch*

$$v_h(x) = \sum_{j \in \mathcal{I}} M_j(\bar{u})\, \psi_j(x) \;,$$

mit $M_j(\cdot)$ nach (4.16). Dann gilt $v_h \in \mathcal{U}_{ad}^h$ und es existiert eine von $\bar{u}$ und h unabhängige Konstante C , so daß

$$\|\bar{u} - v_h\|_\Omega \;\; \leq \;\; Ch \left(\sum_{j \in \mathcal{I}} \|\bar{u}\|_{1,\Omega_j}^2 \right)^{1/2} , \tag{4.33}$$

$$\|\bar{u} - v_h\|_{-1,\Omega} \;\; \leq \;\; Ch^2 \left(\sum_{j \in \mathcal{I}} \|\bar{u}\|_{1,\Omega_j}^2 \right)^{1/2} , \tag{4.34}$$

$$\left(\sum_{j \in \mathcal{I}} \|\bar{u} - v_h\|_{-1,\Omega_j}^2 \right)^{1/2} \;\; \leq \;\; Ch^2 \left(\sum_{j \in \mathcal{I}} \|\bar{u}\|_{1,\Omega_j}^2 \right)^{1/2} . \tag{4.35}$$

Beweis: Ganz analog wie in Lemma 3 aus [22] erhält man lokale Abschätzungen

$$\begin{aligned} \|\bar{u} - v_h\|_{\Omega_j} &\leq \tilde{C}h \, \|\bar{u}\|_{1,\Omega_j} \, , \\ \|\bar{u} - v_h\|_{-1,\Omega_j} &\leq \tilde{C}h^2 \, \|\bar{u}\|_{1,\Omega_j} \, . \end{aligned}$$

Wenn man die Ungleichungen über alle $j \in \mathcal{I}$ addiert, so erhält man (4.33) und (4.35). Nun analog wie in [22] kann man eine Funktion $\chi \in W^1(\Omega)$ finden, so daß für beliebiges $s_h \in \mathcal{U}^h$ gilt

$$\|\bar{u} - v_h\|_{-1,\Omega} = \frac{(\bar{u} - v_h, \, \chi - s_h)_\Omega}{\|\chi\|_{1,\Omega}} \, . \tag{4.36}$$

Wir wählen

$$s_h(x) = \sum_{j \in \mathcal{I}} M_j(\chi) \, \psi_j(x) \, ,$$

und bekommen in (4.36) mit einer Cauchy–Schwarz Ungleichung die für (4.34) ausschlaggebende Abschätzung

$$\|\bar{u} - v_h\|_{-1,\Omega} \leq \frac{\| \, \bar{u} - v_h \, \|_\Omega \, \tilde{C}h \, \|\chi\|_{1,\Omega}}{\|\chi\|_{1,\Omega}} \, .$$

□

Lemma 15 *Sei v_h eine beliebige Approximation aus $\mathcal{U}^h_{ad}$, dann existiert eine Funktion $v^* \in \mathcal{U}_{ad} \cap \bigoplus_{j \in \mathcal{I}} W^1(\Omega_j)$, so daß*

$$M_j(v^*) = M_j(v_h) \qquad \forall \, j \in \mathcal{I}$$

und

$$\|v^*\|^2_{1,\Omega_j} \leq \|\xi_0\|^2_{1,\Omega_j} + \|\xi_1\|^2_{1,\Omega_j} \qquad \forall \, j \in \mathcal{I} \, .$$

Lemma 16 *Sei u_h die Lösung von $\mathbf{P}^h_{i,s}$ und v ergibt sich für u_h als entsprechendes v^* nach Lemma 15, dann existiert eine von ξ_0 , ξ_1 und h unabhängige Konstante C und es gilt*

$$\|u_h - v\|_{-1,\Omega} \leq Ch^2 \sqrt{\|\xi_0\|^2_{1,\Omega} + \|\xi_1\|^2_{1,\Omega}} \, , \tag{4.37}$$

$$\left(\sum_{j \in \mathcal{I}} \|u_h - v\|^2_{-1,\Omega_j} \right)^{1/2} \leq Ch^2 \sqrt{\|\xi_0\|^2_{1,\Omega} + \|\xi_1\|^2_{1,\Omega}} \, . \tag{4.38}$$

Beweis: Die Beziehung (4.37) stimmt mit (17) in [22] überein. Analog wie in Lemma 5 aus [22] stellen wir neue finite Elemente $\tau_j(x)$, $j \in \mathcal{I}$ zusammen. Für beliebiges $j \in \mathcal{I}$ gelte: supp $\tau_j = \Omega_j$ und

$$(\tau_j,\, \psi)_{1,\Omega_j} = (u_h - v,\, \psi)_{\Omega_j} \qquad \forall\, \psi \in W^1(\Omega_j)\ . \tag{4.39}$$

Nach der Definition von $\|\cdot\|_{-1,\mathrm{M}}$ erhält man

$$\|u_h - v\|_{-1,\Omega_j} = \sup_{\phi \in C^\infty(\overline{\Omega}_j)\setminus\{0\}} \frac{(\,u_h - v,\, \phi\,)_{\Omega_j}}{\|\phi\|_{1,\Omega_j}} \quad = \tag{4.40}$$

$$\sup_{\phi \in C^\infty(\overline{\Omega}_j)\setminus\{0\}} \frac{(\,\tau_j,\, \phi\,)_{1,\Omega_j}}{\|\phi\|_{1,\Omega_j}} \le \sup_{\phi \in C^\infty(\overline{\Omega}_j)\setminus\{0\}} \frac{\|\tau_j\|_{1,\Omega_j}\, \|\phi\|_{1,\Omega_j}}{\|\phi\|_{1,\Omega_j}} \quad = \quad \|\tau_j\|_{1,\Omega_j}\ .$$

Aufgrung der Definition von $\tau_j(x)$ in (4.39) erhält man außerdem die Gleichheiten

$$\|\tau_j\|^2_{1,\Omega_j} = (\,u_h - v,\, \tau_j\,)_{\Omega_j} \tag{4.41}$$

und

$$(\,\tau_j,\, 1\,)_{1,\Omega_j} = (\,u_h - v,\, 1\,)_{\Omega_j}\ .$$

Die letzte Gleichung kann auch explizit geschrieben werden

$$\int_{\Omega_j} \tau_j(x)\, dx = \int_{\Omega_j} (u_h(x) - v(x))\, dx\ . \tag{4.42}$$

Aus (4.41) folgt

$$\|\tau_j\|^2_{1,\Omega_j} = \int_{\Omega_j} (u_h(x) - v(x))\, \tau_j(x)\, dx =$$

$$\int_{\Omega_j} (u_h(x) - v(x)) \left(\tau_j(x) - \frac{1}{\text{mes }\Omega_j} \int_{\Omega_j} \tau_j(x)\, dx \right) dx +$$

$$\frac{1}{\text{mes }\Omega_j} \left(\int_{\Omega_j} (u_h(x) - v(x))\, dx \right)^2\ .$$

Der letzte Term bildet sich entsprechend (4.42). Nach der Wahl von $v(x)$ muß er aber verschwinden und es gilt mit (4.16)

$$\|\tau_j\|^2_{1,\Omega_j} \quad = \quad \int_{\Omega_j} (M_j(v^*) - v^*(x))\, (M_j(\tau_j) - \tau_j(x))\, dx \le$$

$$\left(\int_{\Omega_j} (M_j(v^*) - v^*(x))\, dx \right)^{1/2} \left(\int_{\Omega_j} (M_j(\tau_j) - \tau_j(x))\, dx \right)^{1/2}\ .$$

Gemäß der Interpolationstheorie [88], [89] lassen sich die obigen Integrale durch die W^1-Normen der Funktionen v^* und τ_j abschätzen. Schließlich erhählt man

$$\|\tau_j\|^2_{1,\Omega_j} \;\leq\; \tilde{C}h\,\|v^*\|_{1,\Omega_j}\,\tilde{C}h\,\|\tau_j\|_{1,\Omega_j}\;. \tag{4.43}$$

Aus Lemma 15 und (4.43) folgt

$$\|\tau_j\|_{1,\Omega_j} \;\leq\; \tilde{C}^2h^2\,\|v^*\|_{1,\Omega_j} \leq Ch^2\,\sqrt{\|\xi_0\|^2_{1,\Omega_j}+\|\xi_1\|^2_{1,\Omega_j}}\;.$$

Diese Ungleichung zusammen mit (4.40) beweist die Behauptung (4.38).

□

Das folgende Lemma beruht auf den Lemmata 14, 15 und 16.

Lemma 17 *Auf jeder Iteration des MSR-Verfahrens (Abb. 4.4) gilt folgende Beziehung*

$$\left\|u^{i,s}-\bar{u}^{i,s}\right\|_{\Omega} \;\leq\; \varepsilon_i \tag{4.44}$$

mit

$$\varepsilon_i=\sqrt{\frac{2\,o_i}{\chi_i}}+\sqrt{\frac{2\,\eta_i}{\chi_i}}\;, \tag{4.45}$$

wobei

$$\begin{aligned} o_i \;=\; \max_{s=\overline{1,\,s(i)}}\Bigg\{ & C\,\chi_i h_i^2\,\left(\frac{C}{\chi_i^2}+\left\|u^{i,s-1}\right\|^2_{\Omega}+C_1\right)+ \\ & C\,\chi_i h_i^2\left(\left\|u^{i,s-1}\right\|_{\Omega}+\sqrt{\frac{C}{\chi_i^2}+\|u^{i,s-1}\|^2_{\Omega}+C_1}\;+\frac{C_2}{\chi_i}\right)\times \\ & \left(\sqrt{C_1}+\sqrt{\frac{C}{\chi_i^2}+\|u^{i,s-1}\|^2_{\Omega}+C_1}\;\right)+ \\ & C\,h_i^4\,\left(\frac{C}{\chi_i^2}+\left\|u^{i,s-1}\right\|^2_{\Omega}+C_1\right)+\left(C_3\,h_i^2\right)^2+C_4\,h_i^2\Bigg\}\;. \end{aligned} \tag{4.46}$$

Die Werte C, C_1, .., C_4 sind von $\|u^{i,s-1}\|_{\Omega}$, χ_i, η_i und h_i unabhängig.

Beweis: Wir fahren fort mit der Untersuchung der Ungleichung (4.27). Der erste Term auf der rechten Seite in (4.27) läßt sich nach einer verallgemeinerten Cauchy–Schwarz Ungleichung abschätzen. Wir übernehmen diese Technik aus [22]. Es gilt beispielweise für beliebiges $j\in\mathcal{I}$

$$\left(u^{i,s-1},\,v-u_h\right)_{\Omega_j}\leq\left\|u^{i,s-1}\right\|_{1,\Omega_j}\,\|v-u_h\|_{-1,\Omega_j}\;.$$

Wenn man derartige Ungleichungen addiert, ergibt sich aus (4.27) die folgende Beziehung

$$\begin{aligned}
\chi_i \, \|\bar{u} - u_h\|_{\Omega}^2 + \|\bar{y}(\bar{u}) - y_h(u_h)\|_{\Omega'}^2 \quad &\leq \\
\chi_i \sum_{j \in \mathcal{I}} \Big(\left\|u^{i,s-1}\right\|_{1,\Omega_j} \|v - u_h\|_{-1,\Omega_j} + \left\|u^{i,s-1}\right\|_{1,\Omega_j} \|v_h - \bar{u}\|_{-1,\Omega_j} \Big) \quad &+ \\
\chi_i \, \Big((\bar{u},\, v - u_h)_{\Omega} + (\bar{u},\, v_h - \bar{u})_{\Omega} + (u_h - \bar{u},\, v_h - \bar{u})_{\Omega} \Big) \quad &+ \\
(\bar{y}(\bar{u}),\, (\bar{y}(v) - \bar{y}(u_h)) + (\bar{y}(v_h) - \bar{y}(\bar{u})))_{\Omega'} \quad &+ \\
(\bar{y}(\bar{u}),\, (\bar{y}(u_h) - y_h(u_h)) + (y_h(v_h) - \bar{y}(v_h)))_{\Omega'} \quad &+ \\
((y_h(u_h) - \bar{y}(\bar{u})),\, (y_h(v_h) - \bar{y}(v_h)) + (\bar{y}(v_h) - \bar{y}(\bar{u})))_{\Omega'} \quad &.
\end{aligned} \tag{4.47}$$

Für die affine stetige Abbildung $\bar{y}(\cdot) : \mathcal{Y}' \to \mathcal{Y}$ gelten die Abschätzungen (vgl. [22, S. 44])

$$\begin{aligned}
\|\bar{y}(v) - \bar{y}(u_h)\|_{1,\Omega} \quad &\leq \quad (\, 1 + c_F^2 \,) \, \|v - u_h\|_{-1,\Omega} \, , \\
\|\bar{y}(v_h) - \bar{y}(\bar{u})\|_{1,\Omega} \quad &\leq \quad (\, 1 + c_F^2 \,) \, \|v_h - \bar{u}\|_{-1,\Omega} \, .
\end{aligned}$$

