Analysis und Approximation der Cahn-Hilliard Gleichung mit Hindernispotential

Carsten Gräser

Diplomarbeit im Fach Mathematik

eingereicht am
Fachbereich Mathematik und Informatik
der
Freien Universität Berlin
August 2004

Inhaltsverzeichnis

Einleitung **5**

1 Physikalische Motivation der Cahn-Hilliard Gleichung **7**
1.1 Die Ginzburg-Landau Energie 7
1.2 Die Cahn-Hilliard Gleichung 8

2 Funktionalanalytischer Hintergrund **11**
2.1 Schwache Konvergenz 11
2.2 Konvexe Analysis 15
2.3 Maximal monotone Operatoren 19
2.4 Gelfand'sche Dreier 21
2.5 Sobolevräume 23
2.6 Funktionenräume für Evolutionsgleichungen 26

3 Die Cahn-Hilliard Gleichung **34**
3.1 Die Cahn-Hilliard Gleichung als Variationsproblem 34
3.2 Existenz und Eindeutigkeit von Lösungen 35
3.2.1 Beweisskizze 35
3.2.2 Galerkin-Approximation von $H^1(\Omega)$ 35
3.2.3 Ein differenzierbares Energiefunktional 37
3.2.4 Regularisierung der Variationsungleichung 39
3.2.5 Lösungen des regularisierten Problems 39
3.2.6 Lösungen der Variationsungleichung 48
3.2.7 Regularität der Lösungen 51

4 Eine Diskretisierung der Cahn-Hilliard Gleichung **53**
4.1 Die Methode der finiten Elemente 53
4.2 Definitionen 57
4.3 Die semi-implizite Diskretisierung 59
4.4 Existenz und Eindeutigkeit von Lösungen 60
4.5 Konvergenz 63

5 Lösung der diskreten Probleme **69**
5.1 Die algebraische Darstellung der diskreten Cahn-Hilliard Gleichung 69
5.2 Der Splitting-Algorithmus von Lions-Mercier 71
5.2.1 Der Lions-Mercier Algorithmus für maximal monotone Operatoren . . 71
5.2.2 Die Cahn-Hilliard Gleichung als Operator-Inklusion 72
5.2.3 Die Auswertung der Operatoren 74
5.2.4 Implementierung des Algorithmus 76
5.3 Das Gauß-Seidel Verfahren 77
5.3.1 Das Gauß-Seidel Verfahren für lineare Probleme 78
5.3.2 Das Gauß-Seidel Verfahren für Variationsungleichungen 79
5.3.3 Ein Gauß-Seidel Verfahren für die diskrete Cahn-Hilliard Gleichung . . 81

5.3.4 Implementierung des Algorithmus 84
5.4 Die Cahn-Hilliard Gleichung als Minimierungsproblem 85
5.4.1 Aufstellen eines Energiefunktionals 86
5.4.2 Das Minimierungsproblem 87
5.4.3 Das Sattelpunktproblem 89
5.5 Der Uzawa Algorithmus 91
5.5.1 Der Uzawa Algorithmus für Minimierungsprobleme 91
5.5.2 Die Teilprobleme im Uzawa Algorithmus 92
5.5.3 Erzwingen von Massenerhaltung 93
5.5.4 Der Uzawa Algorithmus für die diskrete Cahn-Hilliard Gleichung . . . 95
5.5.5 Implementierung des Algorithmus 97

6 Numerische Experimente **99**
6.1 Implementierung 99
6.2 Resultate 99
6.3 Fazit 101
6.4 Ergebnisse der numerischen Experimente 102
6.4.1 Die Lösung in 1-D 102
6.4.2 Die Lösung in 2-D 103
6.4.3 Verlauf der Iteration in 1-D 104
6.4.4 Verlauf der Iteration in 2-D 105
6.4.5 Anzahl der Iterationsschritte 106

A Notation **107**

Literatur **109**

Einleitung

Die Methoden der Mathematik werden vielfach eingesetzt, um komplexe Vorgänge physikalischer, chemischer, biologischer oder volkswirtschaftlicher Natur zu modellieren und zu simulieren. Zur Beschreibung physikalischer Prozesse haben sich insbesondere partielle Differentialgleichungen als hilfreich erwiesen. Sie gestatten es, Erhaltungssätze sowie Zustands- und Materialgesetze elegant zu formulieren und damit Modelle zur Beschreibung dieser Prozesse zu entwickeln.

In der Materialwissenschaft kommt Phasenübergangs- und Phasenseparationsprozessen besondere Bedeutung zu. Diese beschreiben, wie sich verschiedene Phasen ineinander umwandeln, wie es bei Schmelz- oder Reaktionsvorgängen der Fall ist, oder sich entmischen. Die Entmischung von Legierungen als Beispiel für Phasenseparation ist ein Effekt, der die Haltbarkeit von Materialien und Werkstücken deutlich beeinflussen kann. Deshalb ist man an einer Simulation der dabei auftretenden Prozesse interessiert. Zur Modellierung solcher Vorgänge werden die Gesetze der Thermodynamik angewandt. Das Resultat sind Phasenfeldmodelle, die auf partiellen Differentialgleichungen beruhen und verschiedene Aspekte des jeweiligen Vorgangs beschreiben. Ein einfaches Modell für Separationsprozesse ist die Cahn-Hilliard Gleichung.

Im Rahmen dieser Arbeit soll ein Einblick in die Struktur der Cahn-Hilliard Gleichung mit Hindernispotential und ein Überblick über die Techniken zu ihrer numerischen Lösung mit dem Computer gegeben werden. Dazu wird in Kapitel 1 eine kurze physikalische Motivation entwickelt, um ein Verständnis für die Bedeutung der Gleichung zu schaffen.

In Kapitel 2 werden die mathematischen Hilfsmittel, die zur analytischen und numerischen Behandlung der Gleichung benötigt werden, eingeführt. Neben Begriffen aus der linearen Funktionalanalysis sind dies insbesondere solche aus der nichtlinearen Funktionalanalysis und der konvexen Analysis. Sie ermöglichen es, auch nichtlineare und nicht glatte Probleme wie die Cahn-Hilliard Gleichung mit Hindernispotential elegant zu behandeln.

Die kontinuierliche Cahn-Hilliard Gleichung wird in Kapitel 3 analytisch behandelt. In diesem Rahmen werden mit den eingeführten Techniken und Begriffen Existenz- und Eindeutigkeitsresultate gezeigt. Grundlage der Untersuchungen ist dabei eine Formulierung der Gleichung als Variationsproblem in einem Sobolevraum.

Kapitel 4 untersucht eine Diskretisierung der Gleichung mittels finiter Element Techniken. Es wird die Existenz und Eindeutigkeit von diskreten Lösungen gezeigt und ein Konvergenzresultat zitiert. Die darin gegebene Fehlerabschätzung stellt den Zusammenhang zwischen den Lösungen der kontinuierlichen und der diskreten Gleichung her.

Die Lösung der dabei entstehenden diskreten Probleme ist Thema von Kapitel 5. Da die üblichen Verfahren zur Lösung linearer Probleme nicht anwendbar sind, werden nichtlineare Löser betrachtet. Es wird der Operator-Splitting Algorithmus von Lions und Mercier und eine von Barrett vorgeschlagene nichtlineare Block-Gauß-Seidel Iteration präsentiert. Da die Geschwindigkeit dieser Algorithmen unbefriedigend ist, wird eine Formulierung der diskreten Cahn-Hilliard Gleichung als Minimierungs- und Sattelpunktproblem entwickelt. Ziel ist dabei die Anwend-

barkeit schneller Mehrgitterverfahren. Es wird jedoch eine zusätzliche Nebenbendingung eingeführt, die die Anwendung des ebenfalls langsamen Uzawa Algorithmus notwendig macht.

Der Gauß-Seidel und der Uzawa Algorithmus werden schließlich in Kapitel 6 anhand exemplarischer Rechnungen für ein- und zweidimensionale Probleme verglichen. Es wird insbesondere auf die Entwicklung der Geschwindigkeit der Algorithmen bei wachsender Anzahl von Unbekannten eingegangen.

Für die in der Arbeit angeschnittenen Themenkomplexe der Mathematik werden Referenzen zu detaillierteren Darstellungen gegeben. Die präsentierten Sätze werden entweder bewiesen oder es werden, falls die Beweise den Rahmen der Arbeit sprengen würden, Referenzen auf Arbeiten oder Bücher gegeben, in denen sie zu finden sind. In wenigen Fällen wird eine kurze Beweisskizze gegeben, um auf einzelne Aspekte hinzuweisen, ohne den gesamten Beweis auszuführen.

An dieser Stelle möchte ich herzlich Prof. Ralf Kornhuber für seine engagierte, geduldige und allzeit hilfsbereite Betreuung bei der Bearbeitung des Themas danken.

Berlin, August 2004 Carsten Gräser

1 Physikalische Motivation der Cahn-Hilliard Gleichung

Die Cahn-Hilliard Gleichung beruht als Phasenfeldmodell auf den Gesetzen der Thermodynamik. Sie beschreibt dabei die Separation zweier Spezies. Eine Einführung in die Thermodynamik und eine detaillierte thermodynamische Fundierung von Phasenfeldmodellen kann und soll allerdings im Rahmen dieser Arbeit nicht erfolgen. Ziel ist es vielmehr, durch eine kurze Motivation einen Eindruck von den physikalischen Prozessen zu vermitteln, die die Gleichung beschreiben soll. Eine detaillierterer Einblick in die Modellierung von Phasenseparationsprozessen ist beispielsweise bei [BS96] zu finden.

1.1 Die Ginzburg-Landau Energie

Die Cahn-Hilliard Gleichung beschreibt die Separation eines Gemisches oder einer Legierung zweier Spezies A und B in einem beschränkten räumlichen Gebiet Ω. Für solche Legierungen lässt sich beobachten, dass sie oberhalb einer kritischen Temperatur T_c stabil sind, das heißt es findet keine Entmischung statt. In diesem Zustand können die Spezies uniform über das Gebiet verteilt sein. Wird die Temperatur T jedoch unter die kritische Temperatur T_c gesenkt, so setzen Entmischungsprozesse ein. Einen guten Überblick über diese Thematik verschafft der Artikel [Ell89].

Zunächst werden jedoch stationäre Gleichgewichtszustände beschrieben. Hierzu betrachtet man im einfachsten Fall eine lokale freie Energie $\psi(u, T)$ für jeden Punkt x im Gebiet, wobei u den Zustand im Punkt x beschreibt. Es ist jedoch zu beobachten, dass der Separationsprozess in räumlich getrennten Punkten nicht unabhängig voneinander stattfindet. Es bilden sich Regionen aus, in denen nur jeweils eine Phase vorliegt, was sich durch Oberflächenenergien erklären lässt. Die Grenze zwischen diesen Regionen wird als Interface bezeichnet. Um nun den Zustand des Systems zu beschreiben, wird ein Ordnungsparameter u eingeführt. Für eine Legierung beschreibt dabei $u(x)$ die Konzentration der Spezies im Punkt x. Dabei entspricht $u(x) = 1$ dem reinen Vorliegen von Spezies A und $u(x) = -1$ dem reinen Vorliegen von Spezies B im Punkt x. Liegen beide Spezies gemischt vor, nimmt $u(x)$ Werte in $(0; 1)$ an. Wird von einer nur lokal definierten freien Energie ausgegangen, so lässt sich die Gesamtenergie des Systems mit

$$\mathcal{E}(u) = \int_\Omega \psi(u(x), T)dx$$

beschreiben. Vor dem Hintergrund dieser Systemenergie wird $\psi(u(x), T)$ auch als Energiedichte bezeichnet. Um auch die Oberflächenenergie am Interface zu modellieren, haben Cahn und Hilliard in [CH58] den Term $\frac{1}{2}\gamma|\nabla u|^2$ eingeführt und zu $\psi(u, T)$ addiert. Dies führt zu

$$\Psi(u(x), T) = \psi(u(x), T) + \frac{1}{2}\gamma|\nabla u(x)|^2$$

und zur Gesamtenergie

$$\mathcal{E}_\gamma(u) = \int_\Omega \psi(u(x), T) + \frac{1}{2}\gamma|\nabla u(x)|^2 dx.$$

Ein solches Energiefunktional wird als Ginzburg-Landau Energiefunktional bezeichnet. Dabei lässt sich zeigen, dass $\gamma > 0$ proportional zur Dicke des Interfaces im Gleichgewichtszustand ist. Für einen Gleichgewichtszustand ist auch zu erwarten, dass der Ordnungsparameter u dieses Funktional minimiert. Es gibt prinzipiell zwei Arten von Phasenfeldmodellen: Modelle mit und Modelle ohne Massenerhaltung. Die Separation ist der klassische Fall mit Massenerhaltung, da sich die Spezies nur trennen und nicht ineinander umwandeln. Deshalb ist im Gleichgewichtszustand zusätzlich noch die Massenerhaltung

$$\int_\Omega u(x)dx = u_m|\Omega|$$

zu fordern. Soll nun die Dynamik des Systems modelliert werden, so ist u auch von der Zeit abhängig zu betrachten, also als $u(x,t)$. Für Systeme ohne Massenerhaltung nimmt man meistens an, dass die Dynamik sich in Richtung der Minimierung von $\mathcal{E}_\gamma$ entwickelt, also in Richtung der negativen Funktionalableitung von $\mathcal{E}_\gamma$ nach u. Dies führt auf die Evolutionsgleichung

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial u}\mathcal{E}_\gamma(u).$$

1.2 Die Cahn-Hilliard Gleichung

Im Fall der Massenerhaltung wird die Funktionalableitung

$$\langle w, v\rangle = \left\langle \mathcal{E}'_\gamma(u), v\right\rangle = (\psi'(u), v) + \gamma(\nabla u, \nabla v) \tag{1.2.1}$$

als chemisches Potential bezeichnet und als eine thermodynamische Kraft interpretiert. Für den Massenfluss J wird nun $J = -M\nabla w$ gefordert. Dabei ist M ein Mobilitätsfaktor. Damit gilt

$$\frac{\partial u}{\partial t} = \operatorname{div}(M\nabla w). \tag{1.2.2}$$

Die dadurch gegebene Gleichung wird als Cahn-Hilliard Gleichung bezeichnet. Durch Einsetzen der Beziehungen (1.2.1) und (1.2.2) ergibt sich die partielle Differentialgleichung vierter Ordnung

$$\frac{\partial u}{\partial t} = \operatorname{div}(M\nabla(\psi'(u) - \gamma\triangle u)) \qquad x \in \Omega, t > 0.$$

Falls es sich um ein abgeschlossenes System handelt, existiert kein Fluss über den Rand. Damit ergibt sich die Randbedingung

$$M(\nabla w)\cdot\nu = 0 \qquad \text{auf } \partial\Omega.$$

Für u werden natürliche Randbedingungen gewählt, die sich für die Variationsformulierung (1.2.2) ergeben:

$$\gamma(\nabla u)\cdot\nu \qquad \text{auf } \partial\Omega.$$

Des Weiteren ist für ein Evolutionsproblem ein Anfangswert vorzugeben. Dies geschieht durch

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

Die Frage ist nun, wie die Funktion ψ aussieht. Da uniforme Zustände für $T > T_c$ stabil sind, nimmt die Funktion ihr Minimum im Inneren des Intervalls $(-1, 1)$ an.

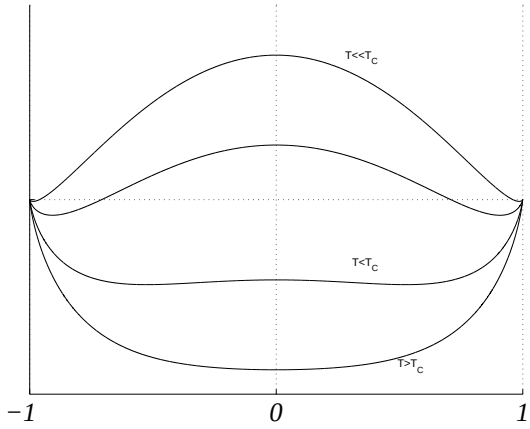

Abbildung 1.2.1: Die Funktion ψ für verschiedene T

Eine oft verwendete Wahl für ψ ist die Funktion

$$\psi(u) = \psi(0, T) + \frac{1}{2}KT_c u^2 + \frac{1}{2}KT\left[(1-u)\log_e(1-u) + (1+u)\,(\,log)_e(1+u)\right],$$

mit

$$\psi(0, T) = \frac{1}{2}KT_c - kT\log_e 2,$$

wobei K die Boltzmannkonstante ist. Sie besitzt die Eigenschaft, für $T > T_c$ in $(-1; 1)$ konvex zu sein. Somit sind uniforme Zustände energiearm. Für $T < T_c$ besitzt ψ zwei Minima $-1 < a < 0 < b < 1$. Diese führen zu Separationsprozessen, da der Zustand $u = 0$ nicht mehr minimale Energie besitzt. Für kleiner werdendes T rücken die Minima immer näher an -1 und 1. Für $u \to \pm 1$ gilt jedoch immer $\psi(u) \to \infty$. Damit ist sichergestellt, dass u nicht größer oder kleiner als 1 wird, was der Interpretation als Konzentration widerspräche. Der Grenzfall, in dem T sehr klein gegenüber T_c wird und gegen 0 geht, ist Thema dieser Arbeit. Beim so genannten deep-quench-limit wird davon ausgegangen, dass das System in einem uniformen Zustand bei $T > T_c$ mit

$$u_0(x) = u_m + \xi(x)$$

für $|\xi(x)| \ll 1$ und $\int_\Omega \xi(x)dx = 0$ ist. Nun wird die Temperatur sprunghaft rapide gesenkt, was $T \ll T_c$ entspricht. Damit wird aus dem differenzierbaren das nicht differenzierbare Funktional

$$\psi(u) = \begin{cases} \frac{1}{2}(1-u^2) & \text{für } |u| \le 1 \\ \infty & \text{für } |u| > 1 \end{cases}.$$

Um $u \in [-1;1]$ sicherzustellen nimmt ψ nun außerhalb dieses Intervalls den Wert ∞ an, womit sich die Bezeichnung Hindernispotential für ψ erklärt. In der Formulierung der Cahn-Hilliard Gleichung tritt die Ableitung von ψ auf, die sich nun nicht mehr bilden lässt. Deshalb wird auf den Begriff des Subdifferentials zurückgegriffen (siehe Abschnitt 2.2). Damit lässt sich das chemische Potential darstellen durch

$$w + \gamma \triangle u + u \in \partial I_{[-1;1]}$$

womit die Cahn-Hilliard Gleichung zu folgendem System wird:

$$\frac{\partial u}{\partial t} = \triangle w, \tag{1.2.3}$$

$$(|u| - 1)(-\gamma \triangle u - u - w) = 0, \tag{1.2.4}$$

$$(\gamma \triangle u + u + w) \operatorname{sign} u \geq 0, \tag{1.2.5}$$

$$|u| \leq 1. \tag{1.2.6}$$

Dabei folgt für $|u(x)| < 1$ aus (1.2.4) wieder genau die Differentialgleichung

$$-\gamma \triangle u - u - w = 0.$$

Multipliziert man diese mit $\eta - u$, wobei $|\eta(x)| \leq 1$ gilt, so ergibt sich

$$(\eta - u)(-\gamma \triangle u - u - w) = 0.$$

Für den Fall $u = 1$ gilt $\eta - u \leq 0$, womit aus (1.2.5)

$$(\eta - u)(-\gamma \triangle u - u - w) \geq 0$$

folgt. Dieselbe Ungleichung folgt auch aus $u = -1$ und $\eta - u \leq 0$. Damit gilt für alle drei Fälle diese Ungleichung. Integration über Ω und Anwenden partieller Integration führt schließlich zu

$$\gamma \int_\Omega \nabla u(x) \cdot \nabla(\eta(x) - u(x))dx - \int_\Omega u(x)(\eta(x) - u(x))dx \geq \int_\Omega w(x)(\eta(x) - u(x))dx.$$

In Kapitel 3 wird diese Gleichung als exakt formuliertes Variationsproblem betrachtet.

2 Funktionalanalytischer Hintergrund

In diesem Kapitel werden einige Resultate der linearen und nichtlinearen Funktionalanalysis zitiert und die benötigten Funktionenräume eingeführt. Ziel ist nicht, eine umfassende Einführung in die jeweiligen Gebiete zu geben, sondern die für die analytische und numerische Behandlung der Cahn-Hilliard Gleichung notwendigen Aussagen zusammenzutragen. Da an den verschiedensten Stellen Aussagen aus den gleichen Themenkreisen verwendet werden, werden diese nicht an der benötigten Stelle zitiert, sondern vorher thematisch geordnet zusammengestellt.

Einfache Korollare und Bemerkungen, die nicht in der benötigten Form in der Literatur zu finden sind, werden bewiesen. Zu den anderen Aussagen sind jeweils Referenzen für die Beweise angegeben. Falls die Sätze in der Referenz in ihrer Allgemeinheit den Rahmen der Arbeit sprengen würden, werden sie in einfachen Versionen zitiert, die offensichtlich in der allgemeinen Formulierung als Spezialfall enthalten sind.

Sofern nicht anders spezifiziert, bezeichnet in diesem Kapitel V einen mit $\|\cdot\|_V$ normierten $\mathbb{R}$-Vektorraum, V' seinen Dualraum, $\langle\cdot,\cdot\rangle : V' \times V \to \mathbb{R}$ die duale Paarung von V und V', H einen Hilbertraum über $\mathbb{R}$, H' dessen Dualraum, $(\cdot,\cdot)_H$ das Skalarprodukt und $\|\cdot\|_H$ die induzierte Norm in H. Nach dem Darstellungssatz von Fréchet-Riesz wird teilweise H mit H' und $\langle\cdot,\cdot\rangle_H$ mit $(\cdot,\cdot)_H$ identifiziert.

Satz 2.0.1 *(Darstellungssatz von Fréchet-Riesz) Sei H ein Hilbertraum, dann existiert zu jedem $x' \in H'$ genau ein $x \in H$ mit $\|x\|_H = \|x'\|_{H'}$ und*

$$x'(y) = \langle x', y\rangle_H = (x,y)_H \qquad \forall y \in H.$$

Die lineare, bijektive, isometrische Abbildung $R_H : H \to H'$ mit $R_H x = x'$ heißt Riesz-Isomorphismus.

BEWEIS. Der Beweis ist zum Beispiel in [Wer00] gegeben. ■

Korollar 2.0.1 *Da R_H linear, normerhaltend und bijektiv ist, ist sofort klar, dass H' ein Hilbertraum ist, dessen Norm vom Skalarprodukt*

$$(x', y')_{H'} = (R_H^{-1}x', R_H^{-1}y')_H = (x,y)_H$$

erzeugt wird. ■

2.1 Schwache Konvergenz

Ein Standardresultat der Analysis ist die Existenz konvergenter Teilfolgen von beschränkten Folgen in endlichdimensionalen normierten Räumen. Sie folgt aus der Kompaktheit der abgeschlossenen Einheitskugel. In unendlichdimensionalen Räumen ist diese Aussage nicht nur falsch, es gilt sogar folgende Charakterisierung endlichdimensionaler Vektorräume:

Satz 2.1.1 *Folgende Aussagen sind äquivalent:*

(1) V *ist endlichdimensional.*

(2) *Die abgeschlossene Einheitskugel* $B_V := \{x \in V : \|x\|_V \leq 1\}$ *ist kompakt.*

(3) *Jede beschränkte Folge in* V *besitzt eine konvergente Teilfolge.*

BEWEIS. Siehe [Wer00]. ■

Um die fehlende Kompaktheit der Einheitskugel in unendlichdimensionalen Räumen auszugleichen, wird ein abgeschwächter Konvergenzbegriff betrachtet, der für bestimmte Räume zu einer ähnlichen Aussage wie (3) führt. Eine ausführliche Darstellung inklusive der hier nicht ausgeführten Beweise ist beispielsweise bei [Wer00] und in anderen Lehrbüchern zur Funktionalanalysis zu finden.