Mit den Ungleichungen

$$\begin{aligned}
2 \, (u_h - \bar{u},\, v_h - \bar{u})_{\Omega} \leq \|u_h - \bar{u}\|_{\Omega}^2 + \|v_h - \bar{u}\|_{\Omega}^2 \quad &, \\
2 \, ((y_h(u_h) - \bar{y}(\bar{u})),\, (y_h(v_h) - \bar{y}(v_h)) + (\bar{y}(v_h) - \bar{y}(\bar{u})))_{\Omega'} \quad &\leq \\
2 \, \|y_h(u_h) - \bar{y}(\bar{u})\|_{\Omega'}^2 + \|y_h(v_h) - \bar{y}(v_h)\|_{\Omega'}^2 + \|\bar{y}(v_h) - \bar{y}(\bar{u})\|_{\Omega'}^2 &
\end{aligned}$$

erhält man in (4.47)

$$\begin{aligned}
\chi_i \, \|\bar{u} - u_h\|_{\Omega}^2 + \|\bar{y}(\bar{u}) - y_h(u_h)\|_{\Omega'}^2 \quad &\leq \\
\chi_i \sum_{j \in \mathcal{I}} \left\|u^{i,s-1}\right\|_{\Omega_j} \Big(\|v - u_h\|_{-1,\Omega_j} + \|v_h - \bar{u}\|_{-1,\Omega_j} \Big) \quad &+ \\
\chi_i \sum_{j \in \mathcal{I}} \|\bar{u}\|_{1,\Omega_j} \Big(\|v - u_h\|_{-1,\Omega_j} + \|v_h - \bar{u}\|_{-1,\Omega_j} \Big) \quad &+ \\
\frac{\chi_i}{2} \Big(\|u_h - \bar{u}\|_{\Omega}^2 + \|v_h - \bar{u}\|_{\Omega}^2 \Big) \quad &+ \\
\|\bar{y}(\bar{u})\|_{\Omega} \, (\, \|\bar{y}(v) - \bar{y}(u_h)\|_{\Omega} + \|\bar{y}(v_h) - \bar{y}(\bar{u})\|_{\Omega}) \quad &+ \\
\|y_h(u_h) - \bar{y}(\bar{u})\|_{\Omega'}^2 + \frac{1}{2} \, \|y_h(v_h) - \bar{y}(v_h)\|_{\Omega'}^2 + \frac{1}{2} \, \|\bar{y}(v_h) - \bar{y}(\bar{u})\|_{\Omega'}^2 \quad &+
\end{aligned} \tag{4.48}$$

$$(\bar{y}(\bar{u}),\, (\bar{y}(u_h) - y_h(u_h)) + (y_h(v_h) - \bar{y}(v_h)))_{\Omega'} \ .$$

Man berücksichtigt außerdem, daß die Werte $\|u^{i,s-1}\|_{\Omega_j}$ und $\|u^{i,s-1}\|_{1,\Omega_j}$ zusammenfallen. Der erste und zweite Term auf der rechten Seite der Beziehung (4.48) läßt die Anwendung der Cauchy–Schwarz Ungleichung für endliche Summen zu. Im vierten Term berücksichtigt man, daß die L_2-Normen durch W^1-Normen majorisiert werden können. Nach allen Umformungen ergibt sich

$$\begin{aligned}
\frac{\chi_i}{2}\,\|\bar{u}-u_h\|^2_{\Omega} \le \frac{\chi_i}{2}\,\|v_h-\bar{u}\|^2_{\Omega} \quad &+ \\
\left(\,\chi_i\,\left\|u^{i,s-1}\right\|_{\Omega} + \chi_i\,\left(\sum_{j\in\mathcal{I}}\|\bar{u}\|^2_{1,\Omega_j}\right)^{1/2}\right)\,T_{\Delta} \quad &+ \\
C\,\|\bar{y}(\bar{u})\|_{\Omega}\,(\,\|v-u_h\|_{-1,\Omega} + \|v_h-\bar{u}\|_{-1,\Omega}\,) \quad &+ \\
\frac{1}{2}\,\|y_h(v_h)-\bar{y}(v_h)\|^2_{\Omega} + \frac{C^2}{2}\,\|v_h-\bar{u}\|^2_{-1,\Omega} \quad &+ \qquad (4.49)\\
\|\bar{y}(\bar{u})\|_{\Omega}\left(\|y_h(u_h)-\bar{y}(u_h)\|_{\Omega} + \|y_h(v_h)-\bar{y}(v_h)\|_{\Omega}\right) & \\
\forall\, v\in\mathcal{U}_{ad}\ , \quad \forall\, v_h\in\mathcal{U}^h_{ad} &
\end{aligned}$$

mit

$$T_{\Delta} \;=\; \left(\sum_{j\in\mathcal{I}}\|v-u_h\|^2_{-1,\Omega_j}\right)^{1/2} + \left(\sum_{j\in\mathcal{I}}\|v_h-\bar{u}\|^2_{-1,\Omega_j}\right)^{1/2} \ .$$

Wir wählen nun in (4.49) $v = v^*(u_h) \in \mathcal{U}_{ad}$ nach Lemma 16 und

$$v_h = \sum_{j\in\mathcal{I}} M_j(\bar{u})\,\psi_j \in \mathcal{U}^h_{ad}$$

nach Lemma 14. Auf die letzten Glieder in (4.49) wenden wir zugleich die erhaltene Abschätzung des Diskretisierungsfehlers der Zustandsgleichung (4.12) an. Es ergibt sich

$$\begin{aligned}
\frac{\chi_i}{2}\,\|\bar{u}-u_h\|^2_{\Omega} \;\le\; \frac{\chi_i}{2}\,C\,h_i^2\sum_{j\in\mathcal{I}}\|\bar{u}\|^2_{1,\Omega_j} \quad &+ \\
\left(\,\chi_i\,\left\|u^{i,s-1}\right\|_{\Omega} + \chi_i\,\left(\sum_{j\in\mathcal{I}}\|\bar{u}\|^2_{1,\Omega_j}\right)^{1/2} + C\,\|\bar{y}\|_{\Omega}\right) \quad &\times \\
\left(\,C\,h_i^2\sqrt{\|\xi_0\|^2_{1,\Omega}+\|\xi_1\|^2_{1,\Omega}} + C\,h_i^2\left(\sum_{j\in\mathcal{I}}\|\bar{u}\|^2_{1,\Omega_j}\right)^{1/2}\right) \quad &+
\end{aligned}$$

$$C\,h_i^4 \sum_{j\in\mathcal{I}} \|\bar{u}\|^2_{1,\Omega_j} + \Big(Ch_i^2 \;\|f+v_h\|_\Omega + Ch_i^2\,T_\Sigma \Big)^2 \quad + \tag{4.50}$$

$$\|\bar{y}\|_\Omega\, Ch_i^2 \;(\; \|f+v_h\|_\Omega + \|f+u_h\|_\Omega + 2\,T_\Sigma \;)$$

mit

$$T_\Sigma \;=\; \sum_{m\in\mathcal{M}_N} \|g_m\|_{1/2,\,\Gamma_m} \;+$$

$$\sum_{m\in\mathcal{M}_A} T_m^{1/2}(g_m) + \sum_{m\in\mathcal{M}_B} T_m^{1/2}(g_{m+1}) + \sum_{m\in\mathcal{M}_C} T_m^{1/2}(g_m,\, g_{m+1})\;.$$

Analog mit [22, S. 45] erhält man Abschätzungen

$$\|u_h\|_\Omega \le \sqrt{\|\xi_0\|^2_\Omega + \|\xi_1\|^2_\Omega}\;, \qquad \|v_h\|_\Omega \le \sqrt{\|\xi_0\|^2_\Omega + \|\xi_1\|^2_\Omega}\;.$$

Die Größen $\|\bar{u}\|_{1,\Omega_j}$ in (4.50) lassen sich nach Lemma 12 und die Größe $\|\bar{y}\|_\Omega$ läßt sich mit (4.30) abschätzen. Endgültig mit neuen Konstanten C, C', .. gilt

$$\frac{\chi_i}{2}\,\|\bar{u}-u_h\|^2_\Omega \le C\chi_i h_i^2 \left(\frac{C}{\chi_i^2} + \|u^{i,s-1}\|^2_\Omega + \bar{C}\right) \quad +$$

$$Ch_i^2\left(\chi_i \left\|u^{i,s-1}\right\|_\Omega + \chi_i\sqrt{\frac{C}{\chi_i^2} + \|u^{i,s-1}\|^2_\Omega + \bar{C}} + C\,\tilde{c}_y\right) \quad \times \tag{4.51}$$

$$\left(\sqrt{\bar{C}} + \sqrt{\frac{C}{\chi_i^2} + \|u^{i,s-1}\|^2_\Omega + \bar{C}}\;\right) +$$

$$Ch_i^4\left(\frac{C}{\chi_i^2} + \|u^{i,s-1}\|^2_\Omega + \bar{C}\right) + \left(C'h_i^2\right)^2 + \tilde{c}_y C' h_i^2 \quad .$$

Es ist nicht schwierig zu sehen, daß die Definition von o_i in (4.46) unmittelbar auf (4.51) beruht. Den Ausdruck (4.45) für den Fehler ε_i im Ersatzproblem des MSR-Verfahrens erhält man mit Lemma 11 nach der Dreiecksungleichung. Aus (4.51) folgt

$$\varepsilon_i = \sqrt{\frac{2\,o_i}{\chi_i}} + \sqrt{\frac{2\,\eta_i}{\chi_i}} \;\;\ge\;\; \|\bar{u}^{i,s} - u_h^{i,s}\|_\Omega + \|u^{i,s} - u_h^{i,s}\|_\Omega \ge\; \|\bar{u}^{i,s} - u^{i,s}\|_\Omega\;.$$

□

An dieser Stelle sind wir mit unseren Vorbereitungen fertig und formulieren den zu beweisenden Konvergenzsatz.

Satz 19 *Die Folge der Näherungen* $z^{i,s} = (y^{i,s},\, u^{i,s})$ *nach dem diskretisierten MSR-Verfahren hat immer endlich viele Glieder in* s*:*

$$s(i) \le (d_2 + \varepsilon_i)/\vartheta_i \tag{4.52}$$

und konvergiert **schwach** *im Raum* $W^1(\Omega) \times L_2(\Omega)$ **gegen ein Element** *aus der Lösungsmenge* $\mathcal{Z}^*$ *, wenn*

$$0 \le \underline{\chi} < \chi_i \le \bar{\chi}\,, \qquad 0 < \vartheta_i \le \bar{\vartheta}\,,$$

$$\sum_{i=1}^{\infty} \sqrt{\frac{\eta_i}{\chi_i}} < \infty\,, \qquad \sum_{i=1}^{\infty} \frac{h_i}{\chi_i} < \infty\,, \tag{4.53}$$

$$\delta_i \ge \varepsilon_i + \sqrt{2\,(d_2 + \varepsilon_i)\,(\varepsilon_i + \vartheta_i)}\;, \tag{4.54}$$

wobei ε_i *nach Lemma 17 definiert ist und*

$$d_2 \;=\; 2 \sup_{v \in \mathcal{U}_{ad}} \|v\|_\Omega\,. \tag{4.55}$$

Beweis: Wenn man den Parameter δ_i im MSR-Verfahren (s. S. 107) hinreichend groß wählt, so gilt

$$\left\|\bar{u}^{i,s} - u^{i,s-1}\right\|_\Omega \ge \left\|u^{i,s} - u^{i,s-1}\right\|_\Omega - \left\|u^{i,s} - \bar{u}^{i,s}\right\|_\Omega \ge (\delta_i - \varepsilon_i) > 0\,. \tag{4.56}$$

Wir wenden jetzt das Lemma 8 auf die i-te Iteration des MSR-Verfahrens an. Als Raum $\mathcal{X}$ wählen wir unseren Prozeßraum $\mathcal{Z}$, d.h. $\|\cdot\|_{\mathcal{X}} := \|\cdot\|_{\mathcal{Z}}$. Man führe desweiteren folgende Substitutionen durch:

$$\begin{aligned}
&\mathcal{X}_1 := \{\, z = (y,\, u) \in \mathcal{Z} : y = 0 \;\text{ in } \mathcal{Y}\,\}\,, \quad \Phi(z) := 2\,J(z)\,,\\
&b(z_1,\, z_2) := 0\,, \qquad a(z_1,\, z_2) := (y_1,\, y_2)_{\Omega'}\;, \qquad \ell(z) := 0\,,\\
&\mathcal{G} := \{(y,\, u) \in \mathcal{Y} \times \mathcal{U}_{ad} :\; -\bar{\Delta}\, y = u + f + \sum_{m \in \mathcal{M}_N} g_m\, \delta_{\Gamma_m} \quad \text{in } \mathcal{Y}'\,\}\,,\\
&a^0 := z^{i,s-1}\,, \quad a^1 := \bar{z}^{i,s}\,, \quad x := z^*\,, \quad \delta := \delta_i - \varepsilon_i\,, \quad \chi := \chi_i\,.
\end{aligned}$$

Um Lemma 8 anwenden zu können, sind die Abgeschlossenheit und Konvexität von $\mathcal{G}$ im Raum $\mathcal{X} = \mathcal{Z} = \mathcal{Y} \times \mathcal{U}$ zu zeigen. Durch die Zustandsgleichung in $\mathcal{G}$ wird eine affine abgeschlossene Mannigfaltigkeit in $\mathcal{Y} \times \mathcal{Y}'$ erzeugt, da der verallgemeinerte Laplace-Operator einen isometrischen Isomorphismus zwischen $\mathcal{Y}$ und $\mathcal{Y}'$ festlegt. Somit ist der Schnitt dieser Mannigfaltigkeit mit dem topologisch stärkeren Raum $\mathcal{Z}$ in $\mathcal{X}$ abgeschlossen. Die Abgeschlossenheit der ganzen Menge $\mathcal{G}$ gilt wegen des zusätzlichen Schnitts der affinen Menge mit der in $\mathcal{X}$ abgeschlossenen Menge $\mathcal{Y} \times \mathcal{U}_{ad}$.