Definition 2.1.1 Eine Folge (x_n) in V heißt schwach konvergent (Notation: $x_n \rightharpoonup x$) gegen $x \in V$, wenn gilt:

$$x'(x_n) \xrightarrow[n\to\infty]{} x'(x) \qquad \forall x' \in V'.$$

In einem Hilbertraum lässt sich die schwache Konvergenz aufgrund des Darstellungssatzes von Fréchet-Riesz eleganter darstellen durch:

$$x_n \rightharpoonup x \quad \Leftrightarrow \quad (x_n, y)_H \to (x, y)_H \qquad \forall y \in H.$$

Die Eindeutigkeit des Grenzwertes in normierten Räumen ist klar. Auch die Tatsache, dass Grenzwerte beschränkter Folgen durch die gleiche Konstante beschränkt sind, folgt direkt aus der Konvergenz. Um diese Eigenschaften für schwach konvergente Folgen zu zeigen, wird folgendes Korollar aus dem Satz von Hahn-Banach benötigt:

Satz 2.1.2 *Es sei* V *ein normierter Raum. Zu* $x \in V \setminus \{0\}$ *existiert ein* $x' \in V'$ *mit* $\|x'\|_{V'} = 1$ *und* $x'(x) = \|x\|_V$. *Zu* $x_1, x_2 \in V \setminus \{0\}$ *mit* $x_1 \neq x_2$ *existiert ein* $x' \in V'$ *mit* $x'(x_1) \neq x'(x_2)$.

BEWEIS. Siehe [Wer00]. ■

Bemerkung 2.1.1

(1) *Konvergente Folgen konvergieren schwach gegen denselben Grenzwert.*

(2) *Der schwache Grenzwert einer Folge ist eindeutig bestimmt.*

(3) *Der schwache Grenzwert einer beschränkten Folge ist durch die gleiche Konstante beschränkt wie die Folge.*

BEWEIS. (1) ist klar, da für normierte Räume $f \in V'$ folgenstetig ist.
(2) Es gelte $x_n \rightharpoonup x$ und $x_n \rightharpoonup \tilde{x}$ in V, dann gilt $x = \tilde{x}$, sonst existiert nach Satz 2.1.2 $x' \in V'$ mit

$$\lim_{n\to\infty} x'(x_n) = x'(x) \neq x'(\tilde{x}) = \lim_{n\to\infty} x'(x_n).$$

(3) Es gelte $\|x_n\| \leq C$ und $x_n \rightharpoonup x$ in V. Ist $x = 0$, so ist $\|x\|_V \leq C$ klar. Sei nun $x \neq 0$ und $x' \in V'$ gemäß Satz 2.1.2 so gewählt, dass $x'(x) = \|x\|_V$ und $\|x'\|_{V'} = 1$ gilt, dann gilt

$$\|x\|_V = x'(x) = \lim_{n\to\infty} x'(x_n) \leq \limsup_{n\to\infty} |x'(x_n)| \leq \limsup_{n\to\infty} \|x'\|_{V'}\|x_n\|_V \leq C.$$

■

Nun zitieren wir das wesentliche Resultat, dass die schwache Konvergenz für die zu betrachtenden Funktionenräume interessant macht.

Satz 2.1.3 *Sei V ein reflexiver normierter Raum und (x_n) eine beschränkte Folge in V, dann existiert eine schwach konvergente Teilfolge (x_{n_k}) von (x_n).*

BEWEIS. Siehe [Wer00]. ■

Die schwache Konvergenz „ersetzt“ also in reflexiven Räumen in gewissem Sinne die Kompaktheit der Einheitskugel. In der Literatur ist dieser Satz oft für reflexive Banachräume zu finden, was nur scheinbar eine zusätzliche Voraussetzung ist. Da Dualräume und Bidualräume von normierten Räumen stets vollständig sind, muss auch ein reflexiver Raum vollständig sein. Die umgekehrte Beziehung von Beschränktheit und schwacher Konvergenz zeigt das folgende Korollar aus dem Satz von Banach-Steinhaus.

Satz 2.1.4 *Sei V ein normierter Raum, dann ist für $M \subset V$ äquivalent:*

(1) *M ist beschränkt.*

(2) *Für jedes $x' \in V'$ ist $\{x'(x) : x \in M\}$ beschränkt.*

BEWEIS. Siehe [Wer00]. ■

Da konvergente Folgen $(x'(x_n))$ in $\mathbb{R}$ beschränkt sind, zeigt dies insbesondere, dass schwach konvergente Folgen beschränkt sind. Damit lässt sich auch eine etwas allgemeinere Charakterisierung schwacher Konvergenz zeigen.

Bemerkung 2.1.2 *Sei V ein normierter Raum, dann sind für eine Folge (x_n) in V folgende Aussagen äquivalent:*

(1) *$x_n \rightharpoonup x$, d.h. (x_n) ist schwach konvergent.*

(2) *$y'_n(x_n) \to y'(x)$ gilt für alle konvergenten Folgen $y'_n \to y'$ in V'.*

BEWEIS. (2) $\Rightarrow$ (1) ist klar, da konstante Folgen in V' konvergent sind.
Für (1) $\Rightarrow$ (2) gilt, da (x_n) nach Satz 2.1.4 beschränkt ist,

$$\begin{aligned} |y'_n(x_n) - y'(x)| &= |(y'_n - y')(x_n) + y'(x_n - x)| \\ &\leq \|x_n\|_V \|y'_n - y'\|_{V'} + |y'(x_n) - y'(x)| \xrightarrow[n\to\infty]{} 0. \end{aligned}$$

■

Auch in nicht reflexiven Räumen gibt es noch ein Konzept, das die fehlende Kompaktheit der Einheitskugel in gewissem Maße „ausgleicht“: die schwach-* Konvergenz.

Definition 2.1.2 Es sei V ein normierter Vektorraum. Eine Folge (x'_n) in V' heißt schwach-* konvergent gegen $x' \in V'$ (Notation: $x'_n \rightharpoonup^* x'$), falls gilt:

$$x'_n(y) \to x'(y) \qquad \forall y \in V.$$

Bemerkung 2.1.3

(1) *Konvergente Folgen in V' konvergieren schwach-* gegen denselben Grenzwert.*

(2) *Der schwach-* Grenzwert einer Folge ist eindeutig bestimmt.*

(3) *Der schwach-* Grenzwert einer beschränkten Folge ist durch die gleiche Konstante beschränkt wie die Folge.*

BEWEIS. (1) Es gelte $x'_n \to x'$ in V', also $\|x'_n - x'\|_{V'} \to 0$, dann gilt für $y \in V$

$$|(x'_n - x')(y)| \leq \|x'_n - x'\|_{V'} \|y\|_V \to 0,$$

also $x'_n(y) \to x'(y)$.
(2) ist klar, da $\lim x'_n(y)$ in $\mathbb{R}$ eindeutig bestimmt ist und da zwei Funktionale identisch sind, wenn sie auf allen Elementen von V übereinstimmen.
(3) Es gelte $\|x'_n\|_{V'} \leq C$ und $x'_n \rightharpoonup^* x'$ in V', dann gilt für jedes $y \in V$ und für die reelle Folge $|x'_n(y)| \in \mathbb{R}$

$$|x'_n(y)| \leq \|x'_n\|_{V'} \|y\|_V \leq C\|y\|_V$$

und damit für deren Grenzwert $|x'(y)| = |\lim x'_n(y)| \leq C\|y\|_V$, woraus nach Definition der Norm in V' $\|x'\|_{V'} \leq C$ folgt. ■

Satz 2.1.5 *Sei V ein separabler normierter Raum und (x'_n) eine beschränkte Folge in V', dann existiert eine schwach-* konvergente Teilfolge (x'_{n_k}) von (x'_n).*

BEWEIS. Siehe [Wer00]. ■

Auch hier zeigt ein Korollar aus dem Satz von Banach-Steinhaus die umgekehrte Beziehung von Beschränktheit und schwacher Konvergenz, doch ist nun zusätzlich die Vollständigkeit von V zu fordern.

Satz 2.1.6 *Sei V ein Banachraum, dann ist für $M \subset V'$ äquivalent:*

(1) *M ist beschränkt.*

(2) *Für jedes $x \in V$ ist $\{x'(x) : x' \in M\}$ beschränkt.*

BEWEIS. Siehe [Wer00]. ■

Analog zur schwachen Konvergenz folgt nun aus der Beschränktheit konvergenter Folgen in $\mathbb{R}$ die Beschränktheit schwach-* konvergenter Folgen für Banachräume V. Auch hier lässt dies eine allgemeinere Charakterisierung schwach-* konvergenter Folgen zu.

Bemerkung 2.1.4 *Sei V ein Banachraum, dann sind für eine Folge (x'_n) in V' folgende Aussagen äquivalent:*

(1) $x'_n \rightharpoonup^* x'$.

(2) $x'_n(y_n) \to x'(y)$ *gilt für alle konvergenten Folgen* $y_n \to y$ *in* V.

BEWEIS. Der Beweis läuft analog zum Beweis von Bemerkung 2.1.2. Es sind lediglich die Elemente aus V und V' zu tauschen und die Beschränktheit schwach-* konvergenter Folgen für Banachräume V zu nutzen. ■

Wie die Begriffe schwache Konvergenz und schwach-* Konvergenz vermuten lassen, lassen sich beide auch durch Topologien, die sogenannte schwache und die schwach-* Topologie, erklären. Beide sind lokalkonvexe Topologien auf dem jeweiligen Vektorraum. Mit diesem Ansatz lässt sich folgender allgemeinerer Satz formulieren:

Satz 2.1.7 *Sei V ein normierter Raum und B_V beziehungsweise $B_{V'}$ die abgeschlossene Einheitskugel im jeweiligen Raum, dann gilt:*

(1) $B_{V'}$ *ist schwach-*-kompakt.*

(2) B_V *ist genau dann schwach-kompakt, wenn V reflexiv ist.*

(3) B_V *ist genau dann $\|\cdot\|_V$-kompakt, wenn V endlichdimensional ist.*

BEWEIS. Siehe [Wer00]. ■

Es ist jedoch zu beachten, dass in (1) und (2) von Kompaktheit in topologischen Räumen gesprochen wird, die nicht mit Folgenkompaktheit gleichzusetzen ist. Deshalb folgt aus (1) und (2) für beschränkte Folgen auch nicht direkt die Existenz schwach und schwach-* konvergenter Teilfolgen. Bei (2) lässt sich diese trotzdem zeigen; bei (1) ist zusätzlich Separabilität von V zu fordern.

2.2 Konvexe Analysis

Eine klassiches Problemstellung der reinen und angewandten Mathematik ist die Minimierung einer Funktion $f : V \to \mathbb{R}$ in einer Teilmenge M eines Vektorraums V. Für $M = V$ und „glattes“ f stellt der klassische Ableitungsbegriff (siehe etwa [Rud64]) ein gutes Werkzeug zur Untersuchung solcher Probleme dar. Ist f jedoch nicht glatt oder M eine echte Teilmenge von V, so ist diese Theorie meist weniger hilfreich. Die Theorie der konvexen Analysis stellt auch für nichtglatte Funktionen elegante Werkzeuge zur Handhabung von Minimierungsproblemen

bereit. Sie basiert im Wesentlichen auf konvexen Funktionen $f : V \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\infty, -\infty\}$ und erlaubt, auf die Differenzierung der nichtglatten Anteile von f zu verzichten.

Eine detaillierte Einführung in dieses Gebiet ist etwa bei [ET76] und [Roc70], die Beweise der hier zitierten Sätze bei [ET76] zu finden. In diesem Abschnitt bezeichnet V einen Banachraum.

Definition 2.2.1 Es sei eine Abbildung $f : V \to \overline{\mathbb{R}}$ und eine Menge $K \subset V$ gegeben. Definiere dann:

(1) K heißt konvex, wenn $\lambda x + (1-\lambda)y \in K$ für jedes $x, y \in K$ und jedes $\lambda \in [0;1]$ gilt.

(2) Die Indikatorfunktion der Menge K ist $I_K : V \to \overline{\mathbb{R}}$ mit

$$I_K(x) = \begin{cases} 0 & \text{für } x \in K \\ \infty & \text{für } x \notin K \end{cases}.$$

(3) f heißt konvex, wenn für jedes $x, y \in V$ und jedes $\lambda \in [0;1]$ gilt:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

f heißt strikt konvex, wenn die Ungleichung mit $<$ gilt.

(4) f heißt proper, wenn $f > -\infty$ gilt und ein $x \in V$ mit $f(x) \neq \infty$ existiert.

(5) f heißt unterhalbstetig, wenn für jede konvergente Folge (x_n) in V gilt:

$$f(\lim_{n\to\infty} x_n) \leq \liminf_{n\to\infty} f(x_n).$$

(6) $\operatorname{dom} f := \{x \in V : f(x) \text{ ist endlich}\}$

Funktionen die konvex, proper und unterhalbstetig sind, müssen keinesfalls differenzierbar sein, wie zum Beispiel die Betragsfunktion $|\cdot| : \mathbb{R} \to \mathbb{R}$ und die Indikatorfunktion I_K einer konvexen Menge K zeigen. Deshalb wird der Begriff des Subdifferentials eingeführt, der eine einfache Charakterisierung der Minimalstellen solcher Funktionen zulässt.

Definition 2.2.2 Es sei eine Abbildung $f : V \to \overline{\mathbb{R}}$ gegeben. Ein Funktional $f'(x) \in V'$ heißt Subgradient von f an der Stelle x, falls $f(x)$ endlich ist und

$$f(x) + f'(x)(y-x) = f(x) + \langle f'(x), y-x \rangle_V \leq f(y) \qquad \forall y \in V$$

gilt. Die Menge $\partial f(x)$ aller solcher Subgradienten $f'(u) \in V'$ heißt Subdifferential von f an der Stelle x.

Für Hilberträume H werden die Subgradienten in H' nach dem Darstellungssatz von Fréchet-Riesz mit Elementen aus H identifiziert und statt der dualen Paarung $\langle \cdot, \cdot \rangle_H$ das Skalarprodukt $(\cdot, \cdot)_H$ betrachtet, womit $\partial f(x) \subset H$ gilt. Ein besonderes Beispiel ist der Subgradient ∂I_K der Indikatorfunktion einer nichtleeren konvexen Teilmenge K eines Hilbertraums H. Nach Definition des Subgradienten gilt hier

$$\begin{aligned} f \in \partial I_K(x) \qquad &\Leftrightarrow \qquad (f, y-x)_H + I_K(x) \leq I_K(y) \qquad && \forall y \in V \\ &\Leftrightarrow \qquad x \in K \text{ und } (f, y-x)_H \leq 0 && \forall y \in K. \end{aligned}$$

Bemerkung 2.2.1 *Es sei $K \subset V$ konvex, abgeschlossen und nicht leer, dann ist I_K konvex, proper und unterhalbstetig.*

BEWEIS. Da $K \neq \varnothing$ gefordert war, ist I_K proper und da K konvex ist, zeigt eine einfache Fallunterscheidung, dass auch I_K konvex ist. Aufgrund der Abgeschlossenheit von K kann $I_K(\lim x_n) = \infty$ nur gelten, falls nur endlich viele der x_n in K liegen, woraus folgt, dass $\liminf I_K(x_n) = \infty = I_K(\lim x_n)$ gilt. Mit $I_K(\lim x_n) = 0$ ist die Bedingung $I_K(\lim x_n) \leq \liminf I_K(x_n)$ trivial erfüllt. Also ist auch I_K unterhalbstetig. ■

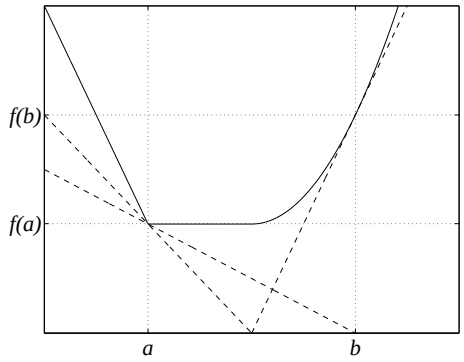

Abbildung 2.2.1: Subgradienten von f an den Punkten $(a, f(a))$ und $(b, f(b)$

Im Gegensatz zum klassischen Ableitungsbegriff, der nach geeigneter Verschiebung eine affin-lineare Tangente an einen Punkt beschreibt, ist ein Subgradient nach geeigneter Verschiebung eine affin-lineare Minorante einer Funktion. Die Menge aller dieser Minoranten ist das Subdifferential. Falls f an einem Punkt eine Tangente besitzt, ist zu erwarten, dass das Subdifferential nur diese enthält. Besitzt f jedoch einen „Knick“, ist zu erwarten, dass es mehrere Subgradienten gibt. Der Zusammenhang zur klassischen Ableitung wird noch genauer betrachtet. Besondere Bedeutung besitzen Subdifferentiale, da sich mit ihnen Minimalstellen sehr einfach und allgemein wie folgt charakterisieren lassen:

Bemerkung 2.2.2 *Für Abbildungen $f : V \to \overline{\mathbb{R}}$ gilt*

$$f(x) \leq f(y) \quad \forall y \in V \quad \Leftrightarrow \quad 0 \in \partial f(x).$$

■

Ferner gelten für Subdifferentiale folgende Regeln.

Satz 2.2.1 *Für Abbildungen $f_1, f_2 : V \to \overline{\mathbb{R}}$, $\lambda > 0$ und $x \in V$ gilt:*

(1) $\partial(\lambda f_1)(x) = \lambda \partial f_1(x)$.

(2) *Falls $f_1(V), f_2(V) \subset \mathbb{R}$, so gilt $\partial f_1(x) + \partial f_2(x) \subset \partial(f_1 + f_2)(x)$.*

(3) *Sind f_1 und f_2 konvex, proper und unterhalbstetig und existiert ein Punkt $\bar{x} \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$ an dem f_1 stetig ist, so gilt $\partial f_1(x) + \partial f_2(x) = \partial(f_1 + f_2)(x)$.*

BEWEIS. Siehe [ET76]. ■

Nun wird der Zusammenhang zwischen der Gateaux-Ableitung und dem Subdifferential von Funktionen spezifiziert.

Definition 2.2.3 Es sei eine Abbildung $f : V \to \overline{\mathbb{R}}$ gegeben. Dann heißt $f_y(x) \in \overline{\mathbb{R}}$ Richtungsableitung von F an der Stelle x in Richtung y, falls

$$f_y(x) = \lim_{h \to 0} \frac{f(x + hy) - f(x)}{h}$$

gilt. Gilt $Df(x) := f_{(\cdot)}(x) \in V'$, so heißt $Df(x)$ Gateaux-Ableitung von f an der Stelle x.

Satz 2.2.2 *Für konvexe Abbildungen $f : V \to \overline{\mathbb{R}}$ gilt:*

(1) *Ist f an einer Stelle $x \in V$ Gateaux-differenzierbar, so gilt $\partial f(x) = \{Df(x)\}$.*

(2) *Ist f an einer Stelle $x \in V$ stetig, endlich und besitzt dort nur einen Subgradienten, so ist f an dieser Stelle Gateaux-differenzierbar und es gilt $\partial f(x) = \{Df(x)\}$.*

BEWEIS. Siehe [ET76]. ■

Korollar 2.2.1 *Sei $K \subset V$ eine abgeschlossene, konvexe, nichtleere Menge und $f : V \to \mathbb{R}$ konvex, stetig und Gateaux-differenzierbar, dann sind f, I_K und $f + I_K$ konvex, proper und unterhalbstetig und es gilt $Df + \partial I_K = \partial(f + I_K)$. Ferner sind folgende Aussagen äquivalent:*

(1) *Es gilt $0 \in \partial(f + I_K)(x)$.*

(2) *Es gilt $-Df(x) \in \partial I_K(x)$.*

(3) *Es gilt $x \in K$ und $f(x) \leq f(y)$ für alle $y \in K$.*

(4) *Es gilt $x \in K$ und $\langle Df(x), y - x \rangle_V \geq 0$ für alle $y \in K$.*

BEWEIS. Aus der Stetigkeit von f folgt die Unterhalbstetigkeit von f und wegen $f(V) \subset \mathbb{R}$ ist f offensichtlich proper. I_K ist nach Bemerkung 2.2.1 konvex, unterhalbstetig und proper. Damit ist auch $f + I_K$ konvex und proper. Mit

$$f(\lim_{n\to\infty} x_n) + I_K(\lim_{n\to\infty} x_n) \leq \lim_{n\to\infty} f(x_n) + \liminf_{n\to\infty} I_K(x_n) = \liminf_{n\to\infty}(f(x_n) + I_K(x_n))$$

ist auch die Unterhalbstetigkeit gezeigt. Nach Satz 2.2.1 und Satz 2.2.2 gilt nun

$$Df + \partial I_K = \partial f + \partial I_K = \partial(f + I_K).$$

Damit ist die Äquivalenz von (1) und (2) klar. (1) ist nun nach Definition des Subdifferentials äquivalent zu

$$x \in V : \qquad f(x) + I_K(x) \leq f(y) + I_K(y) \qquad \forall y \in V,$$

also zu (3). Analog dazu ist (2) äquivalent zu

$$x \in V: \qquad \langle -Df(x), y - x\rangle_V + I_K(x) \leq I_K(y) \qquad \forall y \in V,$$

also zu (4). ■

Nun wird noch ein Existenzresultat für die Lösung von restringierten Minimierungsproblemen zitiert.

Satz 2.2.3 *Es sei V ein reflexiver Banachraum und $K \subset V$ konvex, abgeschlossen und nicht leer. $f : K \to \mathbb{R}$ sei konvex, proper und unterhalbstetig. Ferner sei K beschränkt oder es gelte $|f(x)| \to \infty$ für $x \in K$ und $\|x\|_V \to \infty$. Dann besitzt das Minimierungsproblem*

$$x \in K: \qquad f(x) \leq f(y) \qquad \forall y \in K$$

eine Lösung. Ist f strikt konvex in K, so ist die Lösung eindeutig bestimmt.

BEWEIS. Siehe [ET76]. ■

Die Eigenschaft

$$\|x\|_V \to \infty \Rightarrow |f(x)| \to \infty$$

wird auch als Koerzivität bezeichnet und ist nicht mit der Koerzivität einer symmetrischen, stetigen Bilinearform auf $V \times V$

$$a(x, x) \geq \alpha \|x\|_V \qquad \forall x \in V, \alpha > 0$$

zu verwechseln. Ist jedoch eine symmetrische Bilinearform $a(\cdot, \cdot)$ koerziv, so ist für jedes $l \in V'$ das Funktional $f(x) = \frac{1}{2}a(x, x) - l(x)$ konvex und koerziv.

2.3 Maximal monotone Operatoren

Im Hinblick auf das in der konvexen Analysis entwickelte Subdifferential erscheint es sinnvoll, mengenwertige Operatoren zu untersuchen. Besondere Bedeutung für die Betrachtung nichtlinearer Operatoren besitzt die Theorie der maximal monotonen Operatoren. Eine detaillierte Einführung in diese Theorie inklusive der hier nicht ausgeführten Beweise ist zum Beispiel bei [Zei85a] zu finden.

Im Folgenden sei V ein reeller reflexiver Banachraum. Der Dualraum H' des Hilbertraums H wird nach dem Darstellungssatz von Fréchet-Riesz mit H und die duale Paarung $\langle \cdot, \cdot \rangle_H$ mit dem Skalarprodukt $(\cdot, \cdot)_H$ identifiziert. $2^Y = \{y \subset Y\}$ bezeichne die Potenzmenge von Y. Statt der „üblichen“ Operatoren $A : V \to V'$ werden nun mengenwertige Operatoren $A : V \to 2^{V'}$ und $A : H \to 2^H$ betrachtet.

Definition 2.3.1 Für eine Abbildung $A : X \to 2^Y$ bezeichnet

$$D(A) := \{x \in X : Ax \neq \varnothing\}$$

den effektiven Definitionsbereich von A,

$$R(A) := \bigcup_{x \in X} Ax$$

den Wertebereich von A und

$$G(A) := \{(x, y) \in X \times Y : x \in D(A), y \in Ax\}$$

den Graphen von A. Es wird die Notation $(x, y) \in A :\Leftrightarrow (x, y) \in G(A)$ verwendet.A heißt einwertig, wenn das Bild jedes $x \in X$ eine einelementige Menge $Ax = \{\bar{A}x\}$ ist. In diesem Fall wird A mit $\bar{A} : X \to Y$ identifiziert. Insbesondere gilt dann $D(A) = X$. Die „inverse“ Abbildung $A^{-1} : Y \to 2^X$ zu A ist durch

$$A^{-1}y := \{x \in X : y \in Ax\}$$

definiert. Sie existiert immer, aber es kann $A^{-1}y = \varnothing$ gelten.

Definition 2.3.2 Sei $X \subset V$, dann heißt ein Operator $A : X \to 2^{V'}$ monoton, falls

$$\langle y_1 - y_2, x_1 - x_2 \rangle_V \geq 0 \qquad \forall (x_1, y_1), (x_2, y_2) \in A$$

gilt. A heißt maximal monoton, wenn zusätzlich $(x_1, y_1) \in A$ aus

$$(x_1, y_1) \in X \times V' : \qquad \langle y_1 - y_2, x_1 - x_2 \rangle_V \geq 0 \qquad \forall (x_2, y_2) \in A,$$

folgt.

Nun werden einige Resultate für maximal monotone Operatoren präsentiert.

Satz 2.3.1 *Es sei $f : V \to \overline{\mathbb{R}}$ konvex, proper und unterhalbstetig, dann ist $\partial f : V \to 2^{V'}$ maximal monoton.*

BEWEIS. Siehe [Zei85a]. ■

Definition 2.3.3 Die Dualitätsabbildung $J_V : V \to 2^{V'}$ von V ist durch $J_V = \partial\phi$ definiert, wobei $\phi(u) = \frac{1}{2}\|u\|_V^2$ auf V gilt.

Bemerkung 2.3.1 *Für einen Hilbertraum H ist J_H einwertig und es gilt $J_H = R_H$.*

BEWEIS. Es gilt $D\phi(x) = (x, \cdot)_H$, denn für $x, y \in H$ und $h \in \mathbb{R}$ gilt

$$\lim_{h \to 0} \frac{\phi(x + hy) - \phi(x)}{h} = \lim_{h \to 0} \left((x, y)_H + \frac{1}{2} h \|y\|_H^2 \right) = (x, y).$$

Da ϕ konvex ist, folgt aus Satz 2.2.2 $\partial\phi = \{D\phi\}$, also $\langle J_H x, y \rangle = (x, y)_H$. ■

Identifiziert man einen Hilbertraum H nach dem Darstellungssatz von Fréchet-Riesz mit seinem Dualraum und betrachtet $J_H : H \to H$, so gilt $J_H = I$.

Der folgende Satz gilt für strikt konvexe Räume, also für Räume, in denen aus $\|x_1\|_V = 1 = \|x_2\|_V$ und $\frac{\|x_1 + x_2\|_V}{2} = 1$ $x_1 = x_2$ folgt.

Satz 2.3.2 *Sei V strikt konvex und die Abbildung $A : V \to 2^{V'}$ maximal monoton, dann ist für jedes $\lambda > 0$ die Abbildung*

$$(J_V + \lambda A)^{-1} : V' \to V$$

einwertig und maximal monoton.

BEWEIS. Siehe [Zei85a]. ■

Bemerkung 2.3.2 *Hilberträume sind strikt konvex.*

BEWEIS. Es seien $x_1, x_2 \in H$ mit $\|x_1\|_H = \|x_2\|_H = \frac{\|x_1+x_2\|_H}{2} = 1$ gegeben, dann gilt

$$\|x_1 + x_2\|_H^2 + \|x_1 - x_2\|_H^2 = 2\|x_1\|_H^2 + 2\|x_2\|_H^2 = 4 = \|x_1 + x_2\|_H^2.$$

■

2.4 Gelfand'sche Dreier

Für die Lösbarkeit von Evolutionsgleichungen ist die Wahl des Raumes, in dem die Lösungen liegen sollen, entscheidend. Sehr häufig werden Gleichungen der Form $\frac{\partial u}{\partial t} + A(u) = 0$ in einem Banachraum V betrachtet, wobei $A(u) \in V'$ gilt. Dies legt nahe, die Zeitableitung in einem anderen Raum als u zu suchen. Eine wichtige Rolle bei der allgemeinen Betrachtung solcher Zusammenhänge spielen bestimme Tripel von Räumen, so genannte Gelfand'sche Dreier, die im Folgenden eingeführt werden.

Lemma 2.4.1 *Seien X und Y Banachräume und $T : X \to Y$ ein linearer, stetiger Operator, so gilt für den Kern des adjungierten Operators T'*

$$\overline{\operatorname{im} T} = (\ker T')_\perp,$$

wobei $U_\perp = \{x \in Y : x'(x) = 0 \quad \forall x' \in U\}$ für $U \subset Y'$ der Annihilator von U in Y genannt wird. Insbesondere ist T' genau dann injektiv, wenn $\operatorname{im} T$ dicht in Y liegt.

BEWEIS. Der Beweis ist in [Wer00], S.112, gegeben. ■

Das Lemma impliziert insbesondere, dass $\operatorname{im} T \subset Y$ genau dann dicht liegt, wenn $x' \in (\ker T')$ alle $x \in Y$ auf 0 abbildet, also wenn T' injektiv ist.

Lemma 2.4.2 *Sei V ein reflexiver Banachraum, H ein Hilbertraum und $V \hookrightarrow H$ stetig, injektiv und dicht eingebettet, dann ist auch $H \hookrightarrow V'$ stetig, injektiv und dicht eingebettet.*

BEWEIS. Es sei $i : V \to H$ die gegebene Einbettung. Da V und der Hilbertraum H reflexiv sind, sind die kanonischen Isometrien $i_V : V \to V''$ und $i_H : H \to H''$ bijektiv und es gilt $i'' \circ i_V = i_H \circ i$, also ist mit i auch i'' injektiv. Da $\operatorname{im} i \subset H$ dicht ist, liefert die Anwendung

von Lemma 2.4.1 auf i, dass i' injektiv ist. Die Anwendung auf i' liefert, da i'' injektiv ist, dass $\operatorname{im} i' \subset V'$ dicht ist. Sei nun $R_H : H \to H'$ der isometrische Riesz-Isomorphismus, der durch den Darstellungssatz von Frechét-Riesz gegeben ist, dann ist auch $(i' \circ R_H) : H \to V'$ stetig. Aus der Bijektivität von R folgt nun, da i' injektiv ist, die Injektivität von $i' \circ R_H$ und ferner, dass $\operatorname{im}(i' \circ R_H) = \operatorname{im} i' \subset V'$ dicht liegt. ■

Definition 2.4.1 Unter den Bedingungen von Lemma 2.4.2 heißt das Tripel

$$V \hookrightarrow H \hookrightarrow V'$$

Gelfand'scher Dreier.

Für Evolutionsgleichungen wird nun im Allgemeinen gefordert, dass die Lösung zu jedem Zeitpunkt in einem Raum V und die Zeitableitung der Lösung in dessen Dualraum V' liegt, wobei $V \hookrightarrow H \hookrightarrow V'$ einen Gelfand'schen Dreier bilden. Besondere Bedeutung haben stetige Einbettungen auch für konvergente Folgen, wie die nächste Bemerkung zeigt:

Bemerkung 2.4.1 *Es seien V und H normierte Räume und $V \hookrightarrow H$ stetig eingebettet.*

(1) *Es ist auch $H' \hookrightarrow V'$ stetig eingebettet.*

(2) *Es gelte $x_n \to x$ in V, dann gilt auch $x_n \to x$ in H.*

(3) *Es gelte $x_n \rightharpoonup x$ in V, dann gilt auch $x_n \rightharpoonup x$ in H.*

(4) *Es gelte $x'_n \rightharpoonup^* x'$ in H', dann gilt auch $x'_n \rightharpoonup^* x'$ in V'.*

In (2), (3) und (4) wurde jeweils $x \in V$ mit $i(x) \in H$ und $x' \in H'$ mit $i'(x') \in V'$ identifiziert, wobei $i : V \to H$ die stetige Einbettung von V in H und $i' : H' \to V'$ die stetige Einbettung von H' in V' nach (1) ist.

BEWEIS. Um die Übersicht zu wahren, wird im Beweis die Identifizierung mittels i nicht implizit vorgenommen.
(1) Es sei $x' \in H'$ gegeben. Nach Definition des adjungierten Operators i' von i gilt dann $i' : H' \to V'$ und $i'(x') := x' \circ i$. Dabei ist i' stetig mit $\|i'(x')\|_{V'} = \|x' \circ i\| \leq \|x'\|_{H'}\|i\|$.
(2) Es gelte $x_n \to x$ in V, dann gilt

$$\|i(x_n) - i(x)\|_H \leq \|i\|\|x_n - x\|_V \to 0.$$

(3) Es gelte $x_n \rightharpoonup x$ in V. Sei nun $y' \in H'$ gegeben, dann gilt $i'(y') \in V'$ und damit

$$y'(i(x_n)) = (i'(y'))(x_n) \to (i'(y'))(x) = y'(i(x)).$$

(4) Es gelte $x'_n \rightharpoonup^* x'$ in H'. Sei nun $y \in V$ gegeben, dann gilt $i(y) \in H$ und damit

$$(i'(x'_n))(y) = x'_n(i(y)) \to x'(i(y)) = (i'(x'))(y).$$

■

2.5 Sobolevräume

Für partielle Differentialgleichungen werden Lösungen oftmals nicht hinsichtlich der Formulierung mit klassischen Ableitungen, sondern hinsichtlich Variationsformulierungen in Hilberträumen gesucht. Wie die Sätze aus den vorherigen Abschnitten zeigen, gelten hier teilweise fast so starke Aussagen wie in endlichdimensionalen Räumen. Die Räume klassisch differenzierbarer Funktionen sind jedoch im Allgemeinen keine Hilberträume, weshalb aufbauend auf den L^p-Räumen die so genannten Sobolevräume betrachtet werden. Zu jedem Zeitpunkt soll die Phasenvariable in einem "geeigneten" Sobolevraum liegen. Eine ausführliche Darstellung dieser Theorie ist etwa bei [Wlo82] und [Ada75] zu finden.

Definition 2.5.1 Es sei $\Omega \subset \mathbb{R}^d$, $d \in \{1,2,3\}$, offen, nicht leer und beschränkt. Definiere dann:

(1) Der Träger einer Funktion $\phi : \Omega \to \mathbb{R}$ ist $\mathrm{supp}(\phi) := \overline{\{x \in \Omega : \phi(x) \neq 0\}}$. Der Raum der Testfunktionen ist

$$\mathcal{D}(\Omega) = \{\phi \in C^\infty(\Omega) : \mathrm{supp}(\phi) \subset \Omega \text{ und } \mathrm{supp}(\phi) \text{ ist kompakt}\}.$$

Er ist versehen mit einer auf partiellen Ableitungen basierenden lokalkonvexen Topologie (siehe etwa [Wer00]). Sein Dualraum $(\mathcal{D}(\Omega))'$ bezüglich dieser Topologie ist der Raum der Distributionen. Existiert zu $f \in (\mathcal{D}(\Omega))'$ ein $\tilde{f} \in L^2(\Omega)$ mit $(\tilde{f}, \phi) = f(\phi)$ für alle $\phi \in \mathcal{D}(\Omega)$, so heißt f reguläre Distribution und wir schreiben $f = \tilde{f} \in L^2(\Omega)$.

(2) Die distributionelle Ableitung von $f \in (\mathcal{D}(\Omega))'$ in Richtung x_i ist die Distribution

$$\frac{\partial f}{\partial x_i} : \mathcal{D}(\Omega) \to \mathbb{R}, \qquad \phi \mapsto -f\left(\frac{\partial \phi}{\partial x_i}\right),$$

wobei $\partial\phi/\partial x_i$ die klassische Ableitung von ϕ ist. Gilt $f \in L^2(\Omega)$ und $\frac{\partial f}{\partial x_i} \in L^2(\Omega)$, so heißt $\frac{\partial f}{\partial x_i}$ schwache Ableitung von f.

(3) Ω genüge zusätzlich der gleichmäßigen Kegelbedingung (siehe [Wlo82]). Die Sobolevräume $H^1(\Omega)$ und $H^2(\Omega)$ sind dann definiert durch

$$H^1(\Omega) := \left\{ f \in L^2(\Omega) : \frac{\partial f}{\partial x_i} \text{ existiert in } L^2(\Omega) \text{ für } i = 1, ..., d \right\},$$
$$H^2(\Omega) := \left\{ f \in H^1(\Omega) : \frac{\partial^2 f}{\partial x_j x_i} := \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i} \text{ existiert in } L^2(\Omega) \text{ für } i, j = 1, ..., d \right\}.$$

In der Literatur werden die hier durch partielle Integration definierten Räume $H^m(\Omega)$ meist $W_2^m(\Omega)$ genannt. Die Sobolevräume $H^m(\Omega)$ und die Ableitungen in ihnen sind mittels Fouriertransformation definiert. Für beliebiges offenes Ω ist der Raum $W_2^m(\Omega)$ im Allgemeinen echt größer als $H^m(\Omega)$. Unter der hier gegebenen Voraussetzung, dass Ω beschränkt ist und der gleichmäßigen Kegelbedingung genügt, folgt aus dem Satz von Calderon-Zygmund jedoch die Gleichheit beider Räume (siehe [Wlo82]).

Die angegebenen Sätze für $H^1(\Omega)$ gelten teilweise nur, wenn Ω diesen Forderungen genügt. Deshalb werden in der gesamten Arbeit stets die genannten Bedingungen vorausgesetzt und mit den hier gegebenen Definitionen gearbeitet, ohne dass darauf explizit hingewiesen wird.

Satz 2.5.1

(1) $H^1(\Omega)$ *und* $H^2(\Omega)$ *sind separable Hilberträume, versehen mit den Skalarprodukten*

$$(f,g)_1 := (f,g)_{H^1(\Omega)} := (f,g) + (\nabla f, \nabla g) = \int_\Omega f(x)g(x)dx + \int_\Omega \nabla f(x) \cdot \nabla g(x)dx,$$

$$(f,g)_2 := (f,g)_{H^2(\Omega)} := (f,g)_1 + \sum_{i,j\in\{1,\dots,d\}} \int_\Omega \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \frac{\partial^2 g(x)}{\partial x_i \partial x_j} dx$$

und den induzierten Normen $\|f\|_{H^m(\Omega)} := (f,f)_m^{\frac{1}{2}}$, $m \in \{1,2\}$.

(2) *(Rellich-Kondrachov) Die Einbettung*

$$H^1(\Omega) \hookrightarrow L^2(\Omega)$$

durch die Identität ist kompakt.

(3) *Es gilt* $\mathcal{D}(\Omega) \subset H^1(\Omega)$ *und* $\mathcal{D}(\Omega)$ *liegt dicht in* $L^2(\Omega)$.

(4) *(Sobolev) Es gelte* $m > \frac{d}{2}$*, dann sind die Elemente aus* $H^m(\Omega)$, $m \in \{1,2\}$*, stetig und beschränkt auf* $\overline{\Omega}$ *und die Einbettung*

$$H^m(\Omega) \hookrightarrow C(\overline{\Omega})$$

ist stetig.

(5) $H^1(\Omega)$ *liegt dicht in* $L^2(\Omega)$.

BEWEIS. (1), (2), (3), (4) siehe [Ada75] oder [Wlo82].
(5) folgt direkt aus (3). ■

Insbesondere sind für den in der Arbeit betrachteten Fall $\Omega \subset \mathbb{R}^d$ mit $d \in \{1,2,3\}$ die Elemente von $H^2(\Omega)$ stetig.

Die verwendeten Bezeichnungen für die benutzten Normen und Skalarprodukte sind in der Notationstabelle (Anhang A) am Ende der Arbeit zu finden. Besondere Bedeutung kommt auch der L^2-Norm $\|\cdot\|_0 = (\cdot,\cdot)^{1/2}$ und der H^1-Halbnorm $|\cdot|_1 = (\nabla\cdot, \nabla\cdot)^{1/2}$ zu. Zunächst gilt dabei für die induzierte Norm auf $H^1(\Omega)$

$$\|f\|_1 := \|f\|_0 + |f|_1 = (f,f)^{\frac{1}{2}} + (\nabla f, \nabla f)^{\frac{1}{2}} \neq (f,f)_1^{\frac{1}{2}}.$$

Da für $a,b \geq 0$ aber $(a+b)^2 \leq 3(a^2+b^2)$ und $a^2+b^2 \leq (a+b)^2$ gilt, ist $\|\cdot\|_1$ mit von Ω unabhängigen Konstanten äquivalent zur induzierten Norm. Im Folgenden wird im Allgemeinen mit der Norm $\|\cdot\|_1$ auf $H^1(\Omega)$ gearbeitet.

Satz 2.5.2 *(Poincaré-Ungleichung) Sei* $\Omega \subset \mathbb{R}^d$ *offen, beschränkt und genüge der gleichmäßigen Kegelbedingung, dann existiert eine nur von* Ω *abhängige Konstante* $C > 0$*, so dass für alle* $f \in H^1(\Omega)$ *gilt:*

$$\|f\|^2_{H^1(\Omega)} \leq C \left[|f|_1^2 + \left(\int_\Omega f(x)dx \right)^2 \right].$$

BEWEIS. Siehe [Wlo82]. ■

Korollar 2.5.1 *Unter den Bedingungen von Satz 2.5.2 existiert eine Konstante $C_P > 0$, so dass $\|f\|_0 \leq C_P(|(f,1)| + |f|_1)$ für jedes $f \in H^1(\Omega)$ gilt.*

BEWEIS. Wende Satz 2.5.2 und die Beziehung $a^2 + b^2 \leq (a+b)^2$ für $a, b \geq 0$ an. ■

Bemerkung 2.5.1 *Es sei $f \in \mathcal{F} := \{\eta \in (H^1(\Omega))' : \langle \eta, 1\rangle = 0\}$ gegeben, dann existiert genau ein $u \in H^1(\Omega)$ mit $(u,1) = 0$ und*

$$(\nabla u, \nabla \eta) = \langle f, \eta\rangle \qquad \forall \eta \in H^1(\Omega).$$

BEWEIS. Wegen $\nabla 1 = 0$ und $\langle f, 1\rangle = 0$ genügt es, die Variationsgleichung mit allen η aus $H^1_{(0)}(\Omega) := \{\eta \in H^1(\Omega) : (\eta, 1) = 0\}$ zu testen. Dieser Unterraum von $H^1(\Omega)$ ist mit dem Skalarprodukt $(\cdot,\cdot)_1$ ein Hilbertraum. Für $\eta, \mu \in H^1_{(0)}(\Omega)$ gilt offensichtlich

$$(\nabla \eta, \nabla \mu) \leq |\eta|_1 |\mu|_1 \leq \|\eta\|_1 \|\mu\|_1.$$

Nach der Poincaré-Ungleichung existiert ferner eine Konstante $C > 0$ mit

$$\frac{1}{C}\|\eta\|_1^2 < |\eta|_1^2 + |(\eta,1)|^2 = (\nabla\eta, \nabla\eta).$$

Des Weiteren ist f auch auf $H^1_{(0)}(\Omega)$ stetig. Damit existiert nach dem Darstellungssatz von Frechét-Riesz ein $R^{-1}f \in H^1_{(0)}(\Omega)$ und nach dem Satz von Lax-Milgram (siehe [Wlo82]) ein stetiger, linearer, bijektiver Operator $A : H^1_{(0)}(\Omega) \to H^1_{(0)}(\Omega)$ mit $\|A^{-1}\| < C$ und es gilt

$$\begin{aligned} (\nabla\mu, \nabla\eta) &= (A\mu, \eta)_1 && \forall \mu, \eta \in H^1_{(0)}(\Omega), \\ \langle f, \eta\rangle &= (R^{-1}f, \eta)_1 && \forall \eta \in H^1_{(0)}(\Omega). \end{aligned}$$

Damit sind die Lösungen des Variationsproblems beim Testen mit $\eta \in H^1_{(0)}(\Omega)$ genau von der Form $u = A^{-1}R^{-1}f \in H^1_{(0)}(\Omega)$. Da A^{-1} und R^{-1} linear, bijektiv und stetig sind, existiert genau ein solches u, das zudem linear und stetig von f abhängt. ■

Damit ist der folgende Operator wohldefiniert, linear und stetig:

Definition 2.5.2 Definiere auf $\mathcal{F} := \{\eta \in (H^1(\Omega))' : \langle \eta, 1\rangle = 0\}$ den Green-Operator mit Neumann-Randbedingungen als $\mathcal{G} : \mathcal{F} \to H^1(\Omega)$, wobei $f \mapsto \mathcal{G}f$ erklärt ist durch

$$\begin{aligned} (\nabla \mathcal{G}f, \nabla\eta) &= \langle f, \eta\rangle && \forall \eta \in H^1(\Omega), \\ (\mathcal{G}f, 1) &= 0. \end{aligned}$$

Bemerkung 2.5.2

(1) *Für $f \in \mathcal{F}$ gilt $\|f\|_{(H^1(\Omega))'} \leq |\mathcal{G}f|_1 = \langle f, \mathcal{G}f\rangle^{\frac{1}{2}}$.*

(2) *Für $f \in \mathcal{F}$ gilt $|\mathcal{G}f|_1 \leq C\|f\|_{(H^1(\Omega))'}$ mit nur von Ω abhängigem $C > 0$.*

(3) *Für $f \in \mathcal{F} \cap L^2(\Omega)$ gilt $|\mathcal{G}f|_1 \leq C_P\|f\|_0$.*

(4) *Für $f \in (H^1(\Omega))'$ und $g \in H^1(\Omega)$ mit $\langle f, 1\rangle = 0 = (g, 1)$ gilt $(\mathcal{G}f, g) = \langle f, \mathcal{G}g\rangle$.*

BEWEIS. (1) Es gilt nach Definition von $\mathcal{G}$

$$\begin{aligned}\|f\|_{(H^1(\Omega))'} &= \sup_{\eta\in H^1(\Omega)} \frac{|\langle f,\eta\rangle|}{\|\eta\|_{H^1(\Omega)}} = \sup_{\substack{\eta\in H^1(\Omega)\\ \eta\neq\text{const}}} \frac{|(\nabla\mathcal{G}f,\nabla\eta)|}{\|\eta\|_{H^1(\Omega)}} \\ &\leq \sup_{\substack{\eta\in H^1(\Omega)\\ \eta\neq\text{const}}} \frac{|\mathcal{G}f|_1|\eta|_1}{\|\eta\|_{H^1(\Omega)}} \leq \sup_{\substack{\eta\in H^1(\Omega)\\ \eta\neq\text{const}}} \frac{|\mathcal{G}f|_1|\eta|_1}{|\eta|_1} = |\mathcal{G}f|_1 = \langle f,\mathcal{G}f\rangle^{\frac{1}{2}}.\end{aligned}$$

(2) Es gilt nach der Poincaré-Ungleichung

$$\|f\|_{(H^1(\Omega))'} = \sup_{\eta\in H^1(\Omega)} \frac{|\langle f,\eta\rangle|}{\|\eta\|_{H^1(\Omega)}} \geq \frac{|\langle f,\mathcal{G}f\rangle|}{\|\mathcal{G}f\|_{H^1(\Omega)}} \geq \frac{|\mathcal{G}f|_1^2}{C|\mathcal{G}f|_1} = \frac{1}{C}|\mathcal{G}f|_1.$$

(3) Aus $f \in \mathcal{F} \cap L^2(\Omega)$ folgt mit der Poincaré-Ungleichung

$$|\mathcal{G}f|_1^2 = \langle f,\mathcal{G}f\rangle = (f,\mathcal{G}f) \leq \|f\|_0\|\mathcal{G}f\|_0 \leq \|f\|_0 C_P|\mathcal{G}f|_1.$$

(4) Es gilt $(\mathcal{G}f, g) = (\nabla\mathcal{G}f, \nabla\mathcal{G}g) = \langle f, \mathcal{G}g\rangle$. ■

Dabei wird $f \in (H^1(\Omega))'$ als Element von $L^2(\Omega)$ aufgefasst, wenn es eine reguläre Distribution ist. Dass $|\mathcal{G}\cdot|_1$ die Eigenschaften einer Halbnorm auf $\mathcal{F}$ besitzt, ist klar. Teil (1) und (2) der Bemerkung zeigen, dass $|\mathcal{G}\cdot|_1$ auf $\mathcal{F}$ sogar eine zu $\|\cdot\|_{(H^1(\Omega))'}$ äquivalente Norm ist. Im Allgemeinen wird auf $\mathcal{F}$ diese Norm statt der kanonisch gegebenen verwenden.