Diese zwei Mengen sind außerdem konvex, somit gilt auch die Konvexität für $\mathcal{G}$. Aufgrund von (2.124) liegt folgende Normäquivalenz vor:

$$|(y,\, u)|_{\mathcal{X}} = \|u\|_{\Omega} \ . \tag{4.57}$$

Die im Abschnitt 2.3.2 geforderte Gleichwertigkeit zwischen $|\cdot|_{\mathcal{X}}$ und $\|\cdot\|_{\mathcal{X}}$ ist hier nicht mehr nötig, da der Konvergenzbeweis direkt für die Halbnorm geführt wird. Weil $\Phi(x) = 2\,J(z^*) \le 2\,J(\bar{z}^{i,s}) = \Phi(a^1)$, gilt $\eta(x) = 0$ in Lemma 8. Nach (2.128) erhält man

$$\begin{aligned} |\bar{z}^{i,s} - z^*|_{\mathcal{X}} \ &\le \ |z^{i,s-1} - z^*|_{\mathcal{X}} - \frac{(\delta_i - \varepsilon_i)^2}{2\,(d_2 + \varepsilon_i)} \le \\ &|\bar{z}^{i,s-1} - z^*|_{\mathcal{X}} + |z^{i,s-1} - \bar{z}^{i,s-1}|_{\mathcal{X}} - \frac{(\delta_i - \varepsilon_i)^2}{2\,(d_2 + \varepsilon_i)} \ . \end{aligned} \tag{4.58}$$

Das Resultat im Nenner von (4.58) ergibt sich mit der Beachtung (4.55) nach folgender Überlegung

$$\begin{aligned} &|z^{i,s-1} - z^*|_{\mathcal{X}} \le |z^{i,s-1} - \bar{z}^{i,s-1}|_{\mathcal{X}} + |\bar{z}^{i,s-1} - z^*|_{\mathcal{X}} \le \\ &\varepsilon_i \ + \sup_{(v_1,\, v_2) \in \mathcal{U}_{ad}^2} \|v_1 - v_2\|_{\Omega} = \varepsilon_i + d_2 \ . \end{aligned}$$

Aus (4.57) und (4.58) erhält man für $s = 1, .. \, s(i) - 1$

$$\begin{aligned} |z^{i,s} - z^*|_{\mathcal{X}} - |z^{i,s-1} - z^*|_{\mathcal{X}} \ &\le \\ |\bar{z}^{i,s} - z^{i,s}|_{\mathcal{X}} - \frac{(\delta_i - \varepsilon_i)^2}{2\,(d_2 + \varepsilon_i)} \ &\le \\ \varepsilon_i - \frac{(\delta_i - \varepsilon_i)^2}{2\,(d_2 + \varepsilon_i)} \ &\le \ -\vartheta_i < 0 \ . \end{aligned} \tag{4.59}$$

Daraus folgt, daß die Werte $|z^{i,s} - z^*|_{\mathcal{X}}$ mit einer absteigenden arithmetischen Reihe (in s) majorisiert werden. Da die zu betrachtenden Größen ϑ_i positiv sind und es gilt $|z^{i,s} - z^*|_{\mathcal{X}} \le d_2 + \varepsilon_i$, ist die Anzahl der s-Schritte immer endlich und man erhält dabei die Abschätzung (4.52). Gleichzeitig bekommt man aus (4.59) eine untere Schranke für den Parameter δ_i , d.h. (4.54) . Nach (2.127) mit $\eta(x) = 0$ ist

$$|z^{i,s} - z^*|_{\mathcal{X}} - |z^{i,s-1} - z^*|_{\mathcal{X}} \le \varepsilon_i \ . \tag{4.60}$$

Da die Werte $|z^{i,s} - z^*|_{\mathcal{X}}$ nach (4.59) bei $s = 0, .. \, s(i) - 1$ monoton fallend sind, können wir (4.60) über alle $s = 1, .. \, s(i)$ fortsetzen. Unter Beachtung der Identität $z^{i+1,0} = z^{i,s(i)}$ für alle i , erhält man eine von s unabhängige Abschätzung:

$$|z^{i+1,0} - z^*|_{\mathcal{X}} - |z^{i,0} - z^*|_{\mathcal{X}} \le \varepsilon_i \ . \tag{4.61}$$

Falls die unendliche Summe aus allen ε_i konvergiert, so folgt aus Lemma 2.2.2 in [80] die Konvergenz (in i) von $|z^{i,0} - z^*|_{\mathcal{X}}$ gegen einen positiven endlichen Wert. Die Folge der exakten Lösungen $|\bar{z}^{i,0} - z^*|_{\mathcal{X}}$ konvergiert gegen denselben Wert, denn $\varepsilon_i \to 0$.

Sei $\{\bar{z}^{i_k,\,0}\}$ eine beliebige schwach konvergierende Teilfolge von $\{\bar{z}^{i,0}\}$ mit einem schwachen Häufungspunkt $z_3 \in \mathcal{G} \subset \mathcal{Z}$.[2] Wir benutzen jetzt die Beziehung (2.126) mit $s := s(i)$

$$|\bar{z}^{i,s(i)} - z^*|^2_{\mathcal{X}} - |z^{i,s(i)-1} - z^*|^2_{\mathcal{X}} \leq \frac{4}{\chi_i} \left(J(z^*) - J(\bar{z}^{i,s(i)}) \right) \tag{4.62}$$

und nach der rückwärtigen Fortsetzung des Indexes s wie für (4.60) erhalten wir folgende Beziehungen zwischen negativen Werten:

$$\begin{aligned}
J(z^*) - J(\bar{z}^{i+1,0}) = J(z^*) - J(\bar{z}^{i,s(i)}) \quad &\geq \\
\frac{\chi_i}{4} \left(|\bar{z}^{i,s(i)} - z^*|_{\mathcal{X}} + |z^{i,s(i)-1} - z^*|_{\mathcal{X}} \right) \left(|\bar{z}^{i,s(i)} - z^*|_{\mathcal{X}} - |z^{i,s(i)-1} - z^*|_{\mathcal{X}} \right) \quad &\geq \\
\frac{\bar{\chi}}{4} \left(|\bar{z}^{i,s(i)} - z^*|_{\mathcal{X}} + |z^{i,s(i)-1} - z^*|_{\mathcal{X}} \right) \left(|\bar{z}^{i+1,0} - z^*|_{\mathcal{X}} - |z^{i,0} - z^*|_{\mathcal{X}} \right) \quad &\geq \\
\frac{\bar{\chi}}{4} \left(d_2 + d_2 + \varepsilon_i \right) \left(|\bar{z}^{i+1,0} - z^*|_{\mathcal{X}} - |z^{i,0} - z^*|_{\mathcal{X}} \right) \quad .
\end{aligned}$$

Die Glieder in der letzten Klammer heben sich im Grenzwert gegenseitig auf und wir erhalten mit $i_k \to \infty$, daß $J(z^*) \geq J(z_3)$, da $J(z)$ schwach unterhalbstetig ist. Daraus folgt $z_3 = (y_3,\, u_3) \in \mathcal{Z}^* = \mathcal{Y}^* \times \mathcal{U}^*$. Die Teilfolge $\{z^{i_k,\,0}\}$ konvergiert schwach gegen dasselbe Element z_3 , denn $\varepsilon_{i_k} \to 0$. Zuletzt wenden wir auf die Folge $u^{i,0}$ und auf die Menge $\mathcal{U}^*$ das Lemma 1 in [79] an, woraus man die schwache Konvergenz der Steuerkomponenten des MSR-Verfahrens gegen das Element u_3 folgert.

Jetzt kann man **die starke Konvergenz** der Zustandskomponenten $y^{i,0}$ gegen y_3 bezüglich der W^1-Norm beweisen. Es gilt nach der Dreiecksungleichung

$$\begin{aligned}
\left\| y^{i,0} - y_3 \right\|_{1,\,\Omega} \quad \leq \quad & \left\| \bar{y}^{i,0} - y^{i,0} \right\|_{1,\,\Omega} + \left\| \bar{y}(\bar{u}^{i,0}) - \bar{y}(u_3) \right\|_{1,\,\Omega} \leq \\
& \|\bar{y}(\bar{u}) - \bar{y}(u_h)\|_{1,\,\Omega} + \|\bar{y}(u_h) - y_h(u_h)\|_{1,\,\Omega} + \\
& \|y^{i,s} - y_h^{i,s}\|_{1,\,\Omega} + \|\bar{y}(\bar{u}) - \bar{y}(u_3)\|_{1,\,\Omega} \leq \\
& \sqrt{1 + c_F^2}\, \|\bar{u} - u_h\|_{-1,\,\Omega} + C_5\, h_i^2 + \\
& \sqrt{1 + c_F^2}\, \sqrt{\frac{2\,\eta_i}{\chi_i}} + (\,1 + c_F^2\,) \left\| \bar{u}^{i,0} - u_3 \right\|_{-1,\,\Omega} \leq
\end{aligned}$$

[2] Es ist hier wesentlich, daß die Menge $\mathcal{G}$ im Raum $\mathcal{Z}$ beschränkt ist.

$$C_6 \sqrt{\frac{2\,o_i}{\chi_i}} + C_5\, h_i^2 + \sqrt{1+c_F^2}\,\sqrt{\frac{2\,\eta_i}{\chi_i}} +$$

$$(1+c_F^2)^{3/2}\,\left\|\bar{u}^{i,0} - u^{i,0}\right\|_{\Omega} + (\,1+c_F^2\,)\,\left\|u^{i,0}-u_3\right\|_{-1,\,\Omega} \leq$$

$$C_6 \sqrt{\frac{2\,o_i}{\chi_i}} + C_5\, h_i^2 + \sqrt{1+c_F^2}\,\sqrt{\frac{2\,\eta_i}{\chi_i}} + (1+c_F^2)^{3/2}\,\varepsilon_i +$$

$$(\,1+c_F^2\,)\,\left\|u^{i,0}-u_3\right\|_{-1,\,\Omega}\ .$$

Der Term $C_5\, h_i^2$ ergibt sich mit Rücksicht auf (4.12). Es reicht hin, zu zeigen, daß der letzte Normterm in der oben stehenden Kette gegen Null strebt, falls $i \to \infty$. Dafür definieren wir neue Elemente y_3^i aus $W^1(\Omega)$ nach der folgenden Variationsgleichung

$$\left(y_3^i,\,\phi\right)_{1,\,\Omega} \;=\; \left(u^{i,0}-u_3,\,\phi\right)_{\Omega} \quad \forall\,\phi \in W^1(\Omega)\ .$$

Man setzt $\phi := y_3^i$ und betrachtet den Limes $i \to \infty$

$$\left\|y_3^i\right\|_{1,\,\Omega}^2 \;=\; \left(\,u^{i,0}-u_3,\,y_3^i\,\right)_{\Omega} \to 0\ ,$$

weil $u^{i,0} \xrightarrow{w} u_3$ gilt. Das bedeutet

$$\|u^{i,0}-u_3\|_{-1,\,\Omega} = \sup_{\phi \in C^\infty(\overline{\Omega})\setminus\{0\}} \frac{(\,u^{i,0}-u_3,\,\phi\,)_{\Omega}}{\|\phi\|_{1,\,\Omega}} \;=$$

$$\sup_{\phi \in C^\infty(\overline{\Omega})\setminus\{0\}} \frac{(\,y_3^i,\,\phi\,)_{1,\,\Omega}}{\|\phi\|_{1,\,\Omega}} \leq \|y_3^i\|_{1,\Omega} \to 0\ .$$

Um die Konvergenzbedingungen (4.53) zu bekommen, muß man in (4.45), (4.46) den Limes $h_i \to 0$ betrachten. Die schwächsten Bedingungen ergeben sich bei der Annahme: $\chi = 0$. Anhand der O-Notationen untersuchen wir die Glieder in (4.45) und (4.46):

$$\varepsilon_i \leq O\left(\sqrt{\frac{o_i}{\chi_i}}\right) + O\left(\sqrt{\frac{\eta_i}{\chi_i}}\right) =$$

$$\underline{O\left(\sqrt{\frac{\eta_i}{\chi_i}}\right)} + \frac{1}{\sqrt{\chi_i}}\,\left(\underline{O\left(\frac{h_i^2}{\chi_i}\right)} + O\left(\,\chi_i h_i^2\,\left\|u^{i,s-1}\right\|_{\Omega}^2\right) +\right.$$

$$\left(O\left(\,\chi_i h_i^2\,\left\|u^{i,s-1}\right\|_{\Omega}\right) + O\left(\,h_i^2\,\right)\right)\,\left(O\left(\frac{1}{\chi_i}\right) + O\left(\,\left\|u^{i,s-1}\right\|_{\Omega}\right)\right)$$

$$\left.O\left(\frac{h_i^4}{\chi_i^2}\right) + O\left(\,h_i^4\,\left\|u^{i,s-1}\right\|_{\Omega}^2\right) + O\left(\,h_i^4\,\right) + O\left(\,h_i^2\,\right)\right)^{1/2}\ .$$

Man kann leicht feststellen, daß die unterstrichenen Glieder majorisierend sind und daß aus (4.53) die Konvergenz der unendlichen Summe aller ε_i folgt.

□

Satz 19 ergibt ein a priori Kriterium der Konvergenz der inneren Approximationen des MSR-Verfahrens im Steuerprozeßraum. Dieses für den ersten Augenblick gewöhnliche Resultat ist bei der Behandlung inkorrekt gestellter Kontrollprobleme besonders wichtig, da man in diesem Fall weder über Stopkriterien noch über andere Indikatoren für die Konvergenz verfügt. Die Lösungsmenge ist meistens instabil. Es kann durchaus passieren, daß die Folge der Approximationen irgendwo "daneben landet" oder keine Konvergenz gegen ein Element vorliegt. Wenn die Bedingungen von Satz 19 aber gelten, kann beides nicht mehr auftreten und man könnte sogar im günstigen Fall auch traditionelle hinreichende Kriterien der starken Konvergenz nachprüfen, wie zum Beispiel die Beschränktheit der unendlichen Summe von Abständen, welche zwischen benachbarten Iterierten des MSR-Verfahrens in der Norm des Prozeßraumes berechnet werden. Zuzüglich letzter Betrachtung bekommt man auch ein a postoriori Kriterium der starken Konvergenz. Für die schwächsten Voraussetzungen, unter denen man ein a priori Kriterium der starken Konvergenz formulieren kann, verweisen wir auf [38].

Abschließend sei nur darauf hingewiesen, daß in der OSR-Variante des Algorithmus der Schlupfparameter ϑ_i nicht betrachten zu werden braucht.

4.4 Numerische Behandlung der diskretisierten Ersatzprobleme

Die numerische Untersuchung des MSR-Verfahrens aus Abschnitt 4.2 führen wir mit einem einfachen Modellproblem vom Typ (4.13) durch:

$$\begin{aligned} J(y,\, u) = \frac{1}{2}\, \|y\|_{\Omega}^2 \xrightarrow{\;!\;} & \min_{(y,\, u) \in\, \mathcal{Y} \times \mathcal{U}_{ad}} \\ -\,\bar{\Delta}\, y = u + g\, \delta_{\Gamma} \qquad & \text{in } \mathcal{Y}' \end{aligned} \tag{4.63}$$

mit folgenden Daten

$$\mathcal{Y} = \left\{\, y \in W^1(\Omega) :\ \gamma_0\, y = 0 \quad \text{in } H^{1/2}(\Gamma_D) \,\right\},$$

$$\mathcal{K} = \mathcal{U} = L_2(\Omega)\,, \qquad \mathcal{U}_{ad} = \{u \in \mathcal{U} :\ -1 \le u(x) \le 1 \ \text{ f.ü. in } \Omega\}$$

wobei Ω ein Quadratgebiet ist und $\Gamma = \partial\Omega \setminus \Gamma_D$ eine Seite des Quadrats darstellt. (Die Seiten mit Dirichlet-Randbedingungen sind in Abbildung 4.6 mit einer Doppellinie gezeichnet). Es handelt sich um eine "optimale Reflektierung" einer Delta-Schicht anhand der verteilten Steuerung $u(x)$, deren Amplitude von oben und von

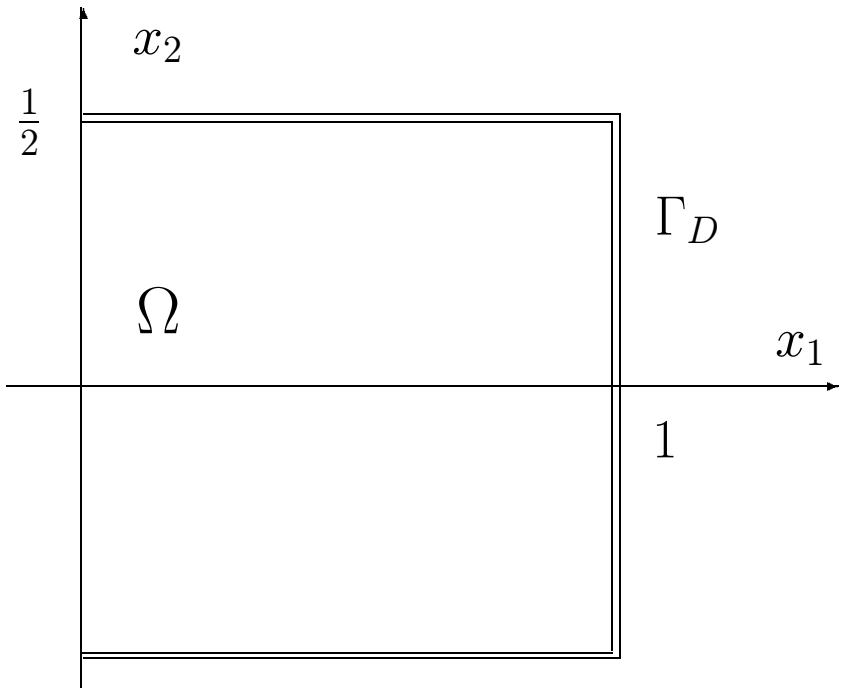


Abbildung 4.6: Das Quadratgebiet

unten beschränkt ist. Es mag sicher unterschiedliche Anwendungen für eine solche Formulierung geben. Die einfachste physikalische Auslegung wäre, wenn man ein homogenes quadratförmiges Bauteil betrachtet, welches von einer Seite mit einer fixierten Wärmequelle erhitzt wird. Die Temperatur der übrigen Seiten wird auf dem Null-Niveau erhalten. Man sucht ein optimales Abkühlungssystem im inneren des Bauteils aufzubauen, daß das Temperaturpotential über dem Quadratgebiet so gering wie möglich vom Null-Niveau auslenkt. Die Konstruktion des Abkühlungssystems wird nur durch die Form der Schaltlinien der optimalen Steuerfunktion $u^*(x)$ bestimmt. Das Problem (4.63) stellt somit ein zweidimensionales Analogon zum klassischen Fuller–Problem dar und man kann erwarten, daß die Anzahl der Schaltlinien in der optimalen Lösung unendlich ist. Die Schaltlinien sammeln sich in der Umgebung einer Kurve, welche ein Analogon zum singulären Punkt bildet. Die Rekonstruktion solcher Lösungen ist vor allem deshalb schwierig, weil die schmalsten Bang-Bang-Formationen in $u^*(x)$ zwischen den benachbarten Schaltungen extrem instabil werden. Ein Beispiel einer solchen Lösung für $g(x_2) = 0.7 \cos \pi x_2$ sieht man auf der Abbildung 4.7.

Die diskretisierten Ersatzprobleme vom Typ (4.17) lassen sich in Form eines quadratischen Optimierungsproblems (QP) in endlichdimensionalen Räumen formulieren. Die freien Variablen sind die Gewichte der finiten Basisfunktionen; diese Gewichte bestimmen anschließend die gesuchten Näherungen $y^{i,s}$ und $u^{i,s}$. Statt der Dreieckszerlegung führen wir eine Rechteckseinteilung des Quadratgebiets durch und statt der "Dach"-Funktionen $\varphi_j(x)$ für $\mathcal{Y}^h$ verwenden wir die bilinearen Ansätze

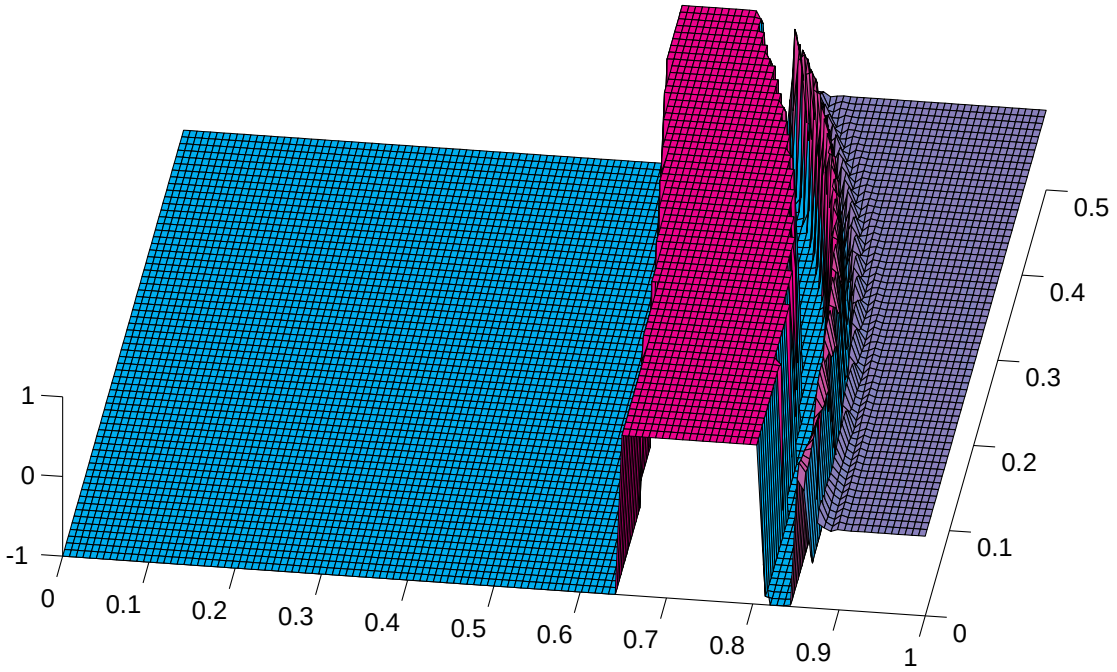


Abbildung 4.7: "Chattering regime" auf dem Quadratgebiet für die optimale Steuerfunktion $u^*(x_1,\ x_2)$

$q_j(x)$ [66]. Auch mit den bilinearen Basisfunktionen kann man die Abschätzung vom Typ (4.12) erreichen [66, Kap. 2, § 5], d. h. die Konvergenzresultate von Abschnitt 4.3 gelten wieder. Gemäß dem MSR-Verfahren kommt man zu folgendem System gekoppelter QP-Probleme:

$$\hat{\Psi}_{i,s}(\mathbf{y},\,\mathbf{u}) \;=\; \frac{1}{2}\,\mathbf{y}^{\mathrm{T}}\,H_y\,\mathbf{y} + \frac{\chi_i}{2}\,\|\mathbf{u}-\mathbf{u}^{i,s-1}\|^2 \to \min_{\mathbf{y},\,\mathbf{u}}$$

$$\mathbf{QP}(\chi_i): \qquad A_y\,\mathbf{y} = A_u\,\mathbf{u} + \mathbf{f}\ ,$$

$$\hat{\mathcal{U}}_{ad}: \qquad -\,\mathbf{e} \le \mathbf{u} \le \mathbf{e}\ ,$$

wobei alle Komponenten des Vektors $\mathbf{e}$ gleich Eins sind. Der Zusammenhang zwischen den Problemen entsteht durch den Proximal-Term und zugleich dadurch, daß man die vorhergehende Näherung als Startpunkt für das nächste QP-Problem benutzt. Der ganze Komplex wird allein mit dem Regularisierungsparameter χ_i gesteuert, was die Einstellung solcher von Strafansätzen freien Verfahren wesentlich vereinfacht.

Der Inhalt der Matrizen und Vektoren in jedem **QP** läßt sich wie folgt beschreiben:

$$(H_y)_{lk} = (q_l,\,q_k)_\Omega\ , \qquad (A_y)_{lk} = (D_{x_1}\,q_l,\,D_{x_1}\,q_k)_\Omega + (D_{x_2}\,q_l,\,D_{x_2}\,q_k)_\Omega\ ,$$

$$(A_u)_{jk} = (\psi_j,\, q_k)_\Omega \;; \qquad (\mathbf{f})_k = (q_k,\, g)_\Gamma \;.$$

Jedes quadratische Optimierungsproblem mit linearen Restriktionen und einfachsten "box-constraints" wird mit einer *dualen Methode* behandelt. Diese Methode besteht darin, eine *duale Funktion* $d(\mathbf{p})$ im Raum der Multiplikatoren $\mathbf{p}$ zu maximieren. Um die duale Funktion einzuführen, betrachten wir die entsprechende Lagrange-Funktion

$$\hat{\mathrm{L}}_{i,s}(\mathbf{y},\, \mathbf{u};\, \mathbf{p}) \;=\; \frac{1}{2}\,\mathbf{y}^{\mathrm{T}}\,H_y\,\mathbf{y} + \frac{\chi_i}{2}\,\|\mathbf{u}-\mathbf{u}^{i,s-1}\|^2 + \mathbf{p}^{\mathrm{T}}\,(A_y\,\mathbf{y} - A_u\,\mathbf{u} - \mathbf{f})\;,$$

und vermerken, daß die Abbildung $\mathbf{p} \to (\mathbf{y}_*,\, \mathbf{u}_*)$,

$$(\mathbf{y}_*(\mathbf{p}), \mathbf{u}_*(\mathbf{p})) \;=\; \operatorname*{argmin}_{(\mathbf{y},\,\mathbf{u})\in\mathbf{R}^J\times\hat{\mathcal{U}}_{ad}} \hat{\mathrm{L}}_{i,s}(\mathbf{y},\, \mathbf{u};\, \mathbf{p}) \tag{4.64}$$

eindeutig ist, weil $\chi_i > 0$ gilt. Jetzt kann man die Werte der dualen Funktion und die Werte ihres Gradients folgendermaßen berechnen:

$$d_{i,s}(\mathbf{p}) \;=\; \hat{\mathrm{L}}_{i,s}(\mathbf{y}_*(\mathbf{p}), \mathbf{u}_*(\mathbf{p}); \mathbf{p})\;, \tag{4.65}$$

$$\vec{\nabla}_{\mathbf{p}}\, d_{i,s}(\mathbf{p}) \;=\; A_y\,\mathbf{y}_*(\mathbf{p}) - A_u\,\mathbf{u}_*(\mathbf{p}) - \mathbf{f}\;. \tag{4.66}$$

Die Berechnung erfolgt leicht, weil sich die Abbildungen $\mathbf{y}_*(\mathbf{p})$ und $\mathbf{u}_*(\mathbf{p})$ explizit ausdrücken lassen:

$$\mathbf{y}_*(\mathbf{p}) = H_y^{-1}A_y^{\mathrm{T}}\,\mathbf{p}\;, \qquad \mathbf{u}_*(\mathbf{p}) = \mathbf{w}_- + \mathbf{w}_+ + \mathbf{w}_0\;,$$

wobei

$$\begin{aligned}
(\mathbf{w}_0)_j &= (\mathbf{u}^{i,s-1})_j + \frac{(A_u^{\mathrm{T}}\,\mathbf{p})_j}{\chi_i}\;,\\
(\mathbf{w}_-)_j &= \max\{\,0,\, -(\mathbf{w}_0)_j - 1\,\}\;,\\
(\mathbf{w}_+)_j &= \max\{\,0,\, -(\mathbf{w}_0)_j + 1\,\}\;.
\end{aligned}$$

Desweiteren ist wesentlich, daß die oben definierte duale Funktion ein eiziges Maximum besitzt und in seier Umgebung stark konkav ist. Es liegt nämlich folgende Abschätzung vor:

$$d_{i,s}^* - d_{i,s}(\mathbf{p}) \;\geq\; \frac{1}{2}\,(\mathbf{p}-\mathbf{p}^*)^{\mathrm{T}}\,A_y\,H_y^{-1}A_y^{\mathrm{T}}\,(\mathbf{p}-\mathbf{p}^*)\;, \tag{4.67}$$

wobei $\mathbf{p}^*$ der Optimalpunkt bezüglich der dualen Variablen darstellt. Es ist offensichtlich, daß die Matrix $A_y\,H_y^{-1}A_y^{\mathrm{T}}$ streng positiv definiert ist.

Zur Maximierung der dualen Funktion benutzt man die Routine e04dgc aus der NAG-Bibliothek [75]. Dieses Standardprogramm minimiert (bzw. maximiert) eine

unrestringierte nichtlineare Funktion von mehreren Variablen mit einem speziellen CG-Verfahren vom Quasi-Newton-Typ. Die Routine e04dgc stützt sich dabei nur auf die Berechnung der Funktions- und Gradientenwerte und benötigt keine Informationen über die zweiten Ableitungen der zu minimierenden Funktion. **Das ist wesentlich, weil die duale Funktion $d_{i,s}(\mathbf{p})$ keine zweite Ableitung besitzt.**

Die in der Routine e04dgc verwendeten Abbruchkriterien sind in [75] ausführlich beschrieben. Unsere Aufgabe besteht darin, den verbleibenden Rechenfehler durch den Parameter η_i des MSR-Verfahrens (s. S. 107) auszudrücken. Dafür müssen wir gemäß dem MSR-Algorithmus die als letzte erhaltene Lösung $(\mathbf{y}_*(\mathbf{p}), \mathbf{u}_*(\mathbf{p}))$ auf die Restriktionsmenge von **QP** exakt projizieren. Das heißt, die endgültigen Näherungen $y^{i,s}$ und $u^{i,s}$ werden durch die Vektoren $\mathbf{y}_f = A_y^{-1}\,(A_u\,\mathbf{u}_*(\mathbf{p}) + \mathbf{f})$ bzw. $\mathbf{u}_f = \mathbf{u}_*(\mathbf{p})$ bestimmt. Weil $\hat{\Psi}^* = d^* \geq d(\mathbf{p})$ gilt, so kann man als Abbruchparameter die Differenz

$$\eta_i \;=\; \hat{\Psi}_{i,s}(\mathbf{y}_f(\mathbf{p}),\, \mathbf{u}_f(\mathbf{p})) - d_{i,s}(\mathbf{p})$$

annehmen. Somit läßt sich das von uns gewählte Abbruchkriterium des MSR-Verfahrens auch in der Praxis nachvollziehen.