Definition 2.5.3 Für $f \in \mathcal{F} = \{f \in (H^1(\Omega))' : \langle f, 1\rangle = 0\}$ definiere

$$\|f\|_{-1} := |\mathcal{G}f|_1 = \langle f,\mathcal{G}f\rangle^{\frac{1}{2}}.$$

2.6 Funktionenräume für Evolutionsgleichungen

Nun werden Sobolevräume eingeführt, die auch die Zeit mit einbeziehen. Es ist wesentlich, dass dabei die Zeit nicht einfach als eine weitere Raumdimension betrachtet, sondern separat behandelt wird. Eine detailliertere Einführung des Begriffes der Bochner-Meßbarkeit, des Bochner-Integrals und vektorwertiger Distributionen ist zum Beispiel bei [Wlo82] zu finden. Eine Darstellung der vektorwertigen L^p-Räume inklusive der hier nicht gegebenen Beweise ist bei [DL92] und [Zei90] zu finden. In diesem Abschnitt werden nur die wesentlichen, benötigten Definitionen und Sätze zitieren.

Definition 2.6.1 Sei H ein Hilbertraum und $1 \leq p < \infty$, definiere dann die Räume

$$L^p(0,T;H) := \{f : (0;T) \to H : f \text{ ist bochner-meßbar, } \|f\|_{L^p(0,T;H)} < \infty\},$$
$$L^\infty(0,T;H) := \{f : (0;T) \to H : f \text{ ist bochner-meßbar, } \|f\|_{L^\infty(0,T;H)} < \infty\}$$

mit den Normen

$$\|f\|_{L^p(0,T;H)} := \left(\int_0^T \|f(t)\|_H^p dt \right)^{\frac{1}{p}},$$
$$\|f\|_{L^\infty(0,T;H)} := \inf\{M : \|f\|_H \leq M \text{ fast überall}\}.$$

Diese Räume besitzen unter anderem folgende, den normalen $L^p(\Omega)$-Räumen ähnliche, Eigenschaften.

Satz 2.6.1 *Es sei H ein Hilbertraum.*

(1) *Für $1 \leq p \leq \infty$ ist $L^p(0,T;H)$ ein Banachraum.*

(2) *Ist H separabel, so ist $L^p(0,T;H)$ für $1 < p < \infty$ reflexiv und sein Dualraum ist $L^q(0,T;H')$, wobei $\frac{1}{p} + \frac{1}{q} = 1$ gilt.*

(3) *$L^\infty(0,T;H')$ ist der Dualraum von $L^1(0,T;H)$.*

(4) *Für $1 \leq p \leq \infty$ ist $L^p(0,T;H)$ stetig in $L^1(0,T;H)$ eingebettet.*

(5) *Der separable Hilbertraum $V \hookrightarrow H$ sei stetig in H eingebettet. Für $1 \leq p \leq \infty$ gilt dann $L^p(0,T;V) \hookrightarrow L^p(0,T;H)$.*

(6) *$L^2(0,T;H)$ ist ein Hilbertraum mit dem Skalarprodukt*

$$(u,v)_{L^2(0,T;H)} := \int_0^T (u(t), v(t))_H dt.$$

(7) *Ist V ein separabler reflexiver Banachraum, so ist $L^1(0,T;V)$ separabel.*

BEWEIS. Siehe [Zei90]. ■

Da $L^\infty(0,T;H)$ nicht reflexiv ist, wird hier mit der schwach-* Konvergenz gearbeitet. Zunächst ist jedoch unklar, was schwach-* Konvergenz in diesem Raum bedeutet, da unbekannt ist, ob $L^\infty(0,T;H)$ Dualraum eines normierten Raumes ist. Deshalb wird die schwach-* Konvergenz in $L^\infty(0,T;V') = (L^1(0,T;V))'$ untersucht, die durch die Dualtitätspaarung gegeben ist als:

Bemerkung 2.6.1 *Sei V ein separabler, reflexiver Banachraum. Eine Folge (x_n') in $L^\infty(0,T;V')$ konvergiert dann schwach-* gegen $x' \in L^\infty(0,T;V')$, wenn gilt*

$$\int_0^T \langle x_n'(t), y(t) \rangle \, dt \to \int_0^T \langle x'(t), y(t) \rangle \, dt \qquad \forall y \in L^1(0,T;V).$$

■

Zu $x \in H$ existiert nach dem Satz von Frechét-Riesz genau ein $x' \in H'$ mit $\langle x', \cdot \rangle = (x, \cdot)_H$ und $\|x'\|_{H'} = \|x\|_H$ auf H. Wird für $x \in L^\infty(0,T;H)$ nun $x'(t)$ mit $x(t)$ für alle $t \in (0;T)$ identifiziert, dann gilt offensichtlich auch $x' \in L^\infty(0,T;H')$ und $\|x\|_{L^\infty(0,T;H)} = \|x'\|_{L^\infty(0,T;H')}$. Damit lässt sich für einen separablen Hilbertraum H auch schwach-* Konvergenz auf $L^\infty(0,T;H)$ erklären. Die punktweisen dualen Paarungen werden hierbei durch die punktweisen Skalarprodukte ersetzt.

Definition 2.6.2 Eine Folge (x_n) in $L^\infty(0,T;H)$ konvergiert schwach-* gegen $x \in L^\infty(0,T;H)$, wenn gilt

$$\int_0^T (x_n(t), y(t))_H dt \to \int_0^T (x(t), y(t))_H dt \qquad \forall y \in L^1(0,T;H).$$

Bemerkung 2.6.2 *Offensichtlich gilt $x_n \rightharpoonup^* x$ in $L^\infty(0,T;H)$ genau dann, wenn $x_n' \rightharpoonup^* x'$ in $L^\infty(0,T;H')$ gilt.* ■

Wegen der Normgleichheit von x und x' lassen sich die Resultate für schwach-* Konvergenz übertragen. Dies liefert insbesondere folgenden Satz:

Satz 2.6.2 *Sei H ein separabler Hilbertraum und $(x_n) \in L^\infty(0,T;H)$ eine beschränkte Folge, dann besitzt (x_n) eine schwach-* konvergente Teilfolge, die durch die gleiche Konstante beschränkt ist.*

BEWEIS. Wende Satz 2.1.5 und Satz 2.6.1 (7) an. ■

Bemerkung 2.6.3 *Aus $x_n \rightharpoonup^* x$ in $L^\infty(0,T;H)$ folgt $x_n \rightharpoonup x$ in $L^2(0,T;H)$.*

BEWEIS. Es ist klar, dass $x_n, x \in L^2(0,T;H)$ gilt. Ferner folgt aus $y \in L^2(0,T;H)$ auch $y \in L^1(0,T;H)$. Damit folgt aus Definition 2.6.2

$$(x_n, y)_{L^2(0,T;H)} \to (x, y)_{L^2(0,T;H)} \qquad \forall y \in L^2(0,T;H).$$

■

Ähnlich wie für die Ortsableitungen wird auch für Zeitableitungen ein verallgemeinertes Konzept eingeführt.

Definition 2.6.3 Definiere für einen Banachraum X:

(1) Der Raum der vektorwertigen Testfunktionen ist

$$\mathcal{D}((0;T),X) := \{f : \mathcal{D}((0;T)) \to X : f \text{ ist linear und stetig}\}.$$

Er ist versehen mit einer auf Ableitungen basierenden Topologie (siehe etwa [Wlo82]). Sein Dualraum $(\mathcal{D}((0;T),X))'$ bezüglich dieser Topologie ist der Raum der vektorwertigen Distributionen. Existiert zu $f \in (\mathcal{D}((0;T),X))'$ ein $\tilde{f} \in L^2(0,T;X)$, so dass $(\tilde{f}, \phi)_{L^2(0,T;X)} = f(\phi)$ für alle $\phi \in \mathcal{D}((0;T),X)$ gilt, so betrachten wir f als Element von $L^2(0,T;X)$ und schreiben $f = \tilde{f} \in L^2(0,T;X)$.

(2) Die verallgemeinerte Zeitableitung von $f \in (\mathcal{D}((0;T),X))'$ ist die vektorwertige Distribution

$$\frac{df}{dt} : \mathcal{D}((0;T)) \to X, \qquad \phi \mapsto -f(\phi'),$$

wobei ϕ' die klassische Ableitung von ϕ nach t ist.

Um die Zeitableitung von den Ortsableitungen zu unterscheiden, wird für die Zeitableitung die Schreibweise $\frac{d}{dt}$ statt $\frac{\partial}{\partial t}$ verwendet.

Definition 2.6.4 Es sei $V \hookrightarrow H \hookrightarrow V'$ ein Gelfand'scher Dreier mit separablen Hilberträumen V und H. Definiere nun mit der verallgemeinerten Zeitableitung $\frac{d}{dt}$ den Raum

$$W(0,T;V) := \left\{ f \in L^2(0,T;V) : \frac{df}{dt} \in L^2(0,T;V') \right\}$$

mit dem Skalarprodukt

$$(u,v)_{W(0,T;V)} = \int_0^T (u(t),v(t))_V \, dt + \int_0^T (\frac{du(t)}{dt}, \frac{dv(t)}{dt})_{V'} dt.$$

Satz 2.6.3

(1) $W(0,T;V)$ *ist ein Hilbertraum.*

(2) *Für* $u, v \in W(0,T;V)$ *gilt*

$$\int_0^T \left\langle \frac{du(t)}{dt}, v(t) \right\rangle dt + \int_0^T \left\langle \frac{dv(t)}{dt}, u(t) \right\rangle dt = (u(t),v(t))_H \Big|_0^T \, dt.$$

(3) *Für* $u \in W(0,T;V)$ *und* $v \in V$ *gilt*

$$\left\langle \frac{du}{dt}(\cdot), v \right\rangle = \frac{d}{dt}(u(\cdot), v) \qquad \text{in } (\mathcal{D}((0;T)))'.$$

(4) *Für* $u \in L^2(0,T;V)$ *ist* $w \in L^2(0,T;V')$ *genau dann die verallgemeinerte Zeitableitung, wenn für alle* $v \in V$ *und alle* $\phi \in \mathcal{D}((0;T))$ *gilt*

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

(5) *Die Identität ist eine stetige Einbettung*

$$W(0,T;V) \hookrightarrow L^2(0,T;V) \hookrightarrow L^2(0,T;H).$$

(6) *Ist die Einbettung $V \hookrightarrow H$ kompakt, so ist die Identität eine kompakte Einbettung*

$$W(0,T;V) \hookrightarrow L^2(0,T;H).$$

(7) *Ist $u \in C^1([0;T];H)$, so ist die verallgemeinerte Zeitableitung die klassische Ableitung von u.*

(8) *Die Identität ist eine stetige Einbettung*

$$W(0,T;V) \hookrightarrow C([0;T];H),$$

es existiert also zu jedem $u \in W(0,T;V)$ ein in der Zeit stetiger Repräsentant.

BEWEIS. (1) siehe [Wlo82].
(2), (3) siehe [DL92].
(4), (7), (8) siehe [Zei90].
(5) ist klar.
(6) siehe [Lio69]. ■

Teil (8) bedeutet insbesondere, dass sinnvoll von $u(t)$ gesprochen werden kann und für Evolutionsprobleme Anfangswerte $u(0) = u_0$ vorgegeben werden können. Aus (6) folgt, dass im Fall einer kompakten Einbettung $V \hookrightarrow H$ aus einer in $W(0,T;V)$ beschränkten Folge eine in $L^2(0,T;H)$ konvergente Teilfolge ausgewählt werden kann. Mit der Charakterisierung in Teil (4) werden nun einige Rechenregeln für verallgemeinerte Zeitableitungen gezeigt.

Korollar 2.6.1 *Für $u \in W(0,T;V)$ und $\psi \in \mathcal{D}((0;T))$ gilt*

$$\frac{du}{dt}\psi + u\psi' = \frac{d}{dt}(u\psi).$$

BEWEIS. Sei $v \in H$ und $\phi \in \mathcal{D}((0;T))$, dann gilt $-(\psi\phi)' + \psi'\phi = -\psi\phi'$ nach der Produktregel für klassische Ableitungen in $\mathcal{D}((0;T))$. Zusammen mit Satz 2.6.3 (4) folgt daraus

$$\begin{aligned}
&\int_0^T \left\langle \frac{du(t)}{dt}\psi(t), v \right\rangle \phi(t)dt + \int_0^T \left(u(t)\psi'(t), v\right)_H \phi(t)dt \\
&\quad = -\int_0^T (u(t), v)_H \, (\psi\phi)'(t)dt + \int_0^T (u(t), v)_H \, \psi'(t)\phi(t)dt \\
&\quad = -\int_0^T (u(t), v)_H \, \psi(t)\phi'(t)dt.
\end{aligned}$$

■

Korollar 2.6.2 *Für $E \in V$ und $u, v \in W(0,T;V)$ gilt*

$$\left\langle \frac{du}{dt}, v \right\rangle E + \left\langle \frac{dv}{dt}, u \right\rangle E = \frac{d}{dt}\left((u,v)_H E\right).$$

BEWEIS. Für $\phi \in \mathcal{D}((0;T))$ und $z \in V$ beliebig gilt

$$\begin{aligned}
&\int_0^T \left(\left\langle \frac{du(t)}{dt}, v(t) \right\rangle E, z \right)_H \phi(t)dt + \int_0^T \left(\left\langle \frac{dv(t)}{dt}, u(t) \right\rangle E, z \right)_H \phi(t)dt \\
&\underset{\text{(a)}}{=} (E,z)_H \left[\int_0^T \left\langle \frac{du(t)}{dt}\phi(t), v(t) \right\rangle dt + \int_0^T \left\langle \frac{dv(t)}{dt}\phi(t), u(t) \right\rangle dt \right] \\
&\underset{\text{(b)}}{=} (E,z)_H \left[\int_0^T \left\langle \frac{d(u\phi)(t)}{dt}, v(t) \right\rangle dt - \int_0^T (u(t),v(t))_H \phi'(t)dt + \int_0^T \left\langle \frac{dv(t)}{dt}, (u\phi)(t) \right\rangle dt \right] \\
&\underset{\text{(c)}}{=} (E,z)_H \left[\underbrace{(u\phi, v)_H \big|_0^T}_{=0} - \int_0^T (u(t),v(t))_H \phi'(t)dt \right] \\
&\underset{\text{(d)}}{=} -\int_0^T \left((u(t),v(t))_H E, z\right)_H \phi'(t)dt.
\end{aligned}$$

Dabei erhält man (a) durch Ausklammern, (b) nach Korollar 2.6.1, (c) nach Satz 2.6.3 (2) und (d) da $\phi(0) = \phi(T) = 0$ gilt. ■

Im Fall $V = H^1(\Omega)$ und $E \equiv 1$ liefert dies eine Produktregel für verallgemeinerte Zeitableitungen. Sie wird in Verbindung mit folgender Aussage benötigt:

Bemerkung 2.6.4 *Es gilt für $u \in W(0,T;H^1(\Omega))$ mit $(u(t),1) = 0$ fast überall in $(0,T)$:*

$$\mathcal{G}\frac{du}{dt} = \frac{d}{dt}\mathcal{G}u.$$

BEWEIS. Nach Satz 2.6.3 (3) gilt $\left\langle \frac{du}{dt}, 1 \right\rangle = 0$. Es sei nun $v \in H^1(\Omega)$ und $\phi \in \mathcal{D}((0;T))$ gegeben. Wegen $(\mathcal{G}f, 1) = 0$ betrachten wir ohne Beschränkung der Allgemeinheit nur den Fall $(v,1) = 0$. Durch Anwendung von Bemerkung 2.5.2 erhalten wir

$$\begin{aligned}
\int_0^T \left(\mathcal{G}\frac{du(t)}{dt}, v\right)\phi(t)dt &= \int_0^T \left\langle \frac{du(t)}{dt}, \mathcal{G}v \right\rangle \phi(t)dt = -\int_0^T (u(t), \mathcal{G}v)\,\phi'(t)dt \\
&= -\int_0^T (\mathcal{G}u(t), v)\,\phi'(t)dt = \int_0^T \left\langle \frac{d}{dt}(\mathcal{G}u(t)), v \right\rangle \phi(t)dt.
\end{aligned}$$

■

Korollar 2.6.3 *Es gilt für* $u \in W(0,T;H^1(\Omega))$ *mit* $(u(t),1)=0$ *fast überall in* $(0;T)$*:*

$$2\left\langle \frac{du}{dt}, \mathcal{G}u \right\rangle = \frac{d}{dt}\,(u, \mathcal{G}u) = \frac{d}{dt}|\mathcal{G}u|_1^2.$$

BEWEIS. Nach Bemerkung 2.6.4 gilt $\left\langle \frac{du}{dt}, \mathcal{G}u \right\rangle = \left\langle \frac{d}{dt}\mathcal{G}u, u \right\rangle$. Benutze nun Korollar 2.6.2 mit u und $\mathcal{G}u$ für $E \equiv 1$. ■

Wie beim Sobolevraum $H^1(\Omega)$ wird auch hier nicht die vom Skalarprodukt induzierte Norm betrachtet, sondern die dazu äquivalente Norm, die man als Summe der induzierten Normen der Skalarprodukte erhält, also

$$\|u\|_{W(0,T;V)} = \|u\|_{L^2(0,T;V)} + \left\| \frac{du}{dt} \right\|_{L^2(0,T;V')}.$$

Der Übersichtlichkeit halber werden die Bezeichnungen einiger Normen und Skalarprodukte nach dem folgenden Schema abgekürzt. Für den Raum $X(0,T;H(\Omega))$ schreibe

$$\begin{aligned}
(u,v)_{XH} &:= (u,v)_{X(0,T;H(\Omega))},\\
\|u\|_{XH} &:= \|u\|_{X(0,T;H(\Omega))},\\
(u,v)_{L^2\nabla} &:= \int_0^T (\nabla u(t), \nabla v(t))dt.
\end{aligned}$$

3 Die Cahn-Hilliard Gleichung

In diesem Kapitel wird die Cahn-Hilliard Gleichung analytisch untersucht. Grundlage der Betrachtungen ist eine Variationsformulierung der Gleichung in geeigneten Sobolevräumen. Wie das vorherige Kapitel zeigt, liegt diesen ein allgemeineres Verständnis des Ableitungsbegriffes zugrunde. Dadurch besitzen sie in verschiedener Hinsicht eine besondere Struktur. Insbesondere die Eigenschaften als separable Hilberträume sind für die Untersuchung der Gleichung wichtig. Um einen Einblick in die Struktur der Gleichung zu gewinnen, werden diese Eigenschaften zur Herleitung von Existenz- und Eindeutigkeitsresultaten ausgenutzt. Viele Methoden und Vorgehensweisen werden später in ähnlicher Form auf eine diskrete Cahn-Hilliard Gleichung angewandt, weshalb die Untersuchung der kontinuierlichen Gleichung auch vor diesem Hintergrund interessant ist.

3.1 Die Cahn-Hilliard Gleichung als Variationsproblem

Mit den definierten Räumen lässt sich die Cahn-Hilliard Gleichung als exakt formuliertes Variationsproblem schreiben durch:

(P) Finde zu $u_0 \in K$ mit $|(u_0,1)| < |\Omega|$ ein $u \in W(0,T;H^1(\Omega)) \cap L^\infty(0,T;H^1(\Omega))$ mit $u \in K$ fast überall in $(0;T)$ und ein $w \in L^2(0,T;H^1(\Omega))$ mit

$$\left\langle \frac{du}{dt}, \eta \right\rangle + (\nabla w, \nabla \eta) = 0 \qquad \forall \eta \in H^1(\Omega), \text{ f.ü. in } (0;T) \tag{3.1.1}$$

$$\gamma(\nabla u, \nabla \eta - \nabla u) - (u, \eta - u) \geq (w, \eta - u) \qquad \forall \eta \in K, \text{ f.ü. in } (0;T) \tag{3.1.2}$$

$$u(0) = u_0.$$

Für diese Formulierung wird im Folgenden die Existenz und Eindeutigkeit von Lösungen gezeigt. Außerdem wird eine weitere Formulierung untersucht, die sich als äquivalent herausstellen wird.