Eine andere wichtige Frage ist die Wahl des Startpunkts für das CG-Verfahren, das zur Maximierung der dualen Funktion $d_{i,s}(\mathbf{p})$ benutzt wird. Dabei verfolgen wir wie immer eine Fortsetzungsstrategie und wählen die letzte Näherung der dualen Variablen als Startpunkt für die Maximierung der nächsten dualen Funktion $d_{i,s+1}(\mathbf{p})$. Wenn das Gitter gewechselt wird, so lassen wir den Vektor $\mathbf{p}$ nach derselben Regel erweitern, wie den Vektor $\mathbf{y}$. Die Erweiterung der Vektoren $\mathbf{y}$ und $\mathbf{u}$ muß die Bedingungen $y^{i+1,0} = y^{i,s(i)}$ und $u^{i+1,0} = u^{i,s(i)}$ exakt unterstützen. Die benutzten Basisfunktionen lassen diese exakte Erweiterung zu.

Das MSR-Verfahren ohne die Straftechnik ergibt eine bessere Konvergenz auf feineren Gittern. Deshalb bezieht sich die numerische Untersuchung eines solchen Verfahrens nur auf die letzten Diskretisierungsniveaus. Weil der Startpunkt für den CG-Solver auf inneren Proximal-Schritten dem Vektor $\mathbf{p}^*$ immer näher kommt, ist es vernünftig zu erwarten, daß die gesamte Singularität der dualen Funktion in der Nähe ihres Maximums mit der quadratischen Funktion aus der rechten Seite von (4.67) beschrieben werden kann. Die Hesse-Matrix dieser quadratischen Funktion ist umso schlechter konditioniert, je größer ihre Dimension ist. Jedoch läßt sich die Matrix $A_y\,H_y^{-1}A_y^{\mathrm{T}}$ gut vorkonditionieren. Die Idee beruht darauf, daß die Massenmatrix H_y eine wesentlich bessere Kondition im Vergleich zu der Steifigkeitsmatrix A_y besitzt, falls sich der Diskretisierungsparameter verkleinert. Man könnte die inverse Steifigkeitsmatrix A_y^{-1}, die im Gegensatz zu A_y dicht besetzt ist und beim eventuellen Abspeichern extrem viel Platz einnehmen würde, durch die iLU-Technik approximieren. Im Falle des Fehlens zweiter Ableitungen der zu maximierenden differenzierbaren Funktion bedeutet diese Technik eine formale Ersetzung der dualen Variablen:

$$\mathbf{p} \;=\; \tilde{U}^{-1}\tilde{L}^{-1}\,\mathbf{b}\,,$$

wobei $\tilde{L}$ und $\tilde{U}$ durch die iLU(0)-Zerlegung von A_y [71] bestimmt werden und $\mathbf{b}$ die neuen dualen Variablen darstellen. In neuen Variablen entstehen andere Berechnungsformeln für (4.65) und (4.66). Offensichtlich hängen der neue und der alte Gradient der dualen Funktion folgendermaßen zusammen:

$$\vec{\nabla}_{\mathbf{b}}\, d_{i,s}(\mathbf{b}) \;=\; \tilde{U}^{-1}\tilde{L}^{-1}\,\vec{\nabla}_{\mathbf{p}}\, d_{i,s}(\mathbf{p})\;.$$

Das heißt, man berechnet einmal für das gegebene Diskretisierungsniveau die dünn besetzten Matrizen $\tilde{L}$ und $\tilde{U}$, speichert sie ab und invertiert sie zweimal auf jeder CG-Iteration. Der gesamte Rechenaufwand verringert sich aber, weil die Gesamtzahl der vorkonditionierten CG-Iterationen kleiner wird. Dies kann man aus folgender Tabelle entnehmen:

h_i	χ_i	η_i	$s(i)$	CG(i)	$J(z^{i,s(i)})\ 10^4$	Prox-Schritt
2^{-5}	$1.02\ 10^{-8}$	$2.66\ 10^{-9}$	2	210	1.33329	1.16
2^{-6}	$6.80\ 10^{-9}$	$8.87\ 10^{-10}$	10	615	1.33261	$4.18\ 10^{-1}$
2^{-7}	$4.53\ 10^{-9}$	$8.58\ 10^{-10}$	40	1939	1.33245	$5.17\ 10^{-1}$
2^{-8}	$3.02\ 10^{-9}$	$2.02\ 10^{-9}$	100	12108	1.33241	$7.63\ 10^{-1}$
2^{-7}	$4.53\ 10^{-9}$	$1.50\ 10^{-10}$	40	1294	1.33245	$1.16\ 10^{-1}$
2^{-8}	$3.02\ 10^{-9}$	$8.25\ 10^{-11}$	100	3958	1.33241	$5.11\ 10^{-2}$

In der Spalte "Prox-Schritt" sind die Werte $\|u^{i,s(i)} - u^{i,s(i)-1}\|_\Omega$ angefürt. Die zwei letzten Zeilen veranschaulichen gut die Vorteile, welche die Strategie der Vorkonditionierung der dualen Funktion erbringen kann. Diese Vorteile sind umso stärker, je feiner das Diskretisierungsniveau ist.

Im Anhang B findet man mehr ausführliche Informationen vom Verhalten des MSR-Verfahrens auf dem 8-ten Diskretisierungsniveau. Die erste Tabelle entspricht dem vorkonditionierten CG-Verfahren und die zweite bezieht sich auf das Verfahren ohne Vorkonditionierung. In der Spalte "Residual" sind die Werte des Betrags $\| A_y\, \mathbf{y}_*(\mathbf{p}) - A_u\, \mathbf{u}_*(\mathbf{p}) - \mathbf{f} \|$ angefürt. Die Anzahl der CG-Iterationen auf späteren Proximal-Schritten wurde künstlich mit 400 beschränkt. Ohne die Vorkonditionierung entstanden ständige Ausbrüche des CG-Verfahrens, da die obere Grenze erreicht wurde. Bemerkenswert ist, daß solche Effekte im Falle eines vorkonditionierten CG völlig ausbleiben. Obwohl der Zielfunktionalswert $J(z^{8,s})$ auf den inneren Proximal-Schritten die ersten sieben Stellen nicht ändert, unterscheiden sich die Näherungen $u^{8,1}$ und $u^{8,100}$ voneinander erheblich. Die letzte Näherung stellt dabei eine wesentlich bessere Reproduktion der Schaltstruktur dar. Solche numerischen Effekte lassen sich leider durch unsere Theorie nicht erklären; den Abbruch der inneren Prox-Schleife muß man immerhin empirisch bestimmen, weil die numerischen Verfahren in der Praxis ein besseres Verhalten aufweisen, als man aus den theoretischen Abschätzungen schließen kann.

Die Basisidee der von uns verfolgten Theorie ist ein Zusammenspiel von Diskretisierung und Regularisierung. Dieses Zusammenspiel kann man auf der Abbildung 4.8 deutlich sehen. Dort wird dargestellt, wie sich die vierte instabile Bang-Bang-Formation graduell unterschiedlich auf einer Folge von Gittern reproduzieren

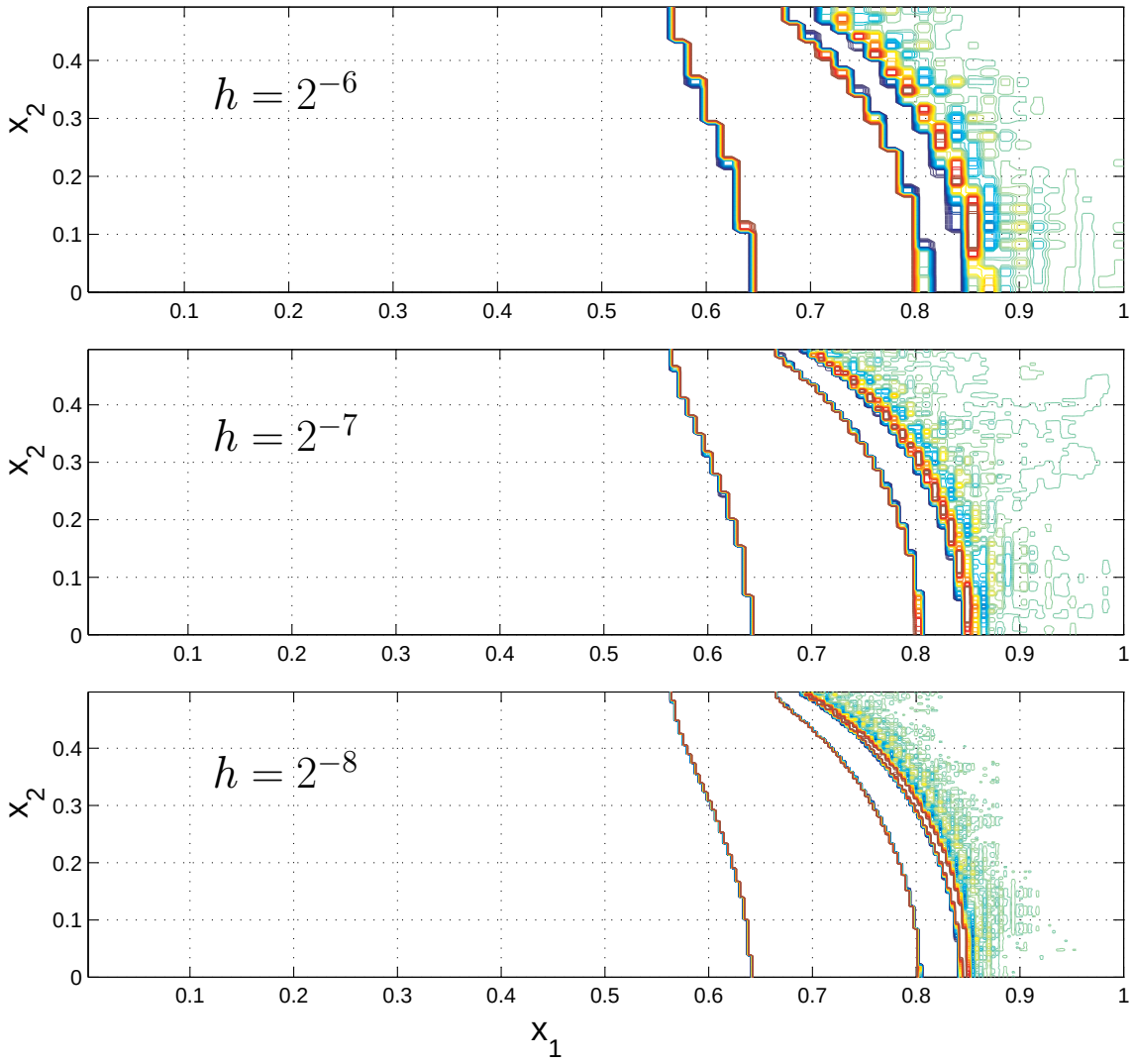


Abbildung 4.8: Konvergenz des MSR-Verfahrens auf Folgen von Gittern, Approximation der Steuerung $u^{i,s(i)}(x_1, x_2)$

läßt. Die Einstellung des Regularisierungsparameters χ_i darf man nicht schneller als erlaubt verkleinern. Die Geschwindigkeit der Verkleinerung muß der Regel (4.53) genügen.

Die Abbildung 4.9 zeigt, welche Vorteile das MSR-Verfahren im Vergleich mit dem herkömmlichen Tichonov-Verfahren hat. Die untere Graphik entspricht der optimalen Zustandsfunktion. Die feinsten Bang-Bang-Formationen der Steuerung lassen sich aufgrund mehrerer Regularisierungsschritte wesentlich besser reproduzieren. Bemerkenswert ist, daß man den Regularisierungsparameter χ_i in beiden Verfahren auf gleiche Weise einstellt. Der sichtbare Nachteil des MSR-Verfahrens besteht in seinem weniger stabilen Verhalten im Bereich der singulären Steuerung. Dort wird der Diskretisierungsfehler schlechter gedämpft, was der Auffassung der schwachen Konvergenz gut entspricht. Um diesen ungewünschten Effekt zu vermeiden, könnte man eventuell ein gemischtes Regularisierungsverfahren anwenden, dessen Untersuchung außerhalb dieser Arbeit liegt.