(Q) Finde zu $u_0 \in K$ mit $|(u_0,1)| < |\Omega|$ ein $u \in W(0,T;H^1(\Omega)) \cap L^\infty(0,T;H^1(\Omega))$, $u \in K_m$ f.ü. $(0;T)$ mit

$$\gamma(\nabla u, \nabla \eta - \nabla u) + \left(\mathcal{G}\frac{du}{dt}, \eta - u \right) \geq (u, \eta - u) \qquad \forall \eta \in K_m, \text{ f.ü. in } (0;T)$$

$$u(0) = u_0.$$

Dabei ist die konvexe Menge $K \subset H^1(\Omega)$ gegeben als

$$K := \{\eta \in H^1(\Omega) : |\eta(x)| \leq 1\}.$$

Da die Massenerhaltung eine wesentliche Rolle bei der Betrachtung der Cahn-Hilliard Gleichung spielt, kommt $m := (u_0, 1)$ besondere Bedeutung zu. In diesem und allen folgenden

Kapiteln wird m als diese Zahl verwendet. Nun lässt sich die ebenfalls konvexe Teilmenge von K, in der Massenerhaltung gilt, definieren.

$$K_m = \{\eta \in K : (\eta, 1) = (u_0, 1) = m\}.$$

3.2 Existenz und Eindeutigkeit von Lösungen

3.2.1 Beweisskizze

Um die Existenz von Lösungen für Evolutionsprobleme zu zeigen, wird oft die Galerkin-Methode verwendet. Für lineare Probleme ist sie zum Beispiel bei [DL92], [Wlo82] und [Zei90] ausführlich dargestellt. Eine Version für nichtlineare Evolutionsgleichungen ist bei [Lio69] zu finden. Für einfache parabolische Variationsungleichungen ist bei [Lio69] ferner folgende Methode beschrieben: Zunächst wird die Variationsungleichung durch eine Folge nichtlinearer Evolutionsgleichungen approximiert. Für diese wird mittels der Galerkin-Methode die Existenz und Eindeutigkeit von Lösungen gezeigt. Dann wird ein Grenzübergang zur Variationsungleichung vorgenommen.

Blowey und Elliot geben in [BE91] einen Existenzbeweis, der diese Ideen für den komplexeren Fall der Cahn-Hilliard Gleichung mit Hindernispotential adaptiert. Im Folgenden wird dieser Beweis detailliert ausgearbeitet. Er lässt sich grob in folgende Schritte gliedern:

(1) Die Approximation der Variationsungleichung in (P) durch Variationsgleichungen führt zu einer Folge regularisierter Probleme (P_ϵ). Funktionswerte $u_\epsilon(t,x) \notin [-1;1]$ werden dabei mit für $\epsilon \to 0$ größer werdender Energie „bestraft“.

(2) Konstruktion ortsdiskreter Galerkin-Approximationen (P_ϵ^k) von (P_ϵ).

(3) Nachweis der Existenz von Lösungen u_ϵ^k der gewöhnlichen Differentialgleichungen in (P_ϵ^k).

(4) Herleitung von Abschätzungen für u_ϵ^k, die von ϵ und k unabhängig sind.

(5) Nachweis, dass der (schwache) Grenzwert u_ϵ einer Teilfolge von u_ϵ^k eindeutige Lösung von (P_ϵ) ist und ebenfalls von ϵ unabhängigen Abschätzungen genügt.

(6) Nachweis, dass der (schwache) Grenzwert u einer Teilfolge von u_ϵ Lösung von (P) ist.

(7) Nachweis der Eindeutigkeit der Lösung u.

3.2.2 Galerkin-Approximation von $H^1(\Omega)$

Es sei $\{z_j\}_{j=1}^\infty$ eine Orthogonalbasis von $H^1(\Omega)$, bestehend aus den Eigenfunktionen des Operators $z - \Delta z$, die der Randbedingung $\frac{\partial z}{\partial \nu} = 0$ genügen. Normiert man diese mit $(z_i, z_j) = \delta_{ij}$, so sind sie die Lösungen von

$$\begin{aligned} (\nabla z_i, \nabla v) + (z_i, v) &= \mu(z_i, v) && \forall v \in H^1(\Omega), \\ (z_i, z_j) &= \delta_{ij}. \end{aligned}$$

Die Existenz einer solchen Basis folgt aus der Anwendung von Spektralsätzen für kompakte Operatoren auf die kompakte Einbettung $H^1(\Omega) \hookrightarrow L^2(\Omega)$. Sie ergibt sich zum Beipiel direkt aus Korollar 19.20 im Kapitel über die Friedrichs-Erweiterung symmetrischer Operatoren in [Zei90]. Eine explizite Untersuchung solcher Eigenwertprobleme ist etwa in [BO91] zu finden.

Definition 3.2.1 Es sei $V^k := span\{z_j\}_{j=1}^k \subset H^1(\Omega)$ und P^k die Projektion

$$P^k : H^1(\Omega) \to V^k, v \mapsto P^k v = \sum_{j=1}^{k} (v, z_j) z_j.$$

Da (z_j) durch die Normierung eine Orthonormalbasis von $L^2(\Omega)$ ist, konvergiert für $v \in L^2(\Omega)$ die Folge der Projektionen von v gegen v, es gilt also $\|P^k v - v\|_0 \to 0$. Aus der Definition als Eigenwert folgt, dass $\eta \in V^k$ und $v \in H^1(\Omega)$ genau dann H^1-orthogonal sind, wenn sie L^2-orthogonal sind. Insbesondere werden im Beweis folgende Aussagen benötigt:

Lemma 3.2.1 *Für $v \in H^1(\Omega)$ und $\eta \in V^k$ gilt*

(1) $(P^k v - v, \eta) = 0$,

(2) $(\nabla(P^k v - v), \nabla\eta) = 0$,

(3) $\|P^k v\|_0 \leq \|v\|_0$,

(4) $|P^k v|_1 \leq |v|_1$.

BEWEIS. Es seien $v \in H^1(\Omega)$ und $\eta \in V^k$ beliebig.
(1) Dann gilt $\eta = P^k \eta$ und damit

$$(P^k v, \eta) = \sum_{j=1}^{k} (v, z_j)(z_j, \eta) = (v, P^k \eta) = (v, \eta).$$

(2) Ferner existiert zu jedem Eigenvektor z_j ein Eigenwert μ_j mit

$$(\nabla z_j, \nabla \underbrace{P^k v - v}_{\in H^1(\Omega)}) = (\mu_j - 1)(z_j, P^k v - v) = 0.$$

(3) Aus der Orthogonalität (1) folgt nun

$$\|P^k \eta\|_0^2 = (P^k \eta, P^k \eta) = (P^k \eta, \eta) \leq \|P^k \eta\|_0 \|\eta\|_0.$$

(4) Dies folgt analog zu (3) aus (2). ■

Offenkundig ist $z \equiv \text{const}$ ein Eigenvektor zum Eigenwert $\mu = 1$. In den weiteren Betrachtungen wird ohne Beschränkung der Allgemeinheit angenommen, dass $z \equiv 1$ in jedem V^k liegt (setze dazu etwa $z_1 :\equiv \text{const}$ geeignet). Damit gilt insbesondere $(P^k v, 1) = (v, 1)$.

3.2.3 Ein differenzierbares Energiefunktional

Zunächst wird eine Variationsgleichung als Regularisierung von (P) hergeleitet. Um das Problem (P), dessen Variationsungleichung vom nicht differenzierbaren Energiefunktional $\psi : \mathbb{R} \to \overline{\mathbb{R}}$ mit

$$\psi(r) = \frac{1}{2}(1-r^2) + I_{[-1;1]} = \begin{cases} \frac{1}{2}(1-r^2) & \text{für} \quad |r| \leq 1 \\ \infty & \text{für} \quad |r| > 1 \end{cases}$$

herrührt, zu einer Variationsgleichung zu regularisieren, wird ein differenzierbares Energiefunktional $\psi_\epsilon(r) = \frac{1}{2}(1-r^2) + \widehat{\psi}_\epsilon(r)$ konstruiert. Es approximiert ψ, indem es $|r| > 1$ durch hohe, aber noch endliche Energie „bestraft“:

$$\widehat{\psi}_\epsilon(r) := \begin{cases} \frac{1}{2\epsilon}(|r| - (1+\frac{\epsilon}{2}))^2 + \frac{\epsilon}{24} & \text{für} \quad |r| \geq 1+\epsilon \\ \frac{1}{6\epsilon^2}(|r|-1)^3 & \text{für} \quad 1 \leq |r| \leq 1+\epsilon \\ 0 & \text{für} \quad |r| \leq 1 \end{cases}$$

$$\beta_\epsilon(r) := \epsilon\widehat{\psi}'_\epsilon(r) = \epsilon(r + \psi'_\epsilon(r)).$$

Bemerkung 3.2.1 *Die Abbildungen ψ_ϵ, $\widehat{\psi}_\epsilon$ und β_ϵ besitzen folgende Eigenschaften:*

(1) $\psi_\epsilon(r) \xrightarrow[\epsilon \to 0]{} \psi(r)$

(2) $\beta_\epsilon(r) = \begin{cases} \left(|r| - \left(1+\frac{\epsilon}{2}\right)\right)\text{sign}(r) & \text{für} \quad |r| \geq 1+\epsilon \\ \frac{1}{2\epsilon}(|r|-1)^2\text{sign}(r) & \text{für} \quad 1 \leq |r| \leq 1+\epsilon \\ 0 & \text{für} \quad |r| \leq 1 \end{cases}$

(3) $\beta'_\epsilon(r) = \begin{cases} 1 & \text{für} \quad |r| \geq 1+\epsilon \\ \frac{1}{\epsilon}(|r|-1) & \text{für} \quad 1 \leq |r| \leq 1+\epsilon \\ 0 & \text{für} \quad |r| \leq 1 \end{cases} \qquad \leq 1$

(4) $\beta_\epsilon \in C^1(\mathbb{R})$, $\psi_\epsilon, \widehat{\psi}_\epsilon \in C^2(\mathbb{R})$ *und* β_ϵ *ist lipschitz-stetig mit* $0 \leq \beta'_\epsilon \leq 1$.

(5) β_ϵ *und* $\widehat{\psi}'_\epsilon = \frac{1}{\epsilon}\beta_\epsilon$ *sind monoton wachsend, also ist* $\widehat{\psi}_\epsilon$ *konvex und es gilt*

$$\widehat{\psi}_\epsilon(s) \leq \widehat{\psi}_\epsilon(r) + \frac{1}{\epsilon}\beta_\epsilon(s)(s-r).$$

Für $|r| \leq 1$ *gilt*

$$\psi_\epsilon(s) \leq \widehat{\psi}_\epsilon(r) + \frac{1}{\epsilon}(s-r)^2 + \frac{1}{2}(1-s^2).$$

(6) *Es gilt* $|\beta_\epsilon(r)| \leq r\beta_\epsilon(r)$ *und* $|\beta_\epsilon(r)| \leq |r|$.

BEWEIS. (1), (2) und (3) sind klar. (4) und (6) folgen aus Teil (2) und (3).
(5) Die Monotonie von β_ϵ folgt aus $0 \leq \beta_\epsilon$ in (4). Für differenzierbare konvexe Funktionen $f : \mathbb{R} \to \mathbb{R}$ gilt $f(s)+f'(s)(r-s) \leq f(r)$, woraus mit $f := \widehat{\psi}_\epsilon$ die erste Ungleichung folgt. Aus der Lipschitz-Stetigkeit von β_ϵ folgt für $|r| \leq 1$ wegen $\beta_\epsilon(r) = 0$, dass $\frac{1}{\epsilon}\beta_\epsilon(s)(s-r) \leq \frac{1}{\epsilon}(s-r)^2$ und somit die zweite Ungleichung gilt. ■

Um schließlich Lösungen von (P) zu erhalten wird ein Grenzübergang $\epsilon \to 0$ durchgeführt. Dafür wird nun noch der Zusammenhang von β_ϵ und seinem punktweisen Grenzwert β untersucht.

Definition 3.2.2 Für $r \in \mathbb{R}$ setze $\beta(r) := \lim_{\epsilon \to 0} \beta_\epsilon(r)$.

Bemerkung 3.2.2 *Die Funktion* β *hat die Eigenschaften*

(1) $\beta(r) = \begin{cases} (|r|-1)sign(r) & \text{für} \quad |r| > 1 \\ 0 & \text{für} \quad |r| \leq 1 \end{cases} = \begin{cases} r-1 & \text{für} \quad r > 1 \\ 0 & \text{für} \quad |r| \leq 1 \\ r+1 & \text{für} \quad r < -1 \end{cases}$,

(2) $|\beta(r) - \beta(s)| \leq |r-s| \qquad \forall r, s \in \mathbb{R}$,

(3) $|\beta(r) - \beta_\epsilon(r)| \leq \frac{\epsilon}{2} \qquad \forall r \in \mathbb{R}$.

BEWEIS. (1) und (2) sind klar. (3) ist für $|r| \leq 1$ auch klar. Für $|r| \geq 1+\epsilon$ gilt

$$|\beta(r) - \beta_\epsilon(r)| = \left|(|r|-1) - \left(|r| - \left(1 - \frac{\epsilon}{2}\right)\right)\right| = \frac{\epsilon}{2}.$$

Setze nun $f(x) := x - \frac{1}{2\epsilon}x^2$, so gilt für $1 \leq |r| \leq 1+\epsilon$ wegen $f'(x) = 1 - \frac{x}{\epsilon}$

$$|\beta(r) - \beta_\epsilon(r)| = \underbrace{f(|r|-1)}_{\geq 0} \leq \max_{1 \leq x \leq 1+\epsilon} f(|x|-1) = \max_{0 \leq x \leq \epsilon} f(x) = f(\epsilon) = \frac{\epsilon}{2}.$$

■

Korollar 3.2.1 *Für* $\epsilon < \frac{1}{4}$ *gilt* $\psi_\epsilon(r) \geq -C_0\epsilon$ *mit einem von* ϵ *unabhängigem* $C_0 > 0$.

BEWEIS. Da ψ_ϵ symmetrisch ist, suchen wir nur für den Fall $r \geq 0$ das Minimum von ψ_ϵ. Für $0 \leq r \leq 1$ gilt $\psi_\epsilon(r) \geq 0$. Für $1 \leq r \leq 1+\epsilon$ ist ψ_ϵ mit $\psi'_\epsilon(r) < \frac{1}{\epsilon}\beta_\epsilon(r) \leq -\frac{1}{2}$ monoton fallend. Da ferner $\psi_\epsilon(r) \xrightarrow[r \to \infty]{} \infty$ gilt, nimmt ψ_ϵ sein Minimum für $r \geq 1+\epsilon$ an. Dort gilt

$$\psi'_\epsilon(r) = \frac{1}{\epsilon}\left(r - \left(1 + \frac{\epsilon}{2}\right)\right) - r = \frac{1-\epsilon}{\epsilon}r - \frac{1+\frac{\epsilon}{2}}{\epsilon},$$

also nimmt ψ_ϵ sein Minimum an der Stelle $r_\epsilon = \frac{1+\frac{\epsilon}{2}}{1-\epsilon} > \frac{1-\epsilon^2}{1-\epsilon} = 1+\epsilon$ an. Nun schätzen wir $\psi_\epsilon(r_\epsilon)$ nach unten ab. Dazu setzen wir $f(\epsilon) = r_\epsilon^2$. Dann gilt für $0 < \epsilon < \frac{1}{4}$

$$f'(\epsilon) = \frac{3(1+\frac{1}{\epsilon})}{(1-\epsilon)^3} > 0, \qquad f''(\epsilon) = \frac{\frac{3}{2}(1-\epsilon) + 9(1+\frac{\epsilon}{2})}{(1-\epsilon)^4} > 0.$$

Also ist die Funktion f zwischen 0 und $\frac{1}{4}$ streng monoton steigend und strikt konvex und wird somit durch ihre Sekante mit einer positiven Steigung C dominiert. Wegen $\widehat{\psi}_\epsilon \geq 0$ und $f(0) = 1$ gilt damit

$$\begin{aligned}\psi_\epsilon(r_\epsilon) &> \frac{1}{2}(1 - r_\epsilon^2) = \frac{1}{2}(f(0) - f(\epsilon)) \\ &\geq \frac{1}{2}\left(f(0) - \left[f(0) + \epsilon\frac{f(\frac{1}{4}) - f(0)}{\frac{1}{4}}\right]\right) = -\frac{1}{2}C\epsilon.\end{aligned}$$

■

3.2.4 Regularisierung der Variationsungleichung

(Schritt 1) Nun wird das aus dem differenzierbaren Energiefunktional ψ_ϵ resultierende Problem als Variationsgleichung aufgestellt.

(P_ϵ) Für $\epsilon > 0$ finde zu $u_0 \in K$ mit $|(u_0, 1)| < |\Omega|$ ein $u_\epsilon \in W(0,T;H^1(\Omega)) \cap L^\infty(0,T;H^1(\Omega))$ und ein $w_\epsilon \in L^2(0,T;H^1(\Omega))$ mit

$$\left\langle \frac{\partial u_\epsilon}{\partial t}, \eta \right\rangle + (\nabla w_\epsilon, \nabla\eta) = 0 \qquad \forall \eta \in H^1(\Omega),\ \text{f.ü. in } (0;T), \tag{3.2.1}$$

$$\gamma(\nabla u_\epsilon, \nabla\eta) + (\psi_\epsilon'(u_\epsilon), \eta) = (w_\epsilon, \eta) \qquad \forall \eta \in H^1(\Omega),\ \text{f.ü. in } (0;T), \tag{3.2.2}$$

$$u_\epsilon(0) = u_0. \tag{3.2.3}$$

Bemerkung 3.2.3 *Offensichtlich ist (3.2.2) äquivalent zu den Gleichungen*

$$\gamma(\nabla u_\epsilon, \nabla\eta) - (u_\epsilon, \eta) + (\widehat{\psi}_\epsilon'(u_\epsilon), \eta) = (w_\epsilon, \eta) \qquad \forall \eta \in H^1(\Omega), \text{f.ü. in } (0;T), \tag{3.2.4}$$

$$\gamma(\nabla u_\epsilon, \nabla\eta) - (u_\epsilon, \eta) + \frac{1}{\epsilon}(\beta_\epsilon(u_\epsilon), \eta) = (w_\epsilon, \eta) \qquad \forall \eta \in H^1(\Omega), \text{f.ü. in } (0;T). \tag{3.2.5}$$

3.2.5 Lösungen des regularisierten Problems

Satz 3.2.1 *Für $0 < \epsilon < \frac{1}{4}$ hat das Problem (P_ϵ) eine Lösung u_ϵ, für die mit von ϵ und T unabhängigen Konstanten gilt:*

$$\begin{aligned}\|u_\epsilon\|_{W(0,T;H^1(\Omega))} &\leq C(1 + T^{\frac{1}{2}}), \\ \|u_\epsilon\|_{L^\infty(0,T;H^1(\Omega))} &\leq C, \\ \|w_\epsilon\|_{L^2(0,T;H^1(\Omega))} &\leq C(1 + T^{\frac{1}{2}}).\end{aligned}$$

(Schritt 2) Es werden zunächst ortsdiskrete gewöhnliche Differentialgleichungen (P_ϵ^k) als Approximation von (P_ϵ) konstruiert. Dazu wird die Galerkin-Approximation aus Definition 3.2.1 verwendet.

(P_ϵ^k) Für $\epsilon > 0$ und $k \in \mathbb{N}$ finde zu $u_0 \in K$ mit $|(u_0, 1)| < |\Omega|$ $u^k(t), w^k(t) \in V^k$, so dass für $t \in [0; T]$ gilt

$$u^k(t) = \sum_{j=1}^{k} U_j(t) z_j, \qquad w^k(t) = \sum_{j=1}^{k} W_j(t) z_j,$$

$$(\frac{du^k}{dt}, \eta) + (\nabla w^k, \nabla \eta) = 0 \qquad \forall \eta \in V^k \tag{3.2.6}$$

$$\gamma(\nabla u^k, \nabla \eta) + \frac{1}{\epsilon}(\beta_\epsilon(u^k), \eta) - (u^k, \eta) = (w^k, \eta) \qquad \forall \eta \in V^k \tag{3.2.7}$$

$$u^k(0) = P^k(u_0). \tag{3.2.8}$$

U und W sind die zeitabhängigen Koeffizienten der Basisvektoren und damit Abbildungen von $\mathbb{R}$ nach $\mathbb{R}^k$.

Lemma 3.2.2 *(P_ϵ^k) besitzt eine eindeutig bestimmte Lösung $U, W \in C^1([0; T]; \mathbb{R}^k)$, für die mit $(u^k(t), 1) = (u_0, 1)$ Massenerhaltung gegeben ist.*

BEWEIS.(Lemma 3.2.2) Im Folgenden seien $X, Y \in \mathbb{R}^k$ beliebig. Setzen wir

$$F_i(X) := \frac{1}{\epsilon}(\beta_\epsilon(\sum_{j=1}^{k} X_j z_j), z_i),$$
$$A_{ij} := (\nabla z_j, \nabla z_i),$$

so liefert das Testen von (P_ϵ^k) mit der bezüglich $(\cdot, \cdot)$ orthonormalen Basis $z_1, ..., z_k$ von V^k

$$U'(t) + AW(t) = 0,$$
$$\gamma AU(t) + F(U(t)) - U(t) = W(t)$$

also $U'(t) = AU(t) - \gamma A^2 U(t) - AF(U(t)) =: G(U(t))$. Falls G lipschitz-stetig ist, folgt aus dem Satz von Picard-Lindelöf (siehe etwa [DB02]) die Existenz einer eindeutig bestimmten Lösung $U \in C^1([0; T]; \mathbb{R}^k)$. Da lineare Abbildungen sowie Linearkombinationen und Kompositionen lipschitz-stetiger Abbildungen lipschitz-stetig sind, ist nur noch zu zeigen, dass F lipschitz-stetig ist:

$$\begin{aligned}
|F_i(X) - F_i(Y)| &\underset{\text{(a)}}{\leq} \frac{1}{\epsilon} \int_\Omega \left| \sum_{j=1}^{k} (X_j - Y_j) z_j(x) \right| |z_i(x)| dx \\
&\leq \frac{1}{\epsilon} \sum_{j=1}^{k} |X_j - Y_j| \int_\Omega |z_j(x)||z_i(x)| dx \\
&\underset{\text{(b)}}{\leq} \frac{1}{\epsilon} \sum_{j=1}^{k} |X_j - Y_j| \|z_j\|_0 \|z_i\|_0 \underset{\text{(c)}}{\leq} \frac{1}{\epsilon} \sum_{j=1}^{k} |X_j - Y_j|.
\end{aligned}$$

Dabei gilt (a), da β_ϵ lipschitz-stetig ist, (b) nach der Cauchy-Schwarz-Ungleichung und (c), da $z_1, ..., z_k$ orthonormal bezüglich $(\cdot,\cdot)$ sind. Damit ist die Lipschitz-Stetigkeit von F mit $\|F(X) - F(Y)\|_{\mathbb{R}^k,1} \leq \frac{k}{\epsilon}\|X - Y\|_{\mathbb{R}^k,1}$ in der 1-Norm $\|X\|_{\mathbb{R}^k,1} = \sum |X_i|$ auf $\mathbb{R}^k$ gezeigt. Um die Massenerhaltung für u^k zu zeigen, testen wir (3.2.6) mit $\eta \equiv 1$ und erhalten

$$\frac{d}{dt}(u^k(t), 1) = \left(\frac{du^k(t)}{dt}, 1\right) = 0,$$

also $(u^k(t), 1) = (u^k(0), 1) = (P^k u_0, 1) = (u_0, 1) = m$. ■

Nun werden Abschätzungen für das zu ψ_ϵ gehörende Ginzburg-Landau Energiefunktional gezeigt, um Abschätzungen für die Lösung von (P_ϵ^k) herzuleiten.

Definition 3.2.3 Das zu ψ_ϵ gehörende Ginzburg-Landau Energiefunktional ist für $v \in H^1(\Omega)$ definiert als

$$\mathcal{E}^\epsilon(v) := \frac{1}{2}\gamma|v|_1^2 + (\psi_\epsilon(v), 1).$$

Lemma 3.2.3 *Die Ginzburg-Landau Energie der Lösung von (P_ϵ^k) erfüllt die Gleichung*

$$\mathcal{E}^\epsilon(u^k(t)) + \int_0^t |w^k(s)|_1^2 ds = \mathcal{E}^\epsilon(P^k(u_0))$$

und es gilt (für k abhängig von ϵ groß genug) $\mathcal{E}^\epsilon(P^k(u_0)) \leq C_1$ mit einer von ϵ und k unabhängigen, positiven Konstante C.

BEWEIS. (Lemma 3.2.3) Sei $(u^k, w^k)^T$ Lösung von (P_ϵ^k). Dann gilt

$$\begin{aligned}\frac{d\mathcal{E}^\epsilon(u^k(t))}{dt} &\underset{\text{(a)}}{=} \gamma(\nabla u^k(t), \nabla\cdot)(\frac{du^k(t)}{dt}) + \int_\Omega \psi_\epsilon'(u^k(t))\frac{du^k(t)}{dt}dx \\ &= \gamma\left(\nabla u^k(t), \nabla\frac{du^k(t)}{dt}\right) + \left(\psi_\epsilon'(u^k(t)), \frac{du^k(t)}{dt}\right) \\ &\underset{\text{(b)}}{=} \left(w^k(t), \frac{du^k(t)}{dt}\right) \underset{\text{(c)}}{=} -|w^k(t)|_1^2\end{aligned}$$

wobei in (a) die Kettenregel, in (b) Gleichung (3.2.7) und in (c) Gleichung (3.2.6) aus (P_ϵ^k) verwendet wurde. Nach Integration über $(0; t)$ erhalten wir

$$\mathcal{E}^\epsilon(P^k(u_0)) = \mathcal{E}^\epsilon(u^k(t)) - \int_0^t \frac{d\mathcal{E}^\epsilon(u^k(s))}{ds}ds = \mathcal{E}^\epsilon(u^k(t)) + \int_0^t |w^k(s)|_1^2 ds.$$

Um $\mathcal{E}^\epsilon(P^k(u_0))$ abzuschätzen, wenden wir Bemerkung 3.2.1 (5), Lemma 3.2.1 (4) und die starke Konvergenz von $P^k v$ gegen v an und erhalten

$$\begin{aligned}\mathcal{E}^\epsilon(P^k u_0) &= \frac{1}{2}\gamma|P^k u_0|_1^2 + (\psi_\epsilon(P^k u_0), 1)\\ &\leq \frac{1}{2}\gamma|u_0|_1^2 + (\widehat{\psi}_\epsilon(u_0), 1) + \frac{1}{2}(1-(P^k u_0)^2, 1) + \frac{1}{\epsilon}\|P^k u_0 - u_0\|_0^2\\ &\xrightarrow[k\to\infty]{} \frac{1}{2}\gamma|u_0|_1^2 + \underbrace{(\widehat{\psi}_\epsilon(u_0), 1)}_{=0,\text{ da } |u_0|\leq 1} + \frac{1}{2}(1-u_0^2, 1) = \mathcal{E}^\epsilon(u_0).\end{aligned}$$

$\mathcal{E}^\epsilon(P^k u_0)$ ist somit nach oben durch eine gegen $\mathcal{E}^\epsilon(u_0)$ konvergente Folge beschränkt. Damit gilt

$$\limsup_{k\to\infty} \mathcal{E}^\epsilon(P^k u_0) \leq \mathcal{E}^\epsilon(u_0).$$

Wegen $|u_0| \leq 1$ ist $\mathcal{E}^\epsilon(u_0)$ unabhängig von ϵ, also gilt

$$\exists C \quad \forall \epsilon > 0 \quad \exists k_\epsilon \quad \forall k > k_\epsilon : \mathcal{E}^\epsilon(P^k u_0) \leq C_1.$$

■

Dieses Resultat wird in der folgenden Form benötigt:

Korollar 3.2.2 *Es gilt mit von T, k und ϵ unabhängigen Konstanten C, C_0, C_1 für $\epsilon < \frac{1}{4}$ und k (abhängig von ϵ) groß genug*

$$\begin{aligned}\frac{1}{2}\gamma|u^k(t)|_1^2 + \int_0^t |w^k(s)|_1^2 ds &= \mathcal{E}^\epsilon(P^k(u_0)) - (\psi_\epsilon(u^k(t)), 1)\\ &\leq C_1 + C_0|\Omega|\epsilon \leq C\end{aligned}$$

Damit gelten insbesondere die Abschätzungen $|u^k(t)|_1^2 \leq C$ und $\int_0^T |w^k(s)|_1^2 ds \leq C$.

BEWEIS. Wende Lemma 3.2.3 und Korollar 3.2.1 an. ■

Nun können Abschätzungen der Lösung von (P_ϵ^k) gezeigt werden.

Lemma 3.2.4 *Für $\epsilon < \frac{1}{4}$ genügt u^k mit von k, ϵ und T unabhängigen Konstanten C den Abschätzungen*

$$\|u^k\|_{W(0,T;H^1(\Omega))} \leq C(1+T^{\frac{1}{2}}), \tag{3.2.9}$$

$$\|u^k\|_{L^\infty(0,T;H^1(\Omega))} \leq C. \tag{3.2.10}$$

BEWEIS. Aus der Massenerhaltung in Lemma 3.2.2 und der Abschätzung von $|u^k(t)|_1$ in Korollar 3.2.2 folgen nach Anwendung der Poincaré-Ungleichung T-unabhängige Abschätzungen

von $\|u^k(t)\|_0$ und $\|u^k(t)\|_1$. Damit ist (3.2.10) sofort und eine Abschätzung in $L^2(0,T;H^1(\Omega))$ durch $CT^{\frac{1}{2}}$ nach Integration über $(0;T)$ gegeben. Für (3.2.9) bleibt noch die Zeitableitung $\frac{du^k}{dt}$ in $L^2(0,T;(H^1(\Omega))')$ abzuschätzen:

$$\begin{aligned}\left\|\frac{du^k(t)}{dt}\right\|^2_{(H^1(\Omega))'} &\underset{\text{(a)}}{\leq} \left\|\frac{du^k(t)}{dt}\right\|^2_{-1} \underset{\text{(b)}}{=} \left(\frac{du^k(t)}{dt}, P^k\mathcal{G}\frac{du^k(t)}{dt}\right)\\ &\underset{\text{(c)}}{=} -\left(\nabla w^k(t), \nabla P^k\mathcal{G}\frac{du^k(t)}{dt}\right) \underset{\text{(d)}}{=} -\left(\nabla w^k(t), \nabla\mathcal{G}\frac{du^k(t)}{dt}\right)\\ &\underset{\text{(e)}}{=} -\left(w^k(t), \frac{du^k(t)}{dt}\right) \underset{\text{(f)}}{=} |w^k(t)|_1^2.\end{aligned}$$

Dabei folgt (a) aus $\frac{du^k(t)}{dt} \in \mathcal{F}$ und Eigenschaften der Norm von $(H^1(\Omega))'$ in Bemerkung 2.5.2, (b) und (d) aus Eigenschaften der Projektion P^k in Lemma 3.2.1, (c) und (f) aus Gleichung (3.2.6) in (P_ϵ^k) und (e) aus der Definition von $\mathcal{G}$. Dies liefert zusammen mit Korollar 3.2.2 eine Abschätzung von $\frac{du^k}{dt}$ in $L^2(0,T;(H^1(\Omega))')$ durch C und damit (3.2.9). ■

Lemma 3.2.5 *Für $\epsilon < \frac{1}{4}$ genügt w^k mit einer von k, ϵ und T unabhängigen Konstante C der Abschätzung*

$$\|w^k\|_{L^2(0,T;H^1(\Omega))} \leq C(1+T^{\frac{1}{2}}). \tag{3.2.11}$$

BEWEIS. Wir schätzen zunächst $|(w^k(t),1)|$ ab. Die Poincaré-Ungleichung aus Bemerkung 2.5.1 und die Abschätzung von $|w^k(t)|_1$ aus Lemma 3.2.2 werden dann die gewünschte Abschätzung von w^k in der $L^2(0,T;H^1(\Omega))$-Norm liefern. Zunächst ergibt das Testen von (3.2.7) in (P_ϵ^k) mit $\eta = 1$ und $\eta = u^k(t)$

$$\begin{aligned}(w^k(t),1) &= \frac{1}{\epsilon}(\beta_\epsilon(u^k(t)),1) - (u^k(t),1),\\ \frac{1}{\epsilon}(\beta_\epsilon(u^k(t)),u^k(t)) &= (w^k(t),u^k(t)) + \|u^k(t)\|_0^2 - \gamma|u^k(t)|_1^2.\end{aligned}$$

Nutzen wir dies, die Beziehungen $|\beta_\epsilon(r)| \leq r\beta_\epsilon(r)$ aus Bemerkung 3.2.1 (6), die Massenerhaltung aus Lemma 3.2.2 und die Abschätzung von $\|u^k(t)\|_0$ aus dem Beweis von Lemma 3.2.4, so erhalten wir

$$\begin{aligned}|(w^k(t),1)| &= |\frac{1}{\epsilon}(\beta_\epsilon(u^k(t)),1) - (u^k(t),1))|\\ &\leq \frac{1}{\epsilon}(|\beta_\epsilon(u^k(t))|,1) + |m|\\ &\leq \frac{1}{\epsilon}(\beta_\epsilon(u^k(t)),u^k(t)) + |m|\\ &= (w^k(t),u^k(t)) + \|u^k(t)\|_0^2 - \gamma|u^k(t)|_1^2 + |m|\\ &\leq C + (w^k(t),u^k(t)).\end{aligned} \tag{3.2.12}$$

Wir benötigen nun eine Abschätzung von $(w^k(t), u^k(t))$. Die Massenerhaltung aus Lemma 3.2.2, die Definition des Green-Operators und Bemerkung 2.5.2 sowie die Abschätzung von $\|u^k(t)\|_0$ aus dem Beweis von Lemma 3.2.4 liefern uns dafür

$$\begin{aligned}(w^k(t), u^k(t)) &= (w^k(t), u^k(t) - \frac{m}{|\Omega|}) + (w^k(t), \frac{m}{|\Omega|}) \\ &= (\nabla w^k(t), \nabla \mathcal{G}(u^k(t) - \frac{m}{|\Omega|})) + \frac{m}{|\Omega|}(w^k(t), 1) \\ &\leq |w^k(t)|_1 |\mathcal{G}(u^k(t) - \frac{m}{|\Omega|})|_1 + \frac{m}{|\Omega|}(w^k(t), 1) \\ &\leq |w^k(t)|_1 C_p \|u^k(t) - \frac{m}{|\Omega|}\|_0 + \frac{m}{|\Omega|}(w^k(t), 1) \\ &\leq |w^k(t)|_1 C + \frac{|m|}{|\Omega|}|(w^k(t), 1)|.\end{aligned} \tag{3.2.13}$$

Aus (3.2.12) und (3.2.13) folgt nun

$$\left(1 - \frac{|m|}{|\Omega|}\right) |(w^k(t), 1)| \leq C + C|w^k(t)|_1.$$

Da wir $|(u_0, 1)| = |m| < |\Omega|$ vorausgesetzt haben, ist $|(w^k(t), 1)|$ somit abgeschätzt und wir erhalten mit der Poincaré-Ungleichung

$$\|w^k(t)\|_0 \leq C_p(|(w^k(t), 1)| + |w^k(t)|_1) \leq C + C|w^k(t)|_1,$$

was nach Integration über $(0; T)$ mit Korollar 3.2.2 $\|w^k\|_{L^2(0,T;H^1(\Omega))} \leq C(1 + T^{\frac{1}{2}})$ ergibt. ■

Nun können wir den Existenzsatz für Lösungen von (P_ϵ) beweisen.

BEWEIS.(Satz 3.2.1) Es sei $(u^k, w^k)^T$ Lösung von (P_ϵ^k). Da dann für k groß genug u^k in $W(0, T; H^1(\Omega))$ und $L^\infty(0, T; H^1(\Omega))$ und w^k in $L^2(0, T; H^1(\Omega))$ unabhängig von ϵ und k beschränkt ist, können wir nacheinander schwach und schwach-* konvergente Teilfolgen in diesen Räumen finden. Nach Satz 2.5.1 (2) und Satz 2.6.3 (6) ist die Einbettung

$$W(0, T; H^1(\Omega)) \hookrightarrow L^2(0, T; L^2(\Omega))$$

kompakt, deshalb können wir auch eine in $L^2(0, T; L^2(\Omega))$ stark konvergente Teilfolge auswählen. Weil natürlich auch alle Teilfolgen den Abschätzungen von u^k und w^k genügen, können wir iterativ schließlich eine Teilfolge finden, die folgende Konvergenzeigenschaften besitzt:

$$u^k \rightharpoonup u_{\epsilon,1} \text{ in } W(0, T; H^1(\Omega)), \tag{3.2.14}$$

$$u^k \rightharpoonup^* u_{\epsilon,2} \text{ in } L^\infty(0, T; H^1(\Omega)), \tag{3.2.15}$$

$$u^k \to u_{\epsilon,3} \text{ in } L^2(0, T; L^2(\Omega)), \tag{3.2.16}$$

$$w^k \rightharpoonup w_\epsilon \text{ in } L^2(0, T; H^1(\Omega)). \tag{3.2.17}$$

Dabei wurde geeignet umnummeriert, um auf eine Doppelindizierung zu verzichten. Mit Bemerkung 2.4.1 folgt aus (3.2.16) nun $u^k \rightharpoonup u_{\epsilon,3}$ in $L^2(0, T; L^2(\Omega))$. Die Einbettung der Räume

$W(0,T;H^1(\Omega)) \hookrightarrow L^2(0,T;H^1(\Omega)) \hookrightarrow L^2(0,T;L^2(\Omega))$ mittels der Identität ist nach Satz 2.6.3 (5) stetig. Damit folgt nach Bemerkung 2.4.1 aus (3.2.14) auch, dass $u^k \rightharpoonup u_{\epsilon,1}$ in $L^2(0,T;L^2(\Omega))$ gilt. Auch aus (3.2.15) folgt mit Bemerkung 2.6.3 und Bemerkung 2.4.1 $u^k \rightharpoonup u_{\epsilon,2}$ in $L^2(0,T;H^1(\Omega))$ und $L^2(0,T;L^2(\Omega))$. Da schwache Grenzwerte eindeutig bestimmt sind, gilt somit

$$u_{\epsilon,1} = u_{\epsilon,2} = u_{\epsilon,3} =: u_\epsilon \in W(0,T;H^1(\Omega)) \cap L^\infty(0,T;H^1(\Omega)),$$
$$w_\epsilon \in L^2(0,T;H^1(\Omega)).$$

Nach Bemerkung 2.1.1 und Bemerkung 2.1.3 ist der Grenzwert der schwach und schwach-* konvergenten Folge durch die gleichen Konstanten beschränkt, es gilt also

$$\begin{aligned} \|u_\epsilon\|_{W(0,T;H^1(\Omega))} &\leq C(1+T^{\frac{1}{2}}), \\ \|u_\epsilon\|_{L^\infty(0,T;H^1(\Omega))} &\leq C, \\ \|w_\epsilon\|_{L^2(0,T;H^1(\Omega))} &\leq C(1+T^{\frac{1}{2}}). \end{aligned}$$

Damit liegen u_ϵ und w_ϵ in den „richtigen“ Räumen und erfüllen die geforderten Abschätzungen. Wir zeigen nun, dass $(u_\epsilon, w_\epsilon)^T$ (P_ϵ) löst. Dafür benötigen wir die Konvergenzresultate in folgender Form:

$$u^k \rightharpoonup u_\epsilon \text{ in } L^2(0,T;L^2(\Omega)) \text{ und } L^2(0,T;H^1(\Omega)), \tag{3.2.18}$$
$$w^k \rightharpoonup w_\epsilon \text{ in } L^2(0,T;L^2(\Omega)) \text{ und } L^2(0,T;H^1(\Omega)). \tag{3.2.19}$$

Nun testen wir die Gleichungen (3.2.6) und (3.2.7) in (P_ϵ^k) mit $P^k\eta \in V^k$ für beliebiges $\eta \in H^1(\Omega)$, multiplizieren sie mit $\phi \in D((0;T))$ und integrieren über $(0;T)$. Da $\eta\phi \in W(0,T;H^1(\Omega))$ gilt, folgt aus (3.2.18) und der Orthogonalität der Projektion P^k aus Lemma 3.2.1

$$\begin{aligned} (u^k, P^k\eta\phi)_{L^2L^2} = (u^k, \eta\phi)_{L^2L^2} &\xrightarrow[k\to\infty]{} (u_\epsilon, \eta\phi)_{L^2L^2}, \\ (u^k, P^k\eta\phi)_{L^2H^1} = (u^k, \eta\phi)_{L^2H^1} &\xrightarrow[k\to\infty]{} (u_\epsilon, \eta\phi)_{L^2H^1}. \end{aligned}$$

Wegen $(\cdot,\cdot)_{L^2\nabla} = (\cdot,\cdot)_{L^2H^1} - (\cdot,\cdot)_{L^2L^2}$ gilt dann auch

$$(u^k, P^k\eta\phi)_{L^2\nabla} = (u^k, \eta\phi)_{L^2\nabla} \xrightarrow[k\to\infty]{} (u_\epsilon, \eta\phi)_{L^2\nabla}.$$

Die gleichen Aussagen für w^k und w_ϵ folgen analog aus (3.2.19). Nach Definition der Zeitableitung und aufgrund der Orthogonalität von P^k gilt ferner

$$\left(\frac{du^k}{dt}, P^k\eta\phi\right)_{L^2L^2} = -(u^k, \eta\phi')_{L^2L^2} \xrightarrow[k\to\infty]{} -(u_\epsilon, \eta\phi')_{L^2L^2} = \int_0^T \left\langle \frac{du_\epsilon}{dt}, \eta \right\rangle \phi(t)dt.$$

Damit bleibt noch der Grenzübergang für den nichtlinearen Term zu zeigen. Aus der Lipschitz-Stetigkeit von β_ϵ in Bemerkung 3.2.1, der Cauchy-Schwartz-Ungleichung und der Abschätzung $\|P^k\eta\|_0 \leq \|\eta\|_0$ aus Lemma 3.2.1 folgt

$$\begin{aligned}
&(\beta_\epsilon(u^k(t)), P^k\eta) - (\beta_\epsilon(u_\epsilon(t)), \eta) \\
&\quad \leq |(\beta_\epsilon(u^k(t)) - \beta_\epsilon(u_\epsilon(t)), P^k\eta)| + |(\beta_\epsilon(u_\epsilon(t)), P^k\eta - \eta)| \\
&\quad \leq \|u^k - u_\epsilon\|_0 \|\eta\|_0 + \|\beta_\epsilon(u_\epsilon(t))\|_0 \|\eta - P^k\eta\|_0.
\end{aligned}$$

Da $L^2(0,T;L^2(\Omega))$ stetig in $L^1(0,T;L^2(\Omega))$ eingebettet ist, konvergiert $u^k \to u_\epsilon$ auch in $L^1(0,T;L^2(\Omega))$ stark. Also konvergiert nach Integration über $(0;T)$ der linke Term gegen 0. Ferner konvergiert $P^k\eta \to \eta$ in $L^2(\Omega)$. Aus $|\beta_\epsilon(r)| \leq |r|$ in Bemerkung 3.2.1, der Abschätzung von u^k in $L^\infty(0,T;H^1(\Omega))$ und $|P^k\eta| \leq |\eta|$ folgt auch, dass der rechte Term unabhängig von k durch eine integrierbare Funktion beschränkt ist. Damit liefert der Konvergenzsatz von Lebesgue (siehe etwa [Rud64])

$$(\beta_\epsilon(u^k), P^k\eta\phi)_{L^2L^2} \qquad \xrightarrow[k\to\infty]{} \qquad (\beta_\epsilon(u_\epsilon), \eta\phi)_{L^2L^2}.$$

Also kann in allen Termen von (P_ϵ^k) ein Grenzübergang durchgeführt werden, der zu

$$\begin{aligned}
&\int_0^T \left\langle \frac{du_\epsilon(t)}{dt}, \eta \right\rangle \phi(t)dt + \int_0^T (\nabla w_\epsilon(t), \nabla\eta)\, \phi(t)dt = 0, \\
&\int_0^T \gamma\, (\nabla u_\epsilon(t), \nabla\eta)\, \phi(t)dt - \int_0^T \left(\psi_\epsilon'(u_\epsilon(t)), \eta\right) \phi(t)dt = \int_0^T (w_\epsilon(t), \eta)\, \phi(t)dt
\end{aligned}$$

für alle $\eta \in H^1(\Omega)$ und alle $\phi \in \mathcal{D}((0;T))$ führt. Somit erfüllen u_ϵ und w_ϵ die Gleichungen aus (P_ϵ). Wir zeigen nun, dass u_ϵ den gewünschten Anfangswert annimmt. Sei dazu $\eta \in H^1(\Omega)$ und $\psi \in C^\infty([0;T])$ mit $\psi(t) = C \neq 0$ in einer Umgebung von 0 und $\psi(t) = 0$ in einer Umgebung von T. Dann gilt $\eta\psi \in W(0,T;H^1(\Omega))$ und damit liefert partielle Integration nach Satz 2.6.3 (2)

$$\begin{aligned}
-(u_\epsilon(0), \eta)\phi(0) &= \int_0^T (u_\epsilon(t), \eta)\phi'(t)dt + \int_0^T \left\langle \frac{du_\epsilon(t)}{dt}, \eta \right\rangle \phi(t)dt \\
&= \lim_{k\to\infty} \left[\int_0^T (u^k(t), P^k\eta\phi'(t))dt + \int_0^T \left(\frac{du^k(t)}{dt}, P^k\eta\phi(t) \right) dt \right] \\
&= \lim_{k\to\infty} \Big[-(u^k(0), P^k\eta)\phi(0) \Big] = -(u_0, \eta)\phi(0).
\end{aligned}$$

Wegen $\phi(0) \neq 0$ und da $\eta \in H^1(\Omega)$ beliebig ist, gilt somit $u_\epsilon(0) = u_0$. ■