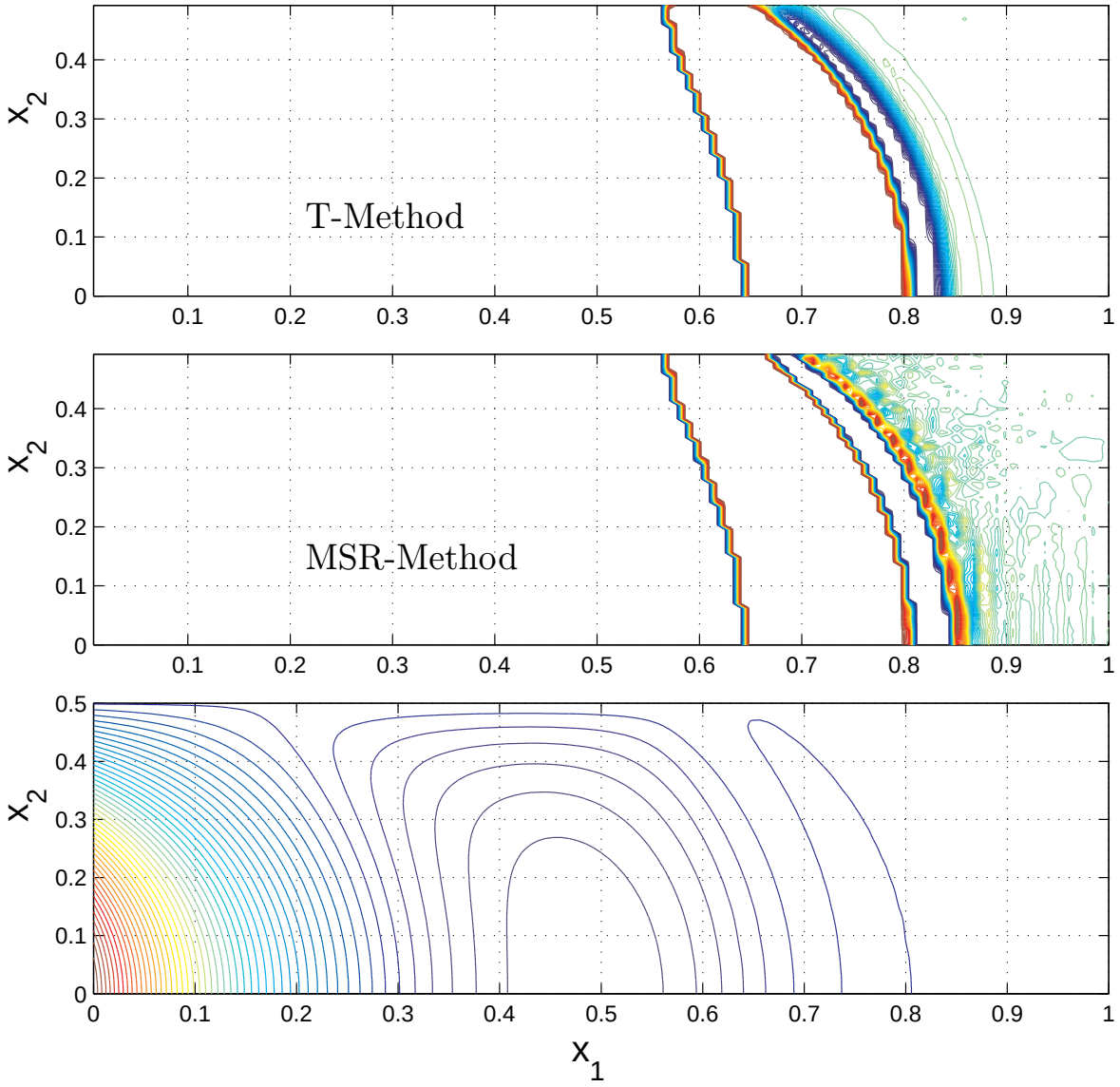


Abbildung 4.9: Der numerische Vergleich von Tichonov- und Proximal-Punkt-Verfahren, Steuerung $u^{7,s(7)}(x_1, x_2)$ und Zustand $y^{7,s(7)}(x_1, x_2)$

Anhang A

Räume

$\mathcal{U}$ — Steuerraum, Hilbertraum mit $\|\cdot\|_{\mathcal{U}}$, $\Lambda_{\mathcal{U}}$ [1]
$\mathcal{U}_0$ — Hilfsraum, Hilbertraum mit $\|G_u \cdot\|_{\mathcal{U}}$, $G_u^* \Lambda_{\mathcal{U}} G_u$
$\mathcal{Y}$ — Zustandsraum, Hilbertraum mit $\|\cdot\|_{\mathcal{Y}}$, $\Lambda_{\mathcal{Y}}$
$\mathcal{Y}_0$ — Hilfsraum, Hilbertraum mit $\|G_y \cdot\|_{\mathcal{Y}}$, $G_y^* \Lambda_{\mathcal{Y}} G_y$
$\mathcal{H}$ — Kopplungsraum, Hilbertraum mit $\|\cdot\|_{\mathcal{H}}$, $\Lambda_{\mathcal{H}}$
$\mathcal{K}$ — Beobachtungsraum, Hilbertraum mit $\|\cdot\|_{\mathcal{K}}$, $\Lambda_{\mathcal{K}}$
$\mathcal{T}$ — Raum der Regularisierung, Hilbertraum mit $\|\cdot\|_{\mathcal{T}}$, $\Lambda_{\mathcal{T}}$
$\mathcal{S}$ — Raum für die Strafauswertung, Hilbertraum mit $\|\cdot\|_{\mathcal{S}}$, $\Lambda_{\mathcal{S}}$
$\mathcal{Z} = \mathcal{Y} \times \mathcal{U}$ — Prozessraum, Hilbertraum mit

$$\|\cdot\|_{\mathcal{Z}} = \sqrt{\|\cdot\|_{\mathcal{Y}}^2 + \|\cdot\|_{\mathcal{U}}^2} \tag{A.1}$$

Operatoren

$A \in \mathcal{L}(\mathcal{Y}, \mathcal{H})$ — Systemoperator
$B \in \mathcal{L}(\mathcal{U}, \mathcal{H})$ — Operator der Kontrolle
$C \in \mathcal{L}(\mathcal{Y}, \mathcal{K})$ — Operator des Zielfunktionals
$W \in \mathcal{L}(\mathcal{H}, \mathcal{S})$ — Operator des Strafterms
$N \in \mathcal{L}(\mathcal{U}, \mathcal{T})$ — Regularisierungsoperator
$G_u \in \mathcal{L}(\mathcal{U}_0, \mathcal{U})$, $G_y \in \mathcal{L}(\mathcal{Y}_0, \mathcal{Y})$ — Einbettungsoperator

[1] Mit verschiedenen Λ.'s bezeichnen wir Operatoren des kanonischen Isomorphismus [57] der entsprechenden Hilberträume. Darüber hinaus bedeutet es, daß alle auftretenden adjungierten Operatoren nach dem verallgemeinerten Schema [99, § 15.2] zusammengestellt werden.

Primäre Konstanten[2]

$\forall\, u \in \mathcal{U}\,,\ \forall\, y \in \mathcal{Y}\,,\ \forall\, u_0 \in \mathcal{U}_0\,,\ \forall\, y_0 \in \mathcal{Y}_0\,,\ \forall\, h \in \mathcal{H}\ :$

$$m_A\ \|G_y\, y_0\|_{\mathcal{Y}} \le \|AG_y\, y_0\|_{\mathcal{H}} \tag{2.4}$$

$$\|A\, y\|_{\mathcal{H}} \le M_A\ \|y\|_{\mathcal{Y}} \tag{2.2}$$

$$m_B\ \|G_u\, u_0\|_{\mathcal{U}} \le \|BG_u\, u_0\|_{\mathcal{H}} \tag{2.5}$$

$$\|B\, u\|_{\mathcal{H}} \le M_B\ \|u\|_{\mathcal{U}} \tag{2.3}$$

$$m_C\ \|y\|_{\mathcal{Y}} \le \sqrt{\|C\, y\|_{\mathcal{K}}^2 + \|A\, y\|_{\mathcal{H}}^2} \tag{2.6}$$

$$\|C\, y\|_{\mathcal{K}} \le M_C\ \|y\|_{\mathcal{Y}} \tag{2.7}$$

$$m_W\ \|h\|_{\mathcal{H}} \le \|W\, h\|_{\mathcal{S}} \le M_W\ \|h\|_{\mathcal{H}} \tag{2.11}$$

$$m_N\ \|u\|_{\mathcal{U}} \le \|N\, u\|_{\mathcal{T}} \le M_N\ \|u\|_{\mathcal{U}} \tag{2.12}$$

$$\exists\, M_{ad} > 0\ :\quad \forall\, u \in \mathcal{U}_{ad}\quad \|u\|_{\mathcal{U}} \le M_{ad} \tag{2.8}$$

$$\exists\, m_{ad} > 0\ :\quad \exists\, \tilde{u} \in \mathcal{U}_{ad}\,,\quad \forall\, y \in \partial\mathcal{Y}_{ad}$$

$$m_{ad} \le \|G_y(AG_y)^{-1}(B\,\tilde{u} + f) - y\|_{\mathcal{Y}} \tag{2.9}$$

[2]Die Nummern in Klammern verweisen auf die Definitionen der Konstanten im Text

Abgeleitete Konstanten

$$\hat{c}_q = c_p \left(\frac{1}{m_W^2} + c_0 \sqrt{\bar{r}} \right) \tag{2.70}$$

$$\hat{\hat{c}}_q = \frac{1}{\tau} \left(c_y^2 M_C^2 \left(1 + \frac{M_A}{m_A} \right) + c_p \left(2 M_B M_{ad} + \frac{\tau}{m_W^2} + c_0 \sqrt{\bar{r}} \right) + 2 \bar{\chi} M_N M_{ad} \left(M_N M_{ad} + \|t_d\|_{\mathcal{T}} \right) \right) \tag{2.80}$$

$$\tau = m_W^2 m_A m_{ad} \tag{2.81}$$

$$\bar{\chi} = \sup_i \chi_i \,, \qquad \bar{r} = \sup_i r_i \tag{2.58}$$

$$c_0 = \sqrt{\frac{M_C^2}{m_A^2} \left(M_B M_{ad} + \|f\|_{\mathcal{H}} \right)^2 + \bar{\chi} \left(M_N M_{ad} + \|t_d\|_{\mathcal{T}} \right)^2}$$

$$c_p = \frac{M_C^2}{m_A} c_y \tag{2.68}$$

$$c_y = \frac{1}{m_C} \sqrt{\left(c_r \frac{M_C^2}{m_A^2} + \frac{2 M_W^2}{m_W^2} \right) (M_B M_{ad} + \|f\|_{\mathcal{H}})^2 + c_r \bar{\chi} \left(M_N M_{ad} + \|t_d\|_{\mathcal{K}} \right)^2} \tag{2.60}$$

$$c_r = \max \left\{ 1 \,, \frac{2 \bar{r}}{m_W^2} \right\}$$

$$c_w = \frac{1}{m_{ad}} \left(c_y + 2 M_{ad} + \frac{1}{m_A} \left(c_y M_A + 2 M_B M_{ad} \right) \right) \tag{2.100}$$

$$c_\kappa(\chi_i, r_i) = \frac{1}{2} \min \left\{ m_C^2 c_\lambda \,, m_N^2 c_u \lambda_i \right\} \tag{2.51}$$

$$c_\lambda = \min \left\{ 1 \,, m_W^2 \lambda_i \right\} \tag{2.48}$$

$$c_u = \min \left\{ 1 \,, m_W^2 \left[\left(\frac{m_B}{m_N} \right)^2 \Big/ \frac{M_B}{M_N} \right]^2 \right\} \tag{2.48}$$

$$\lambda_i = \min \left\{ \chi_i \,, \frac{\chi}{r_i \chi + 2} \right\} \,, \qquad \chi = \frac{m_N^2}{M_B^2 M_W^2} \chi_i \tag{2.44}$$

Anhang B

s	CG(s)	Residual	Proximal-Schritt	$\hat{\Psi}_{8,s}(\mathbf{z}_f) - d_{8,s}(\mathbf{p})$
1	1784	2.189102e-02	7.121102e-02	4.532518e-11
2	211	2.189230e-02	1.817831e-02	1.492000e-11
3	38	2.189185e-02	1.050944e-02	3.549510e-11
4	18	2.189137e-02	6.908398e-03	1.983392e-11
5	2	2.189169e-02	5.756898e-03	1.699375e-11
6	270	2.189156e-02	5.484137e-03	4.017071e-11
7	2	2.189174e-02	4.334813e-03	8.244119e-11
8	12	2.189154e-02	3.364530e-03	6.932330e-11
9	1	2.189131e-02	3.096868e-03	6.510591e-11
10	72	2.189126e-02	4.668582e-03	6.183374e-11
11	46	2.189135e-02	2.477037e-03	3.642434e-11
12	1	2.189128e-02	2.318179e-03	2.484237e-11
13	1	2.189122e-02	2.118570e-03	2.120562e-11
14	40	2.189130e-02	2.913102e-03	3.610770e-11
15	37	2.189120e-02	1.735328e-03	2.818401e-11
16	1	2.189117e-02	1.674001e-03	2.401062e-11
17	1	2.189116e-02	1.619942e-03	2.178703e-11
18	1	2.189118e-02	1.566853e-03	2.076073e-11
19	1	2.189118e-02	1.502496e-03	2.038958e-11
20	1	2.189121e-02	1.427395e-03	2.047349e-11
21	1	2.189117e-02	1.367821e-03	2.079907e-11
22	1	2.189125e-02	1.322562e-03	2.118264e-11
23	39	2.189132e-02	3.884925e-03	3.932881e-11
24	38	2.189134e-02	1.272221e-03	3.480192e-11
25	1	2.189122e-02	1.218102e-03	3.259243e-11

s	CG(s)	Residual	Proximal-Schritt	$\hat{\Psi}_{8,s}(\mathbf{z}_f) - d_{8,s}(\mathbf{p})$
26	1	2.189116e-02	1.179555e-03	3.130792e-11
27	1	2.189124e-02	1.127658e-03	3.063592e-11
28	1	2.189109e-02	1.088168e-03	3.052034e-11
29	1	2.189136e-02	1.047096e-03	3.101629e-11
30	1	2.189099e-02	1.016220e-03	3.193650e-11
31	1	2.189144e-02	9.744332e-04	3.286601e-11
32	1	2.189096e-02	9.516480e-04	3.402353e-11
33	400	2.189092e-02	9.951590e-03	7.843580e-12
34	116	2.189178e-02	2.316425e-03	7.795361e-12
35	1	2.189128e-02	2.073548e-03	1.082199e-11
36	1	2.189124e-02	2.012194e-03	1.610807e-11
37	30	2.189150e-02	2.271217e-03	7.880386e-12
38	1	2.189141e-02	1.846694e-03	1.233439e-11
39	240	2.189345e-02	5.264113e-03	1.491340e-11
40	1	2.189107e-02	3.432706e-03	4.171365e-11
41	26	2.189233e-02	1.240518e-03	2.595869e-11
42	1	2.189119e-02	1.123317e-03	1.886234e-11
43	1	2.189160e-02	1.016374e-03	1.597628e-11
44	1	2.189111e-02	9.301066e-04	1.520009e-11
45	121	2.189314e-02	2.494669e-03	2.189923e-11
46	22	2.189066e-02	1.774204e-03	1.651878e-11
47	1	2.189136e-02	1.752010e-03	1.289062e-11
48	1	2.189109e-02	1.725529e-03	1.043751e-11
49	1	2.189150e-02	1.697211e-03	8.711322e-12
50	1	2.189110e-02	9.285902e-04	8.569236e-12
51	77	2.189246e-02	3.382096e-03	1.069474e-11
52	45	2.189094e-02	1.726410e-03	1.608990e-11
53	1	2.189092e-02	1.698006e-03	2.338661e-11
54	1	2.189136e-02	1.674409e-03	3.211355e-11
55	73	2.189075e-02	2.527489e-03	1.767947e-11
56	1	2.189168e-02	2.444374e-03	1.075031e-11
57	46	2.189001e-02	1.152693e-03	1.268405e-11
58	1	2.189125e-02	1.040852e-03	1.545389e-11
59	27	2.189151e-02	1.003945e-03	1.289975e-11
60	1	2.189169e-02	8.963363e-04	1.197227e-11
61	1	2.189121e-02	8.247758e-04	1.159959e-11
62	32	2.189188e-02	1.364476e-03	1.187895e-11
63	22	2.189143e-02	4.807020e-04	1.169286e-11
64	1	2.189162e-02	4.713910e-04	1.160840e-11
65	1	2.189131e-02	4.522297e-04	1.168062e-11