Satz 3.2.2 *Für $0 < \epsilon < \frac{1}{4}$ ist die Lösung des Problems (P_ϵ) eindeutig bestimmt.*

BEWEIS. Seien $(u_1, w_1)^T$ und $(u_2, w_2)^T$ Lösungen von (P_ϵ). Setze dann $\tilde{u} = u_1 - u_2$ und $\tilde{w} = w_1 - w_2$. Wir werden zeigen, dass $\|\tilde{u}(t)\|_1 = \|\tilde{w}(t)\|_1 = 0$ für $t \in (0;T)$ gilt. Dazu subtrahieren wir die Gleichungen (3.2.1) und (3.2.2) aus (P_ϵ) jeweils für u_1 und u_2 und erhalten

$$\left\langle \frac{d\tilde{u}(t)}{dt}, \eta \right\rangle + (\nabla \tilde{w}(t), \nabla \eta) = 0, \tag{3.2.20}$$

$$\gamma(\nabla \tilde{u}(t), \nabla \eta) - (\tilde{u}(t), \eta) + \frac{1}{\epsilon}(\beta_\epsilon(u_1(t)) - \beta_\epsilon(u_2(t)), \eta) = (\tilde{w}(t), \eta), \tag{3.2.21}$$

$$\tilde{u}(0) = 0$$

für alle $\eta \in H^1(\Omega)$. Testen von (3.2.20) mit $\eta \equiv 1$ liefert $(\tilde{u}(t), 1) = 0$. Testen mit $\mathcal{G}\tilde{u}(t)$ ergibt nach Korollar 2.6.3

$$\frac{1}{2}\frac{d}{dt}|\mathcal{G}\tilde{u}(t)|_1^2 = \left\langle \frac{d\tilde{u}(t)}{dt}, \mathcal{G}\tilde{u}(t) \right\rangle = -(\nabla \tilde{w}(t), \nabla \mathcal{G}\tilde{u}(t)) = -(\tilde{w}(t), \tilde{u}(t))$$

und das Testen von Gleichung (3.2.21) mit $\tilde{u}(t)$ liefert nach Ausnutzen der Monotonie von β_ϵ, also von $(\beta_\epsilon(u_1) - \beta_\epsilon(u_2), u_1 - u_2) \geq 0$,

$$\gamma|\tilde{u}(t)|_1^2 - (\tilde{w}(t), \tilde{u}(t)) \leq \|\tilde{u}(t)\|_0^2. \tag{3.2.22}$$

Durch Einsetzen erhalten wir damit

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|\mathcal{G}\tilde{u}(t)|_1^2 + \gamma|\tilde{u}(t)|_1^2 \leq \|\tilde{u}(t)\|_0^2 &= (\nabla \tilde{u}(t), \nabla \mathcal{G}\tilde{u}(t)) \\ &\leq |\tilde{u}(t)|_1 |\mathcal{G}\tilde{u}(t)|_1 \\ &\leq \frac{1}{2}\gamma|\tilde{u}(t)|_1^2 + \frac{1}{2\gamma}|\mathcal{G}\tilde{u}(t)|_1^2,\end{aligned}$$

wobei wir $ab \leq \frac{\gamma}{2}a^2 + \frac{1}{2\gamma}b^2$ genutzt haben. Nach Multiplikation mit $2e^{-t/\gamma}$ folgt daraus

$$\frac{d}{dt}\underbrace{\left(e^{-t/\gamma}|\mathcal{G}\tilde{u}(t)|_1^2\right)}_{:=f(t)} + \underbrace{e^{-t/\gamma}\gamma|\tilde{u}(t)|_1^2}_{:=g(t)} \leq 0,$$

was nach Integration über $(0;t)$ wegen $f(t) \geq 0$ und $f(0) = 0$

$$0 \leq \int_0^t g(s)ds \leq \int_0^t g(s)ds + f(t) - f(0) = \int_0^t g(s)ds + \int_0^t f'(s)ds \leq 0$$

liefert. Somit gilt $|\tilde{u}(t)|_1 = 0$ für $t \in (0;T)$, was mit $(\tilde{u}(t), 1) = 0$ und der Poincaré-Ungleichung $\|\tilde{u}(t)\|_1 = 0$, also die Eindeutigkeit von u_ϵ liefert. Testen wir nun (3.2.21) mit $\eta = 1$, so erhalten wir wegen $u_1 = u_2$ auch $(\tilde{w}(t), 1) = -(\tilde{u}(t), 1) = 0$. Um abermals die Poincaré-Ungleichung anzuwenden testen wir noch (3.2.20) mit $\tilde{w}(t)$ und erhalten wegen $\tilde{u} = 0$

$$|\tilde{w}(t)|_1^2 = -\left\langle \frac{d\tilde{u}(t)}{dt}, \tilde{w}(t) \right\rangle = 0.$$

Damit gilt auch $\|\tilde{w}(t)\|_1 = 0$ und die Eindeutigkeit der Lösung ist gezeigt. ■

3.2.6 Lösungen der Variationsungleichung

Auch u_ϵ und w_ϵ erfüllen Abschätzungen, die nun genutzt werden, um durch einen Grenzübergang die Existenz von Lösungen des Problems (P) zu zeigen.

Satz 3.2.3 *Das Problem (P) besitzt eine Lösung.*

BEWEIS. Es sei $(u_\epsilon, w_\epsilon)^T$ die Lösung von (P_ϵ). Da die schwachen Grenzwerte u_ϵ und w_ϵ den ϵ-unabhängigen Abschätzungen aus Satz 3.2.1 unterliegen, existiert analog zum Beweis dieses Satzes wieder eine Teilfolge, so dass nach geeigneter Umnummerierung gilt

$$u_\epsilon \rightharpoonup u \text{ in } W(0,T;H^1(\Omega)), \tag{3.2.23}$$
$$u_\epsilon \rightharpoonup^* u \text{ in } L^\infty(0,T;H^1(\Omega)), \tag{3.2.24}$$
$$u_\epsilon \to u \text{ in } L^2(0,T;L^2(\Omega)), \tag{3.2.25}$$
$$w_\epsilon \rightharpoonup w \text{ in } L^2(0,T;H^1(\Omega)). \tag{3.2.26}$$

Somit liegen u und w in den „richtigen“ Räumen

$$u \in W(0,T;H^1(\Omega)) \cap L^\infty(0,T;H^1(\Omega)),$$
$$w \in L^2(0,T;H^1(\Omega)).$$

Um zu zeigen, dass $(u,w)^T$ eine Lösung von (P) ist, zeigen wir nun zuerst, dass $u \in K$ gilt. Dann zeigen wir, dass $(u_\epsilon, w_\epsilon)^T$ die Variationsungleichung aus (P) erfüllt und gehen schließlich zum Grenzwert über. Zunächst testen wir (3.2.1) aus (P_ϵ) mit $\beta_\epsilon(u_\epsilon(t))$ und erhalten

$$\begin{aligned}
&\gamma(\nabla u_\epsilon(t),\nabla\beta_\epsilon(u_\epsilon(t))) + \frac{1}{\epsilon}\|\beta_\epsilon(u_\epsilon(t))\|_0^2 \\
&\quad = (u_\epsilon(t) + w_\epsilon(t), \beta_\epsilon(u_\epsilon(t))) \\
&\quad \le \|u_\epsilon(t)\|_0\|\beta_\epsilon(u_\epsilon(t))\|_0 + \|w_\epsilon(t)\|_0\|\beta_\epsilon(u_\epsilon(t))\|_0 \\
&\quad \le \frac{2\epsilon}{2}\|u_\epsilon(t)\|_0^2 + \frac{1}{4\epsilon}\|\beta_\epsilon(u_\epsilon(t))\|_0^2 + \frac{2\epsilon}{2}\|w_\epsilon(t)\|_0^2 + \frac{1}{4\epsilon}\|\beta_\epsilon(u_\epsilon(t))\|_0^2 \\
&\quad = \epsilon(\|u_\epsilon(t)\|_0^2 + \|w_\epsilon(t)\|_0^2) + \frac{1}{2\epsilon}\|\beta_\epsilon(u_\epsilon(t))\|_0^2,
\end{aligned} \tag{3.2.27}$$

wobei wir $ab \le \frac{1}{2(2\epsilon)}a^2 + \frac{(2\epsilon)}{2}b^2$ benutzt haben. Ferner gilt $\nabla\beta_\epsilon(u_\epsilon) = \beta_\epsilon'(u_\epsilon)\nabla u_\epsilon$ und nach Bemerkung 3.2.1 (4) $0 \le \beta_\epsilon'(r) \le 1$, woraus

$$(\nabla u_\epsilon(t), \nabla\beta_\epsilon(u_\epsilon(t))) = \int_\Omega \beta_\epsilon'(u_\epsilon(t))(\nabla u_\epsilon(t))^2 dx \ge \int_\Omega \beta_\epsilon'(u_\epsilon(t))^2(\nabla u_\epsilon(t))^2 dx = |\beta_\epsilon(u_\epsilon(t))|_1^2$$

folgt. Zusammen mit (3.2.27) ergibt sich daraus

$$\gamma|\beta_\epsilon(u_\epsilon(t))|_1^2 + \frac{1}{2\epsilon}\|\beta_\epsilon(u_\epsilon(t))\|_0^2 \le \epsilon(\|u_\epsilon(t)\|_0^2 + \|w_\epsilon(t)\|_0^2).$$

Daraus folgt nach Integration über $(0;T)$ mit den (ϵ-unabhängigen) Abschätzungen von u_ϵ und w_ϵ aus Satz 3.2.1

$$\|\beta_\epsilon(u_\epsilon)\|_{L^2(0,T;L^2(\Omega))} \leq C\epsilon$$
$$\|\beta_\epsilon(u_\epsilon)\|_{L^2(0,T;H^1(\Omega))} \leq C\epsilon^{\frac{1}{2}}.$$

Damit gilt nach Bemerkung 3.2.2

$$\begin{aligned}\|\beta(u)\|_{L^2L^2} &\leq \|\beta(u)-\beta(u_\epsilon)\|_{L^2L^2} + \|\beta(u_\epsilon)-\beta_\epsilon(u_\epsilon)\|_{L^2L^2} + \|\beta_\epsilon(u_\epsilon)\|_{L^2L^2} \\ &\leq \|u-u_\epsilon\|_{L^2L^2} + \left\|\frac{\epsilon}{2}\right\|_{L^2L^2} + C\epsilon \xrightarrow[\epsilon\to 0]{} 0,\end{aligned}$$

also fast überall $\beta(u)=0$ und damit $u(t)\in K$ fast überall in $(0;T)$. Für $\eta\in K$ gilt offenkundig $\beta_\epsilon(\eta)=0$. Testen wir nun (3.2.5) aus (P_ϵ) mit $\eta-u_\epsilon$ und nutzen ferner die Monotonie von β_ϵ, so erhalten wir

$$\gamma(\nabla u_\epsilon, \nabla\eta-\nabla u_\epsilon) - (w_\epsilon+u_\epsilon, \eta-u_\epsilon) = \frac{1}{\epsilon}(\underbrace{\beta_\epsilon(\eta)}_{=0} - \beta_\epsilon(u_\epsilon), \eta-u_\epsilon) \geq 0. \tag{3.2.28}$$

Somit erfüllt u_ϵ die Variationsungleichung in (P). Nun wird der Grenzübergang $\epsilon\to 0$ durchgeführt. Mit den Konvergenzeigenschaften von u_ϵ und w_ϵ ist analog zum Beweis von Satz 3.2.1 der Grenzübergang von Gleichung (3.2.1) in (P_ϵ) zu Gleichung (3.1.1) in (P) problemlos möglich. Gleiches gilt für den Startwert u_0. Beim Grenzübergang von Ungleichung (3.2.28) zu Ungleichung (3.1.2) in (P) ist allerdings Vorsicht geboten. Wir multiplizieren zunächst (3.2.28) mit einem beliebigen $\phi\in D((0;T))$ mit $\phi\geq 0$, integrieren wieder über $(0;T)$ und erhalten

$$\int_0^T (w_\epsilon+u_\epsilon, u_\epsilon)\phi dt - \int_0^T (w_\epsilon+u_\epsilon,\eta)\phi dt + \int_0^T (\nabla u_\epsilon, \nabla\eta)\phi dt \geq \gamma\int_0^T |u_\epsilon|_1^2\phi dt.$$

Im zweiten und dritten Term können wir analog zum Beweis von Satz 3.2.1 einen Grenzübergang durchführen. Da $u_\epsilon\to u$ in $L^2(0,T;L^2(\Omega))$ stark und w_ϵ in $L^2(0,T;L^2(\Omega))$ schwach konvergiert, bereitet nach Bemerkung 2.1.2 auch der erste Term keine Probleme. Da in $W(0,T;H^1(\Omega))$ aber nur schwache Konvergenz gegeben ist, ist $|u_\epsilon|_1^2\to|u|_1^2$ nicht zu erwarten. Deshalb bilden wir statt des Grenzwertes den Limes Inferior und erhalten

$$\int_0^T (\nabla u(t), \nabla\eta)\phi(t)dt - \int_0^T (w(t)+u(t), \eta-u(t))\phi(t)dt \geq \liminf_{\epsilon\to 0}\gamma\int_0^T |u_\epsilon(t)|_1^2\phi(t)dt. \tag{3.2.29}$$

Um die Ungleichung in (P) zu erhalten, wenden wir die Beziehung $0\leq(a-b)^2=a^2-2ab+b^2$ beziehungsweise $2ab-b^2\leq a^2$ auf die Bilinearform $(\nabla\cdot,\nabla\cdot)$ an und erhalten

$$2\gamma\int_0^T (\nabla u_\epsilon(t), \nabla u(t))\phi(t)dt - \gamma\int_0^T |u(t)|_1^2\phi(t)dt \leq \gamma\int_0^T |u_\epsilon(t)|_1^2\phi(t)dt,$$

woraus nach dem Bilden des Limes Inferior wegen der Konvergenz von $(\nabla u_\epsilon, \nabla u)$ gegen $|u|_1^2$

$$\gamma \int_0^T |u(t)|_1^2 \phi(t) dt \leq \liminf_{\epsilon \to 0} \gamma \int_0^T |u_\epsilon(t)|_1^2 \phi(t) dt$$

folgt. Zusammen mit (3.2.29) ergibt dies die Ungleichung in (P). ■

Satz 3.2.4 *Die Lösung von (P) ist eindeutig bestimmt.*

BEWEIS. Es seien $(u_1, w_1)^T$ und $(u_2, w_2)^T$ Lösungen von (P). Setze dann $\tilde{u} = u_1 - u_2$ und $\tilde{w} = w_1 - w_2$. Anlog zum Beweis der Eindeutigkeit von u_ϵ lässt sich folgendes zeigen:

$$\frac{1}{2}\frac{d}{dt}|\mathcal{G}\tilde{u}(t)|_1^2 = \left\langle \frac{d\tilde{u(t)}}{dt}, \mathcal{G}\tilde{u}(t) \right\rangle = -(\nabla \tilde{w}(t), \nabla \mathcal{G}\tilde{u}(t)) = -(\tilde{w}(t), \tilde{u}(t)). \tag{3.2.30}$$

Nun testen wir (3.1.2) aus (P) für u_1 mit u_2 und für u_2 mit u_1. Dies liefert

$$\gamma(\nabla u_1(t), -\nabla \tilde{u}(t)) - (u_1(t), -\tilde{u}(t)) \geq (w_1(t), -\tilde{u}(t)),$$
$$\gamma(\nabla u_2(t), \nabla \tilde{u}(t)) - (u_2(t), \tilde{u}(t)) \geq (w_2(t), \tilde{u}(t)).$$

Summieren und Multiplikation mit -1 ergibt

$$\gamma|\tilde{u}(t)|_1^2 - \|\tilde{u}(t)\|_0^2 \leq (\tilde{w}(t), \tilde{u}(t)). \tag{3.2.31}$$

Aus (3.2.30) und (3.2.31) folgt genau wie im Beweis der Eindeutigkeit von u_ϵ die Eindeutigkeit von u. Ebenfalls analog dazu gilt

$$|\tilde{w}(t)|_1^2 = -\left\langle \frac{d\tilde{u}(t)}{dt}, \tilde{w}(t) \right\rangle = 0.$$

Damit gilt $\tilde{w}(t) = \text{const}$, also ist w bis auf Konstanten eindeutig bestimmt. Da in (P) eine Ungleichung vorliegt, ist ein Vorgehen wie bei (P_ϵ), um $(\tilde{w}(t), 1) = 0$ zu zeigen, nicht möglich. Um die Ungleichung zu umgehen, nutzen wir die Tatsache $|(u(t), 1)| = |(u_0, 1)| < |\Omega|$ und zeigen, dass u dort, wo es nicht auf dem Hindernis liegt, sogar eine Differentialgleichung erfüllt. Dafür definieren wir die für fast alle $t \in (0; T)$ nichtleeren Mengen

$$\Omega_0(t) := \{x \in \Omega : |u(x, t)| < 1\}.$$

Wegen $|(u(t), 1)| < |\Omega|$ gilt $|\Omega_0(t)| > 0$. Sei nun $\phi \in D(\Omega_0(t))$ beliebig und $\delta \in \mathbb{R}$ jeweils so gewählt, dass $\eta_\pm = u(t) \pm \delta\phi \in K$ gilt, dann liefert das Testen von (3.1.2) mit η_+ und η_-

$$\gamma(\nabla u(t), \nabla \phi) = (u(t) + w(t), \phi),$$

also wegen $\tilde{u} = 0$

$$0 = \gamma(\nabla \tilde{u}(t), \nabla \phi) - (\tilde{u}(t), \phi) = (\tilde{w}(t), \phi) \forall \phi \in D(\Omega_0(t)),$$

und damit $\tilde{w}(x, t) = 0$ in $\Omega_0(t)$, was mit $\tilde{w}(t) = \text{const}$ in Ω zu $\tilde{w}(t) = 0$ führt. ■

Satz 3.2.5 *Die Lösung u von (P) löst das Problem (Q).*

BEWEIS. Sei $(u, w)^T$ die Lösung von (P), dann erhält man durch das Testen von Gleichung (3.1.1) in (P) mit $\eta \equiv 1$ die Massenerhaltung $\left\langle \frac{du}{dt}, 1 \right\rangle = 0$ und $(u, 1) = m$. Nach Definition des Green-Operators folgt somit aus Gleichung (3.1.1)

$$w = -\mathcal{G}\left(\frac{du}{dt}\right) + \lambda$$

mit $\lambda = \frac{(w,1)}{|\Omega|}$. Das Einsetzen dieser Beziehung in Gleichung (3.1.2) aus (P) und testen mit $\eta \in K_m$ liefert genau die Ungleichung aus (Q), denn es gilt

$$(\lambda, \eta - u) = \lambda\left[(1, \eta) - (1, u)\right] = \lambda(m - m) = 0.$$

■

Satz 3.2.6 *Die Lösung von (Q) ist eindeutig bestimmt.*

BEWEIS. Seien u_1 und u_2 Lösungen von (Q). Setze dann $\tilde{u} = u_1 - u_2$. Testen der Ungleichung in (Q) für u_1 mit u_2 und für u_2 mit u_1 liefert

$$\gamma(\nabla u_1(t), -\nabla\tilde{u}(t)) + \left\langle \mathcal{G}\frac{du_1}{dt}, -\tilde{u}(t) \right\rangle \geq (u_1(t), -\tilde{u}(t)),$$
$$\gamma(\nabla u_2(t), \nabla\tilde{u}(t)) + \left\langle \mathcal{G}\frac{du_2}{dt}, \tilde{u}(t) \right\rangle \geq (u_2(t), \tilde{u}(t)).$$

Summieren und Multiplikation mit -1 liefert

$$\gamma|\tilde{u}(t)|_1^2 + \left\langle \mathcal{G}\frac{d\tilde{u}}{dt}, -\tilde{u}(t) \right\rangle \leq \|\tilde{u}(t)\|_0^2.$$

Nun folgt genau wie in den Beweisen der Eindeutigkeit der Lösungen von (P_ϵ) und (P) die Eindeutig von u. ■

Damit sind (P) und (Q) äquivalent, denn eine Lösung von (P) löst (Q), also muss die eindeutige Lösung von (Q) auch die eindeutige Lösung von (P) sein.

3.2.7 Regularität der Lösungen

In [BE91] wird ferner ein Regularitätsresultat für Lösungen von (P) gezeigt. Dabei wird folgende Forderung an den Rand des Gebietes Ω gestellt:

(Reg_Ω) Der Rand von Ω sei so „glatt“, dass Lösungen z von

$$-\triangle z = f, \qquad \frac{\partial z}{\partial \nu} = 0 \tag{3.2.32}$$

mit $(z, 1) = 0$ und $f \in \mathcal{F} \cap L^2(\Omega)$ die Abschätzung $\|z\|_{H^2(\Omega)} \leq C\|\triangle z\|_0$ erfüllen.

Eine ausführliche Diskussion der Regularität solcher Lösungen ist in [Gri85] zu finden. Ein Kriterium ist zum Beispiel durch die folgenden Sätze gegeben.

Satz 3.2.7 *Sei $\Omega \subset \mathbb{R}^d$ polygonal berandet, dann existiert genau eine Lösung von* (3.2.32)*, und diese liegt in $H^2(\Omega)$.*

BEWEIS. Siehe [Gri85]. ■

Satz 3.2.8 *Sei $\Omega \subset \mathbb{R}^d$ konvex, dann existiert eine nur vom Durchmesser von Ω abhängige Konstante, so dass für alle $z \in H^2(\Omega)$ gilt*

$$\|z\|_2 \leq C\|\triangle z\|_0. \tag{3.2.33}$$

BEWEIS. Siehe [Gri85]. ■

Die Forderung der Konvexität kann auch gelockert werden, führt aber zu komplexeren Anforderungen an Ω.

Satz 3.2.9 *Ω genüge der Bedingung (Reg_Ω). Dann gilt für die Lösung von (P)*

$$u \in L^2(0,T;H^2(\Omega))$$

und $\frac{\partial u}{\partial \nu} = 0$ auf $\partial\Omega$ fast überall in $(0;T)$. Ferner gilt für alle $t \in (0;T)$

$$\min\{t^{\frac{1}{2}},1\}\|u(t)\|_2 \leq C,$$
$$\min\{t^{\frac{1}{2}},1\}\|w(t)\|_1 \leq C,$$

wobei C von u_0 abhängt.

BEWEIS. Siehe [BE91]. ■

Dieses Ergebnis wird insbesondere für die Fehlerabschätzung der Diskretisierung in Kapitel 4 benötigt.

4 Eine Diskretisierung der Cahn-Hilliard Gleichung

Um partielle Differentialgleichungen zu lösen, wird oftmals der Computer eingesetzt. Da eine analytische Lösung im Allgemeinen nicht zugänglich ist, wurden verschiedene Verfahren zur Diskretisierung der kontinuierlichen Gleichungen entwickelt. Allen gemeinsam ist, dass sie auf diskrete Probleme führen, zu deren Lösung ein Rechner herangezogen werden kann.