s	CG(s)	Residual	Proximal-Schritt	$\hat{\Psi}_{8,s}(\mathbf{z}_f) - d_{8,s}(\mathbf{p})$
66	1	2.189143e-02	4.371434e-04	1.189725e-11
67	1	2.189123e-02	4.182133e-04	1.226266e-11
68	1	2.189140e-02	4.098418e-04	1.269956e-11
69	1	2.189122e-02	3.978238e-04	1.314215e-11
70	1	2.189139e-02	3.920973e-04	1.356602e-11
71	1	2.189122e-02	3.769533e-04	1.385556e-11
72	1	2.189139e-02	3.717663e-04	1.407330e-11
73	1	2.189122e-02	3.622581e-04	1.418090e-11
74	1	2.189139e-02	3.441980e-04	1.425313e-11
75	1	2.189122e-02	3.318676e-04	1.423387e-11
76	1	2.189139e-02	3.273669e-04	1.417482e-11
77	1	2.189122e-02	3.189112e-04	1.405693e-11
78	1	2.189139e-02	3.157909e-04	1.393059e-11
79	1	2.189122e-02	3.100439e-04	1.379377e-11
80	1	2.189139e-02	3.044110e-04	1.369518e-11
81	1	2.189122e-02	2.976424e-04	1.362125e-11
82	1	2.189139e-02	2.926054e-04	1.357797e-11
83	1	2.189122e-02	2.849480e-04	1.355801e-11
84	1	2.189139e-02	2.816807e-04	1.354535e-11
85	1	2.189123e-02	2.705404e-04	1.358362e-11
86	1	2.189139e-02	2.658943e-04	1.365710e-11
87	1	2.189123e-02	2.602056e-04	1.371763e-11
88	1	2.189139e-02	2.577879e-04	1.376895e-11
89	1	2.189123e-02	2.460449e-04	1.381442e-11
90	1	2.189139e-02	2.433357e-04	1.384486e-11
91	1	2.189123e-02	2.366703e-04	1.384450e-11
92	1	2.189140e-02	2.334295e-04	1.383170e-11
93	1	2.189123e-02	2.257496e-04	1.382345e-11
94	1	2.189140e-02	2.197716e-04	1.381737e-11
95	1	2.189124e-02	2.100891e-04	1.377744e-11
96	1	2.189139e-02	2.081727e-04	1.373958e-11
97	1	2.189124e-02	2.033023e-04	1.371777e-11
98	1	2.189139e-02	2.022354e-04	1.369597e-11
99	1	2.189124e-02	2.001460e-04	1.367698e-11
100	1	2.189139e-02	1.994122e-04	1.365877e-11

s	CG(s)	Residual	Proximal-Schritt	$\hat{\Psi}_{8,s}(\mathbf{z}_f) - d_{8,s}(\mathbf{p})$
1	1773	2.189156e-02	7.023316e-02	2.020093e-10
2	76	2.189145e-02	1.723643e-02	1.129934e-09
3	1	2.189115e-02	1.144790e-02	2.017076e-09
4	139	2.189109e-02	1.319546e-02	2.720388e-10
5	102	2.189115e-02	6.500195e-03	2.746869e-10
6	1	2.189111e-02	6.056016e-03	3.340836e-10
7	45	2.189075e-02	7.250063e-03	3.111178e-10
8	1	2.189100e-02	6.985254e-03	4.381964e-10
9	59	2.189126e-02	7.878855e-03	1.971597e-10
10	47	2.189107e-02	3.263913e-03	2.005779e-10
11	1	2.189105e-02	3.238229e-03	2.092218e-10
12	1	2.189104e-02	3.208887e-03	2.229004e-10
13	51	2.189107e-02	6.431672e-03	2.966010e-10
14	1	2.189101e-02	6.397844e-03	4.556775e-10
15	65	2.189119e-02	6.999055e-03	2.075847e-10
16	45	2.189112e-02	2.143185e-03	2.048251e-10
17	1	2.189109e-02	2.114729e-03	2.035139e-10
18	1	2.189103e-02	2.084752e-03	2.033966e-10
19	1	2.189104e-02	2.042691e-03	2.026135e-10
20	50	2.189103e-02	5.375445e-03	2.107805e-10
21	1	2.189096e-02	5.353523e-03	2.507796e-10
22	52	2.189105e-02	5.947188e-03	2.135891e-10
23	46	2.189125e-02	2.154870e-03	2.354196e-10
24	1	2.189101e-02	2.131457e-03	2.586563e-10
25	1	2.189104e-02	2.107088e-03	2.830028e-10
26	1	2.189099e-02	2.088725e-03	3.090562e-10
27	400	2.189122e-02	8.866598e-03	1.830700e-10
28	400	2.189124e-02	5.835883e-03	2.168248e-10
29	400	2.189163e-02	5.908656e-03	1.374620e-10
30	400	2.189102e-02	5.067348e-03	1.512855e-10
31	400	2.189103e-02	5.154980e-03	1.238387e-10
32	400	2.189106e-02	5.536551e-03	1.599828e-10
33	39	2.189080e-02	4.435699e-03	1.873998e-10
34	1	2.189084e-02	3.593435e-03	2.369565e-10
35	400	2.189102e-02	7.070582e-03	1.997493e-10
36	400	2.189103e-02	3.756871e-03	1.723938e-10
37	400	2.189165e-02	5.601500e-03	1.507201e-10
38	400	2.189125e-02	4.348254e-03	1.287847e-10
39	40	2.189114e-02	3.576382e-03	1.510735e-10
40	1	2.189113e-02	3.403952e-03	1.834820e-10

s	CG(s)	Residual	Proximal-Schritt	$\hat{\Psi}_{8,s}(\mathbf{z}_f) - d_{8,s}(\mathbf{p})$
41	400	2.189102e-02	7.453520e-03	1.925168e-10
42	400	2.189105e-02	4.156141e-03	1.570018e-10
43	43	2.189092e-02	3.169382e-03	1.626288e-10
44	1	2.189084e-02	3.008674e-03	1.697098e-10
45	400	2.189121e-02	8.441457e-03	1.349426e-10
46	280	2.189110e-02	3.418880e-03	1.610153e-10
47	1	2.189096e-02	2.710603e-03	1.968624e-10
48	1	2.189100e-02	2.436804e-03	2.420631e-10
49	148	2.189112e-02	5.652238e-03	1.040663e-10
50	82	2.189115e-02	3.001832e-03	1.558759e-10
51	1	2.189098e-02	2.922013e-03	2.471523e-10
52	109	2.189105e-02	4.326328e-03	1.391819e-10
53	108	2.189107e-02	1.710455e-03	1.560911e-10
54	1	2.189100e-02	1.675973e-03	1.801618e-10
55	1	2.189098e-02	1.658988e-03	2.107689e-10
56	74	2.189104e-02	3.427873e-03	1.393543e-10
57	1	2.189096e-02	3.414055e-03	1.272737e-10
58	82	2.189101e-02	3.946441e-03	1.710008e-10
59	1	2.189094e-02	3.930426e-03	3.565673e-10
60	86	2.189105e-02	4.409879e-03	1.785610e-10
61	58	2.189103e-02	2.477938e-03	1.381060e-10
62	1	2.189096e-02	2.455386e-03	1.169953e-10
63	145	2.189106e-02	4.785703e-03	1.852066e-10
64	101	2.189105e-02	2.447822e-03	1.497358e-10
65	1	2.189108e-02	2.436725e-03	1.326237e-10
66	400	2.189287e-02	2.179933e-02	2.777292e-10
67	400	2.189143e-02	1.308555e-02	1.287684e-10
68	400	2.189132e-02	9.532598e-03	7.967138e-11
69	400	2.189114e-02	5.304293e-03	8.249249e-11
70	38	2.189124e-02	3.170939e-03	9.208354e-11
71	1	2.189109e-02	2.982296e-03	1.014123e-10
72	400	2.189142e-02	5.120125e-03	7.805593e-11
73	231	2.189148e-02	2.203277e-03	8.086266e-11
74	1	2.189136e-02	1.884110e-03	8.070162e-11
75	1	2.189135e-02	1.781582e-03	7.912835e-11
76	40	2.189123e-02	3.022877e-03	7.777424e-11
77	1	2.189131e-02	3.004280e-03	7.993844e-11
78	50	2.189136e-02	2.881426e-03	8.814519e-11
79	1	2.189129e-02	2.870766e-03	1.016365e-10
80	46	2.189132e-02	2.823512e-03	9.795690e-11

s	CG(s)	Residual	Proximal-Schritt	$\hat{\Psi}_{8,s}(\mathbf{z}_f) - d_{8,s}(\mathbf{p})$
81	1	2.189127e-02	2.812666e-03	9.862749e-11
82	59	2.189135e-02	2.958572e-03	8.395205e-11
83	1	2.189122e-02	2.947289e-03	7.661375e-11
84	47	2.189130e-02	2.984648e-03	7.681031e-11
85	1	2.189122e-02	2.974220e-03	8.365466e-11
86	48	2.189123e-02	2.891389e-03	8.306263e-11
87	1	2.189129e-02	2.880946e-03	8.702295e-11
88	50	2.189133e-02	2.853411e-03	7.536083e-11
89	1	2.189116e-02	2.843454e-03	6.904589e-11
90	46	2.189112e-02	2.787530e-03	6.767817e-11
91	1	2.189127e-02	2.777488e-03	7.060482e-11
92	51	2.189129e-02	2.656952e-03	7.274126e-11
93	1	2.189125e-02	2.648143e-03	8.029665e-11
94	60	2.189121e-02	2.591292e-03	6.851843e-11
95	1	2.189119e-02	2.584042e-03	6.410273e-11
96	53	2.189120e-02	2.647412e-03	7.248659e-11
97	1	2.189120e-02	2.638843e-03	9.008987e-11
98	54	2.189124e-02	2.896101e-03	7.151879e-11
99	1	2.189126e-02	2.884318e-03	6.449576e-11
100	50	2.189127e-02	2.979959e-03	6.443689e-11

Literaturverzeichnis

[1] R. A. Adams. *Sobolev Spaces.* Academic Press, New York, 1978.

[2] W. Alt and U. Mackenroth. Convergence of finite element approximations to state constrained convex parabolic boundary control problems. *SIAM J. Contr. Opt.,* 27:718-736, 1989.

[3] W. Alt. Local stability of solutions to differentiable optimization problems in Banach spaces. *J. Optim. Theory Appl.,* 70:443-467, 1991.

[4] V. I. Agoshkov and V. I. Lebedev. Poincare-Steklov operators and methods for separation of a domain in variational problems. (Russian) *Vyčislitelnye Prozessy Sist.,* 2:173-227, 1985.

[5] J.-P. Aubin. *Approximation of Elliptic Boundary-value Problems.* Robert E. Krieger Publishing Co., Huntington, New York, 1980.

[6] H. T. Banks and K. Kunisch. *Estimation techniques for distributed parameter systems.* Birkhäuser, Boston – Basel – Berlin, 1989.

[7] A. V. Balakrishnan. A new computing technique in system identification. *J. Comput. and System Sci.,* 2:102-116, 1968.

[8] A. V. Balakrishnan. On a new computing technique in optimal control. *SIAM J. Control,* 6:149-173, 1968.

[9] A. V. Balakrishnan. *Introduction to Optimization Theory in a Hilbert Space.* Springer-Verlag, 1971.

[10] A. V. Balakrishnan. *Applied Functional Analysis.* Springer-Verlag, 1976.

[11] M. Bergounioux. A penalization method for optimal control of elliptic problems with state constraints. *SIAM J. Contr. Opt.,* 30:305-323, 1992

[12] M. Bergounioux. Optimal control of parabolic problems with state constraints: A penalization method for optimality conditions. *Appl. Math. Opt.,* 29:285-307, 1994

[13] M. Bergounioux. Optimal control of problems governed by abstract elliptic variational inequalities with state constraints. *SIAM J. Contr. Opt.,* 36(1):273-289, 1998

[14] M. Bergounioux and K. Kunisch. Aaugmented Lagrangian techniques for elliptic state constrained optimalcontroll problems. *SIAM J. Contr. Opt.,* 35(5):1524-1543, 1997

[15] M. Bergounioux and D. Tiba. General optimality conditions for constrained convex control problems. *SIAM J. Contr. Opt.,* 34:698-711, 1996

[16] M. Bergounioux and F. Tröltzsch. Optimality conditions and generalized bang-bang principle for a state-constrained semilinear parabolic problem. *Num. Funct. Anal. Opt.,* 17:517-536, 1996.

[17] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont – Massachusetts, 1995.

[18] I. N. Bronstein, K. A. Semendjajew, G. Musiol, H. Mühlig. *Taschenbuch der Mathematik.* Verlag Harri Deutsch, Frankfurt am Main, 1995.

[19] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* Noth-Holland Publ. Company, Amsterdam, 1979.

[20] P. G. Ciarlet. Basic error estimates for elliptic problems. In P. G. Ciarlet and J. L. Lions editors, *Handbook of Numerical Analysis,* Vol. 2: Finite Element Methods (Part 1). Elsevier/North-Holland, Amsterdam, New York, Oxford, Tokyo, 1991.

[21] G. Eskin. Boundary-value problems for second-oder elliptic equations in domains with corners. *Proccedings of Symposia in Pure Mathematics.,* 43:105-131, 1985.

[22] R. S. Falk. Approximation of a class of optimal control problems with order of convergence estimates. *J. Math. Anal. and Appl.,* 44:28-47, 1973.

[23] R. P. Fedorenko. *Die approximative Lösung der Optimalsteuerprobleme.* (in Russisch) Nauka, Moskau, 1978.

[24] A. Friedman. *Partial Differential Equations.* Robert E. Krieger publishing Co., Huntington – New York, 1976.

[25] K. O. Friedrichs. *Spectral Theory of Operations in Hilbert Space.* ASM 9, Springer-Verlag, New York – Heidelberg – Berlin, 1973.

[26] A. T. Fuller. Relay control systems optimized for various performance criteria, *Automatic and Remote Control, (Proc. First World Congress IFAC, Moscow, 1960),* 1:510-519, 1961.

[27] R. Glowinski, J.-L. Lions and R. Trémolières. *Numerical Analysis of Variational Problems.* North Holland Publishing Company, 1981.

[28] K. Glashoff and E. Sachs. On theoretical and numerical aspects of the bang-bang principle. *Num. Math.,* 29:93-114, 1977.

[29] P. Grisvard. *Elliptic Problems in Nonsmooth Domains.* Pitman Advanced Publishing Program, Boston – London – Melbourne, 1985.

[30] Ch. Grossmann and H.-G. Roos *Numerik partieller Differentialgleichungen.* Teubner Studienbücher Mathematik. Teubner-Verlag, Stuttgart, 1992.

[31] W. Hackbusch. *Theory und Numerik elliptischer Differentialgleichungen.* Teubner Studienbücher Mathematik. Teubner-Verlag, Stuttgart, 1986.

[32] W. Hackbusch. *Integralgleichungen: Theorie und Numerik.* Teubner Studienbücher Mathematik. Teubner-Verlag, Stuttgart, 1989.

[33] W. Hackbusch and Th. Will. A numerical method for a parabolic bang-bang problem. *Optimal Control of Partial Differential Equations* (K. H. Hoffmann, W. Krabs, eds.), ISNM 68, Birkhäuser, Basel. 1984.

[34] M. Heinkenschloss and E. W. Sachs. Numerical solution of a constrained control problem for a phase field model. *Int. Ser. Numer. Math.,* 118:171-187, 1994.

[35] R. Hettich, A. Kaplan and R. Tichatschke. Regularized penalty methods for ill-posed optimal control problems with elliptic equations,
Part I : Distributed control with bounded control sets and state constraints, *Control and Cybernetics,* 1:5-27;
Part II : Distributed and boundary control with unbounded control set and state constraints. *Control and Cybernetics,* 1:29-43, 1997.

[36] L. Hörmander. *Linear Partial Differential Operators.* Springer-Verlag, Berlin, 1976.

[37] A. D. Ioffe, V. M. Tichomirov. *Theorie der Extremalaufgaben.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1979.

[38] A. Kaplan, R. Tichatschke. A note on the paper: Multi-step-prox-regularization methods for solving convex variational problems. *Optimization,* 37:149-152, 1996.

[39] A. Kaplan, R. Tichatschke. Proximal penalty method for non-coercive parabolic optimal control problems. *Control and Cybernetics,* 1, 1998.

[40] A. Kaplan and R. Tichatschke. *Stable Methods for Ill-Posed Variational Problems – Prox-Regularization of Elliptic Variational Inequalities and Semi-Infinite Problems.* Akademie Verlag, Berlin, 1994.

[41] C. T. Kelley and E. W. Sachs. Solution of optimal control problems by a pointweise projected Newton method. *SIAM J. Contr. Opt.,* 33(6):1731-1757, 1995.

[42] C. T. Kelley and E. W. Sachs. Local convergence of the symmetric rank-one iteration. *Comput. Optim. Appl.,* 9:43-63, 1998.

[43] G. Knowles. Finite element approximation of parabolic time optimal control problems. *SIAM J. Contr. Opt.,* 20:414-427, 1982.

[44] A. N. Kolmogorov und S. V. Fomin. *Reelle Funktionen und Funktionalanalysis.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1975.

[45] V. A. Kondrat'ev and O. A. Oleinik. Boundary-value problems for partial differential equations in non-smooth domains. *Russian Math. Surveys.,* 38(2):1-86, 1983.

[46] F.-S. Kupfer and E. W. Sachs. Reduced SQP method for nonlinear heat conduction control problems. *Int. Ser. Numer. Math.,* 111:145-160, 1993.

[47] L. D. Kudrjavtsev. *Kurs der mathematischen Analysis.* (in Russisch) Vysscha Schkola, Moskau, Band 3, 352 S., 1989.

[48] V. D. Kupradze. *Three-dimensional Problems of the Mathematical Theory of Elasticity and Thermoelasticity.* ed. by V. D. Kupradze, Amsterdam, North-Holland Publ. Co., 929 S., 1979.

[49] M. M. Lavrent'ev, V. G. Romanov, S. P. Schischatskii. *Ill-posed Problems of Mathematical Physics and Analysis.* Translations of mathematical monographs, Vol. 64, 290 S., 1986.

[50] O. A. Ladyzhenskaya and N. N. Ural'tseva. *Linear and Quasilinear Elliptic Equations.* New York, Acad. Press, 1968. - XVIII, 495 S.

[51] O. A. Ladyzhenskaya, V. A. Solonnikov, N. N. Ural'tseva. *Linear and Quasilinear Equations of Parabolic Type.* Transl. Amer. Math. Monographs, vol. 23, American Mathematical Society, Providence, Rhode Island, 1968.

[52] O. A. Ladyzhenskaya. *The Boundary Value Problems of Mathematical Physics.* New York, Springer, 322 S., 1985.

[53] I. Lasiecka. Boundary control of parabolic systems, finite element approximation. *Appl. Math. Opt.,* 6:31-62, 1980.

[54] I. Lasiecka. Ritz-Galerkin approximation of the time-optimal boundary control problem for parabolic systems with Dirichlet boundary conditions. *SIAM J. Contr. Opt.,* 22:477-500, 1984.

[55] V. I. Lebedev. The composition method and Poincare-Steklov operators. (Russian) *Vyčislitelnye Prozessy Sist.,* 6:199-211, 1988.

[56] G. M. Lieberman. Mixed boundary value problems for elliptic and parabolic differential equations of second order. *J. Math. Anal. and Appl.,* 113:422-440, 1986.

[57] J.-L. Lions. *Contrôle Optimal des Systèmes Gouvernés par des Équations aux Dérivées Partielles.* Dunod, Gauthier-Villars, Paris, 1968.

[58] J.-L. Lions. *Control of Distributed Singular Systems.* Dunod, Paris., 1985.

[59] J.-L. Lions and E. Magenes. *Problèmes aux Limites non Homogenes et Applications,* v. 1-2. Dunod, Paris, 1968.

[60] L. A. Ljusternik und W. I. Sobolev. *Elemente der Funktionalanalysis,* Verlag Harri Deutsch, Thun-Frankfurt/Main, 1979.

[61] F. V. Lubyshev. On the accuracy of difference approximations and the regularisation of optimal control problems on solutions of elliptic equations. *Comput. Maths. Math. Phys.,* 27(4):490-500, 1987.

[62] F. V. Lubyshev. Approximation and regularisation of optimal control problems for a non-selfadjoint elliptic equation with variable coefficients. *Comput. Maths. Math. Phys.,* 31(1):10-20, 1991.

[63] U. Mackenroth. Some remarks on the numerical solution of bang-bang type optimal control problems. *Num. Funct. Anal. Opt.,* 5:457-484, 1982.

[64] K. Malanowski. Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal control problems. *Appl. Math. Opt.,* 8:69-95, 1981.

[65] K. Malanowski. Stability and sensitivity of solutions to optimal control problems for systems with control appearing linearly. *Appl. Math. Opt.,* 16:73-91, 1987.

[66] G. I. Marchuk and V. I. Agoshkov. *Einführung in die Gitter-Projektionsverfahren.* (Russian) Nauka, Moscow, 1981.

[67] B. Martinet. Règularisation d'inèquations variationelles par approximations successives. *RIRO,* 4:154-159, 1970.

[68] H. Maurer and H. D. Mittelmann. Optimization techniques for solving elliptic control problems with control and state constraints, Part 1: Boundary control. *Technical report "Angewandte Mathematik und Informatik",* 14/98 – N, 1998.

[69] V. Mazja. *Sobolev Spaces.* Springer-Verlag, Berlin – Heidelberg – New York – Tokyo, 1985.

[70] V. Mazja. *Einbettungssätze für Sobolewsche Räume – Teil 2.* Teubner-Texte zur Mathematik, Band 28, Leibzig, 1979.

[71] J. Meijerink and H. van der Vorst. An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. *Math. Comput.* 31(137):148-162, 1977.

[72] A. R. Mitchell and R. Wait. *The Finite Element Method in Partial Differential Equations.* John Wiley & Sons, Chichester – New York – Brisbane – Toronto, 1980.

[73] U. Mosco. Convergence of convex sets and of solutions of variational inequalities. *Advances in Mathematics,* 3:510-585, 1969.

[74] The NAG Numerical Algorithms Group Limited. *The NAG Fortran Library Manual,* 1st edition, Oxford, 1993.

[75] The NAG Numerical Algorithms Group Limited. *The NAG C Library Manual, Mark 3,* 1st edition, Oxford, 1994.

[76] J. Nečas. *Les Méthods Directes en Théorie des Équations Elliptiques.* Masson, Paris, 1967.

[77] S. M. Nikol'skij. *Approximation of Functions of Several Variables and Imbedding Theorems.* Springer-Verlag, Berlin, 1975.

[78] L. A. Oganisjan und L. A. Ruchovez. *Variations- und Differenzverfahren zur Lösung der elliptischen Gleichungen.* (in Russisch) Verlag der Akademie der Wissenschaften der Armenischen Sozialistischen Sowjetrepublik, Jerewan, S. 335, 1979.

[79] Z. Opial. Weak convergence of the successive approximtaions for nonexpansive mappings in Banach spaces. *Bull. Amer. Math. Soc.,* 73:591-597, 1967.

[80] B. T. Polyak. *Introduction to Optimization.* Optimization Software, Inc. Publ. Division, New York, 1987.

[81] A. J. Pryde. Second order elliptic equations with mixed boundary conditions. *J. Math. Anal. and Appl.,* 80:203-244, 1981.

[82] H. Reisman. Second order elliptic boundary value problems in a domain with Edges. *Comm. Partial Diff. Equat.,* 6(9):1023-1042, 1981.

[83] F. Riesz and B. Sz-Nagy. *Leçons D'analyse Fonctionnelle,* Akad. Kiado, Budapest, 1952.

[84] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Contr. Opt.,* 15:877-898, 1976.

[85] R. T. Rockafellar. Augmented Lagrange multiplier funktions and applications of the proximal point algorithm in convex programming. *Math. Oper. Res.,* 1:97-116, 1976.

[86] E. J. P. G. Schmidt and N. Weck. On the boundary behavior of solutions to elliptic and parabolic equations — with applications to boundary control for parabolic equations. *SIAM J. Contr. Opt.,* 16(4):593-598, 1978.

[87] L. N. Slobodeckij. Verallgemeinerte Sobolev-Räume und Anwendungen auf Randwertaufgaben für Differentialgleichungen mit partiellen Ableitungen. (in Russisch) *Uč. zap. Leningr. gos. ped. institt., A. I. Gercena,* 197:54-112, 1958.

[88] G. Strang. Approximation in the finite element method. *Numer. Math.,* 19:81-98, 1972.

[89] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method.* Prentice-Hall, Enlewood Cliffs, N. J., 1973.

[90] M. E. Taylor. *Pseudodifferential Operators.* Princeton University Press, Princeton – New Jersey, 1981.

[91] A. N. Tikhonov. Improper problems of optimal planning and stable methods for their solution. *Soviet Math. Doklady,* 6:1264-1267, 1965.

[92] A. N. Tikhonov. Methods for the regularization of optimal control problems. *Soviet Math. Doklady,* 6:761-763, 1965.

[93] A. N. Tichonov, A. S. Leonov, A. G. Jagola. *Nichtlineare inkorrekt gestellte Aufgaben.* (in Russisch) Nauka, Moskau, 1995.

[94] V. A. Trenogin. *Funktionalanalysis.* (in Russisch) Nauka, Moskau, 1980.

[95] F. Tröltzsch. *Optimality Conditions for Parabolic Control Problems and Applications.* Teubner Texte, Leibzig, 1984.

[96] F. Tröltzsch. Semidiscrete finite element approximation of parabolic boundary control problems – convergence of switching points. *ISNM 78,* Birkhäuser Verlag, Basel, 219-232, 1987.

[97] F. P. Vasil'ev. *Methods for Solving Extremal Problems.* (Russian) Nauka, Moscow, 1981.

[98] A. Voigt and J. Wloka. *Hilberträume und elliptische Differentialoperatoren.* B.I.–Wissenschaftsverlag, Mannheim – Wien – Zürich, 1975.

[99] J. Wloka. *Funktionalanalysis und Anwendungen.* Walter de Gruyter –Berlin – New York, 1971.

[100] J. Wloka. *Partielle Differentialgleichungen.* Teubner-Verlag, Stuttgart, 1982.

[101] J. Wloka, B. Rowley, B. Lawruk. *Boundary Value Problems for Elliptic Systems.* Cambridge University Press, 1995.

[102] K. Yosida. *Functional Analysis.* Springer-Verlag, Berlin – Göttingen – Heidelberg, 1965.

[103] M. I. Zelikin and V. F. Borisov. Regimes with increasingly more frequent switchings in optimal control problems. *Proceedings of the Steklov Institute of Mathematics.,* Issue 1:95-186, 1991.

[104] M. I. Zelikin and V. F. Borisov. *Theory of Chattering Control with Applications to Astronautics, Robotics, Economics, and Engineering.* Birkhäuser Boston, 1994.

[105] W. P. Ziemer. *Weakly Differentiable Functions.* Springer-Verlag, New York – Berlin – Heidelberg, 1989.