Die Methode der finiten Differenzen geht von einer klassischen Formulierung der Differentialgleichung aus und basiert auf der Approximation klassischer Ableitungen durch Differenzenquotienten an endlich vielen Knoten. Die Methode der finiten Volumen bildet durch Integration Mittelwerte der gesuchten Lösung über endlich viele Kontrollvolumina. Als sehr flexibel hat sich die Methode der finiten Elemente erwiesen. Sie geht von einer Variationsformulierung aus und ermittelt Lösungen in endlichdimensionalen Teilräumen. Damit ist sie ein Spezialfall der Galerkin-Approximation, die bereits im Beweis der Existenz und Eindeutigkeit von Lösungen des Problems (P) verwendet wurde. Die benutzten Funktionenräume bestehen im Allgemeinen aus stetigen, stückweise polynomiellen Funktionen und liegen damit in $H^1(\Omega)$. Finite Elemente Methoden, deren endlichdimensionale Räume Unterräume des kontinuierlichen Raums sind, werden als konform bezeichnet. In bestimmten Fällen werden auch Funktionenräume verwendet, die keine Unterräume des ursprünglichen Funktionenraums bilden, diesen aber dennoch auf gewisse Art approximieren. In diesem Fall spricht man von nichtkonformen Methoden. Eine Einführung in die Methode der finiten Elemente ist etwa in [Bra97] zu finden.

Zur numerischen Behandlung der Cahn-Hilliard Gleichung untersuchen Blowey und Elliot in [BE92] drei Diskretisierungen mittels finiter Elemente. Allen gemein ist, dass es sich um volldiskrete Schemata handelt, die die Zeitableitung mittels Differenzenquotienten diskretisieren. Eine ist in der Zeit voll-implizit, eine behandelt den konkaven Term $(u, \eta - u)$ in (P) explizit und die dritte behandelt ihn nach dem Crank-Nicolson-Schema. Da die semi-imlipizite zweite Variante den Vorteil hat, für alle Zeitschrittweiten eine eindeutige Lösung zu besitzen, wird im Folgenden nur diese betrachtet.

4.1 Die Methode der finiten Elemente

Es soll nun ein sehr kurzer Überblick über die Methode der finiten Elemente gegeben werden. Sie dient der Diskretisierung partieller Differentialgleichungen, um diese numerisch zu lösen und basiert auf einer Variationsformulierung der Gleichung in einem Hilbertraum H. Sei etwa $a(\cdot,\cdot) : H \times H \to \mathbb{R}$ eine stetige, koerzive Bilinearform mit

$$\begin{aligned} |a(u,v)| &\leq \Gamma \|u\|_H \|v\|_H && \forall u, v \in H, \\ a(v,v) &\geq \alpha \|v\|_H^2 && \forall v \in H \end{aligned}$$

und $f \in H'$. Gesucht ist nun die Lösung der Variationsgleichung

$$u \in H : \qquad a(u,v) = f(v) \qquad \forall v \in H.$$

Nun soll das Problem zunächst in einem abgeschlossenem Unterraum $V \subset H$ approximiert werden:

$$u_V \in V : \qquad a(u_V, v) = f(v) \qquad \forall v \in V.$$

Das Céa-Lemma macht einen solchen Ansatz sinnvoll:

Satz 4.1.1 *(Céa-Lemma) Unter den obigen Bedingungen gilt*

$$\|u - u_V\|_H \leq \frac{\Gamma}{\alpha} \inf_{v \in V} \|u - v\|_H.$$

BEWEIS. Siehe [Bra97]. ■

Die Lösung im Unterraum ist also „so gut“, wie V den Raum H approximiert. Im Allgemeinen ist H unendlichdimensional und man möchte das Problem in endlichdimensionalen Räumen approximieren, denn endlichdimensionale Probleme lassen sich durch Darstellung in einer Basis auf Probleme in $\mathbb{R}^N$ zurückführen. Da endlichdimensionale Räume abgeschlossen sind, ist das Céa-Lemma anwendbar. Für partielle Differentialgleichungen ist H meist ein Sobolevraum. Das Standardbeispiel ist $a(u, v) = (\nabla u, \nabla v)$ mit $H = H_0^1(\Omega)$ (siehe zum Beispiel [Bra97]). Der Sobolevraum soll nun durch endlichdimensionale Räume approximiert werden. Dazu wird zunächst eine Triangulierung definiert.

Definition 4.1.1 Eine Zerlegung $\mathcal{T}^h$ von $\Omega \subset \mathbb{R}^d$ heißt Triangulierung, wenn gilt:

(1) $\overline{\Omega} = \bigcup_{\tau \in \mathcal{T}^h} \tau$.

(2) Jedes $\tau \in \mathcal{T}^h$ ist ein abgeschlossener, nicht entarteter d-dimensionaler Simplex mit Durchmesser $\operatorname{diam} \tau \leq h$.

(3) Für $\tau_1, \tau_2 \in \mathcal{T}^h$ mit $\tau_1 \neq \tau_2$ ist $\tau_1 \cap \tau_2$ entweder leer oder ein Simplex der Oberfläche von τ_1 und τ_2 mit Dimension kleiner d.

Die Forderungen (1) und (2) implizieren, dass Ω polygonal berandet ist. Anforderung (3) bedeutet für $d = 1$, dass der Schnitt zweier Elemente (Intervalle) in $\mathcal{T}^h$ entweder leer oder eine gemeinsame Ecke ist. Für $d = 2$ (Dreiecke) ist zusätzlich eine gemeinsame Kante und für $d = 3$ (Tetraeder) außerdem eine gemeinsame Oberfläche (Dreieck) möglich.

Definition 4.1.2 Der zu $\mathcal{T}^h$ gehörige Raum linearer finiter Elemente S^h ist definiert durch

$$S^h = \left\{ \eta \in C(\overline{\Omega}) : \eta|_\tau \text{ ist affin linear } \forall \tau \in \mathcal{T}^h \right\}.$$

Da die Elemente von S^h stückweise linear, also auch stückweise unendlich oft differenzierbar sind, zeigt der folgende Satz, dass $S^h \subset H^1(\Omega)$ gilt.

Satz 4.1.2 *Ω sei beschränkt und $\eta : \overline{\Omega} \to \mathbb{R}$ stückweise unendlich oft differenzierbar. Dann gilt $\eta \in H^1(\Omega)$ genau dann, wenn $\eta \in C(\overline{\Omega})$ gilt.*

BEWEIS. Siehe [Bra97]. ■

Oft werden Familien von Triangulierungen $(\mathcal{T}^h)$ für $h \to 0$ betrachtet. Insbesondere dann ist es wichtig, dass diese gewissen Bedingungen genügen.

Definition 4.1.3

(1) Eine Familie von Triangulierungen $(\mathcal{T}^h)$ heißt quasiuniform (shape regular), wenn ein $\kappa > 0$ existiert, so dass jedes $\tau \in \mathcal{T}^h$ einen Kreis vom Radius $\rho_\tau \geq \frac{h_\tau}{\kappa}$ enthält, wobei h_τ der halbe Durchmesser von τ ist.

(2) Eine Triangulierung $\mathcal{T}^h$ in $\mathbb{R}^d$ heißt schwach spitzwinklig, wenn entweder $d = 1$ gilt oder alle Innenwinkel α_τ der Elemente $\tau \in \mathcal{T}^h$ die Bedingung $\alpha_\tau \leq \frac{\pi}{2}$ erfüllen.

Eine Basis von S^h ist die Knotenbasis:

Definition 4.1.4 Jedes $x \in \overline{\Omega}$, das Extremalpunkt eines $\tau \in \mathcal{T}^h$ ist, heißt Knoten der Triangulierung $\mathcal{T}^h$. Die Menge

$$\mathcal{N}^h = \{x_1, ..., x_N\} = \left\{x \in \overline{\Omega} : x \text{ ist Knoten von } \mathcal{T}^h\right\} \tag{4.1.1}$$

heißt Knotenmenge von $\mathcal{T}^h$. Sei nun $\phi_i \in S^h$ das Element mit $\phi_i(x_j) = \delta_{ij}$, dann heißt die Familie $(\phi_i)_{i=1,...,N}$ Knotenbasis von S^h.

Dabei heißt ein Punkt $x \in K$ in einer konvexen Menge $K \subset \mathbb{R}^d$ Extremalpunkt von K, wenn

$$y_1, y_2 \in K,\ 0 < \lambda < 1,\ \lambda y_1 + (1-\lambda)y_2 = x \Rightarrow y_1 = x = y_2$$

gilt. $\tau \in \mathcal{T}^h$ ist konvex und die Extremalpunkte von τ sind genau die Eckpunkte des Simplex. Da eine affin lineare Funktion $\phi : \mathbb{R}^d \to \mathbb{R}$ durch Vorgabe der Funktionswerte an den $d+1$ Eckpunkten eines d-dimensionalen Simplex eindeutig bestimmt ist, ist $(\phi_i)_{i=1,...,N}$ tatsächlich eine Basis von S^h.

Nach Satz 4.1.2 gilt $S^h \subset H^1(\Omega)$. Wird nun eine lineare Variationsgleichung in $H^1(\Omega)$ approximativ in S^h gelöst, dann zeigt das Céa-Lemma, dass der Fehler der approximativen Lösung u_h von der Approximationsgüte des Unterraums abhängt. Die Frage ist nun, „wie gut“ S^h den Raum $H^1(\Omega)$ approximiert. Um dies zu untersuchen wird der Interpolationsoperator von S^h definiert.

Definition 4.1.5 Der Interpolationsoperator $I^h : C(\overline{\Omega}) \to S^h$ ist definiert durch

$$I^h(u)(x) = \sum_{i=1}^{N} u(x_i)\phi_i(x).$$

Wegen $I^h u \in S^h$ gilt

$$\inf_{v \in S^h} \|u - v\|_1 \leq \|u - I^h u\|_1.$$

Liegt die Lösung u des Problems in $H^2(\Omega) \subset C(\overline{\Omega})$ (siehe Satz 2.5.1), so ist der Interpolationsfehler des Operators I^h abzuschätzen. Dies leistet beispielsweise der folgende Satz:

Satz 4.1.3 *Ist die Familie $(\mathcal{T}^h)$ von Triangulierungen quasiuniform mit Konstante κ, und besitzt $\Omega \subset \mathbb{R}^2$ einen lipschitz-stetigen Rand (siehe [Wlo82]), dann gilt mit einer nur von κ und Ω abhängigen Konstante C für $u \in H^2(\Omega)$ und $r \in \{0, 1\}$*

$$\|u - I^h u\|_r \leq C h^{2-r} \|u\|_2.$$

BEWEIS. Siehe [Bra97]. ■

Sollen nun Variationsprobleme wie das Beispielproblem in S^h gelöst werden, wird u_h in der Knotenbasis dargestellt und die Variationsformulierung mit allen Basisvektoren getestet. Dies führt auf ein lineares Gleichungssystem in den Koeffizienten von u_h. Es besitzt die Form

$$A\underline{u}_h = B, \qquad u_h = \sum_{j=1}^{N} (\underline{u}_h)_j \phi_j,$$
$$A_{ij} = a(\phi_j, \phi_i), \qquad B_i = f(\phi_i).$$

Besonders oft tritt die Bilinearform $(\nabla\cdot, \nabla\cdot)$ auf. Die Matrix zu dieser Bilinearform wird als Steifigkeitsmatrix bezeichnet. Für ihre Diagonalenelemente gilt $A_{ii} = |\phi_i|_1^2 > 0$. Für eine schwach spitzwinklige Triangulierung lässt sich über die anderen Einträge folgendes feststellen:

Bemerkung 4.1.1 *Ist $\mathcal{T}^h$ schwach spitzwinklig, so gilt für die Knotenbasis $\phi_1, ..., \phi_N$ von S^h $(\nabla\phi_i, \nabla\phi_j) \leq 0$ für $i \neq j$.*

BEWEIS. Für $d = 1$ ist mit $i \neq j$ der Schnitt $\mathrm{supp}(\phi_i) \cap \mathrm{supp}(\phi_j)$ entweder leer, eine Nullmenge oder ein Intervall aus $\mathcal{T}^h$, auf dem eine der Funktionen monoton fallend und die andere monoton steigend ist. Damit gilt in diesem Intervall $\phi_i'(x)\phi_j'(x) \leq 0$. Für $d > 1$ ist der Schnitt $\mathrm{supp}(\phi_i) \cap \mathrm{supp}(\phi_j)$ entweder leer, eine Nullmenge, oder er besteht aus Simplizes aus $\mathcal{T}^h$. $\nabla\phi_i(x)$ und $\nabla\phi_j(x)$ sind in diesem Simplex jeweils nach innen gerichtete Normalen von zwei verschiedenen Oberflächenelementen. Da der Winkel zwischen diesen Oberflächenelementen kleiner oder gleich $\frac{\pi}{2}$ ist, ist der Winkel zwischen $\nabla\phi_i(x)$ und $\nabla\phi_j(x)$ größer oder gleich $\frac{\pi}{2}$ und es gilt $\nabla\phi_i(x) \cdot \nabla\phi_j(x) \leq 0$. ■

Da der Schnitt der Träger zweier Basisfunktionen von S^h klein ist, ist die Anzahl der von Null verschiedenen Einträge in A auch klein. Ist $(\mathcal{T}^h)$ quasiuniform, so ist der Innenwinkel der Dreiecke unabhängig von h beschränkt. Somit besitzt jeder Knoten nur eine unabhängig von N beschränkte Anzahl von Nachbarknoten. Damit ist die Anzahl der von Null verschiedenen Einträge in jeder Zeile von A unabhängig von N beschränkt. Matrizen mit dieser Eigenschaft werden „dünn besetzt“ oder „sparse“-Matrix genannt.

Diese Eigenschaft ermöglicht es oft, die entstehenden Gleichungssysteme effizient durch iterative Verfahren zu lösen. Bekannte Beispiele sind für symmetrische Matrizen das CG- und das

PCG-Verfahren, für nichtsymmetrische Matrizen das BICG-, BICGSTAB- und das GMRES-Verfahren (siehe [Mei99]). Allen gemein ist, dass sie dünn besetzte Systeme deutlich schneller lösen als exakte Löser wie die Gauß-Elimination. Da die Lösung durch die Diskretisierung, die Approximation des Problems in einem diskreten Unterraum, ohnehin nicht exakt ist, ist es auch unnötig, die diskreten Probleme wesentlich genauer zu lösen, als es die Größenordnung des Diskretisierungsfehlers ist. Insbesondere im symmetrischen Fall basieren die Verfahren auf Minimierungsansätzen. Besonders hervorzuheben sind auch die Mehrgitterverfahren (siehe zum Beipiel [Bra93], [Bra97]), die eine Gitterhierarchie nutzen und sich in vielen Fällen als überlegen erweisen.

4.2 Definitionen

Jetzt werden einige Begriffe eingeführt, die für die Formulierung der Diskretisierung von (P) benötigt werden.

Es sei im folgenden $(\mathcal{T}^h)$ eine quasiuniforme Familie von Triangulierungen von Ω. In der Diskretisierung wird die Bilinearform $(\nabla\cdot,\nabla\cdot)$ auf $S^h \times S^h$ exakt ausgewertet. Für das L^2-Skalarprodukt $(\cdot,\cdot)$ wird nur eine Näherung $(\cdot,\cdot)_h$ betrachtet. An diese Näherung werden für $\eta,\chi \in S^h$ und stetiges ξ folgende Forderungen mit h-unabhängigen Konstanten $C_1, C_2, C_3 > 0$ gestellt:

(1) $(\cdot,\cdot)_h$ ist ein Skalarprodukt auf S^h und definiert eine Norm $\|\eta\|_h^2 = (\eta,\eta)_h$.

(2) Für $r \in \{0,1\}$ gilt $|(\eta,\chi) - (\eta,\chi)_h| \le C_1 h^{1+r} \|\eta\|_1 \|\chi\|_r$.

(3) $\|\cdot\|_h$ und $\|\cdot\|_0$ sind äquivalent, d.h. es gilt $C_2\|\eta\|_h^2 \le \|\eta\|_0^2 \le C_3\|\eta\|_h^2$.

(4) $(\eta,\chi)_h = (\eta\chi,1)_h$.

(5) $(\eta,1)_h = (\eta,1)$.

(6) $\xi \ge 0 \quad \Rightarrow \quad (\xi,1)_h \ge 0$.

In [BE92] wird zunächst nur die Forderung $(1,1)_h = |\Omega|$ statt (5) gestellt und dafür die Existenz diskreter Lösungen gezeigt. Jedoch wird die Konvergenz der Diskretisierung nur unter der Bedingung (5) gezeigt, weshalb diese im folgenden generell angenommen wird. Forderung (2) wird nur für den Beweis der Fehlerabschätzung der diskreten Lösung verwendet und ist für die Definition und Lösbarkeit der diskreten Probleme irrelevant.

Beispiele für ein solches Skalarprodukt sind das normale L^2-Skalarprodukt $(\cdot,\cdot)$ und das „gelumpte“ L^2-Skalarprodukt. Man erhält es, wenn vor der Integration eine Interpolation in S^h (siehe Definition 4.1.5) durchgeführt wird, also

$$(\eta,\chi)_h := \int_\Omega I^h(\eta\chi)(x)dx = \sum_{i=1}^N \eta(x_i)\chi(x_i) \int_\Omega \phi_i(x)dx,$$

$$\|\eta\|_h^2 := \int_\Omega I^h(\eta^2)(x)dx = \sum_{i=1}^N \eta^2(x_i) \int_\Omega \phi_i(x)dx.$$

Das Integral des L^2-Skalarproduktes wird also durch eine Quadratur an den Knoten der Triangulierung mit den Integralen über die jeweiligen Knotenbasisfunktionen als Gewichten approximiert. Diese Quadratur integriert genau die Funktionen aus S^h exakt, das Produkt solcher Funktionen jedoch im Allgemeinen nicht.

Nun werden die zur Diskretisierung von (P) benötigten Räume und Mengen definiert.

Definition 4.2.1

(1) P^h ist die diskrete L^2-Projektion in S^h mit

$$(P^h\chi, \eta)_h = (\chi, \eta) \qquad \forall \chi \in L^2(\Omega), \eta \in S^h.$$

(2) $S_m^h := \{\eta \in S^h : (\eta, 1)_h = (P^h u_0, 1)_h = (u_0, 1) = m\}$.

(3) $S_0^h := \{\eta \in S^h : (\eta, 1)_h = (P^h u_0, 1)_h = 0\}$.

(4) $K^h := \{\eta \in S^h : |\eta(x)| \leq 1\}$.

(5) $K_m^h := \{\eta \in S_m^h : |\eta(x)| \leq 1\}$.

Zum Skalarprodukt $(\cdot, \cdot)_h$ existiert auch eine diskrete Form der Poincaré-Ungleichung.

Bemerkung 4.2.1 *(Diskrete Poincaré-Ungleichung) Für h klein genug und $\eta \in S^h$ gilt*

$$\|\eta\|_h \leq C_p^h(|\eta|_1 + |(\eta, 1)_h|).$$

BEWEIS. Aus der Äquivalenz von $\|\cdot\|_0$ und $\|\cdot\|_h$ auf S^h und der Poincaré-Ungleichung für $H^1(\Omega)$ folgt

$$\|\eta\|_h \leq C\|\eta\|_0 \leq CC_p\left(|\eta|_1 + |(\eta, 1)|\right) = CC_p\left(|\eta|_1 + |(\eta, 1)_h|\right).$$

■

Ferner wird eine diskrete Form des Green-Operators benötigt. Dieser wird analog zum kontinuierlichen Green-Operator definiert. Die Existenz folgt ebenso analog aus der diskreten Poincaré-Ungleichung und dem Satz von Lax-Milgram (siehe [Wlo82]).

Definition 4.2.2 Es sei $\mathcal{F}^h := S_0^h$. Definiere nun den Green-Operator mit Neumann-Randbedingungen $\mathcal{G}^h : \mathcal{F}^h \to S^h, f \mapsto \mathcal{G}^h f$, wobei gilt:

$$\begin{aligned}(\nabla \mathcal{G}^h f, \nabla \eta) &= (f, \eta)_h \qquad \forall \eta \in S^h,\\ (\mathcal{G}^h f, 1)_h &= 0.\end{aligned}$$

Auch durch den diskreten Green-Operator wird eine Norm auf $\mathcal{F}^h$ erklärt mittels:

$$\|\eta\|_{-h}^2 := |\mathcal{G}^h\eta|_1^2 = (\mathcal{G}^h\eta, \eta)_h.$$

4.3 Die semi-implizite Diskretisierung

Nun wird die semi-implizite Diskretisierung von (P) aufgestellt. Für die Zeitdiskretisierung wird $(0;T)$ in M Intervalle $J^n = (t^{n-1}, t^n]$ der Länge $\tau = \frac{T}{M}$ unterteilt. Es gilt also $t^n = n\tau$ für $n = 0, ..., M$. Es sind nun Lösungen $(U^n, W^n)^T \in S^h \times S^h$ zu den Zeitpunkten t^n für $n = 1, ..., M$ gesucht. Dabei wird die Zeitableitung diskretisiert durch

$$\frac{\partial u}{\partial t} \approx \frac{U^n - U^{n-1}}{\tau}.$$

Dies führt auf das diskrete Problem:

(P^h) Zu gegebenem u_0 mit $|(u_0, 1)| < |\Omega|$ finde für $n = 1, ..., M$ $(U^n, W^n)^T \in K^h \times S^h$ mit

$$(U^n, \eta)_h + \tau(\nabla W^n, \nabla \eta) = (U^{n-1}, \eta)_h \qquad \forall \eta \in S^h, \tag{4.3.1}$$

$$\gamma(\nabla U^n, \nabla \eta - \nabla U^n) - (W^n, \eta - U^n)_h \geq (U^{n-1}, \eta - U^n)_h \qquad \forall \eta \in K^h, \tag{4.3.2}$$

$$U^0 = P^h u_0. \tag{4.3.3}$$

Diese Diskretisierung ist nur semi-implizit, da in der rechten Seite von (4.3.2) auf U^{n-1} statt U^n zurückgegriffen wird. Das hat den Vorteil, dass das Problem für alle Schrittweiten τ eine Lösung besitzt. Analog zur kontinuierlichen Gleichung wird nun eine zweite Version des Problems aufgestellen.

(Q^h) Zu gegebenem u_0 mit $|(u_0, 1)| < |\Omega|$ finde für $n = 1, ..., M$ $\quad U^n \in K_m^h$ mit

$$\gamma(\nabla U^n, \nabla \eta - \nabla U^n) + (\mathcal{G}^h(\frac{U^n - U^{n-1}}{\tau}), \eta - U^n)_h \geq (U^{n-1}, \eta - U^n)_h \qquad \forall \eta \in K_m^h \tag{4.3.4}$$

$$U^0 = P^h u_0. \tag{4.3.5}$$

Hier wurde W^n durch $\mathcal{G}^h(\frac{U^n - U^{n-1}}{\tau})$ ersetzt. Damit ist die Massenerhaltung in U^n nicht mehr implizit gegeben, weshalb sie nun explizit durch $U^n \in K_m^h$ gefordert wird. Da außerdem $(\mathcal{G}^h V, 1)_h = 0$ gilt, jedoch keineswegs $(W^n, 1)_h = 0$ gelten muss, wird die Darstellung

$$W^n = W_0^n + \underbrace{\frac{(W^n, 1)_h}{|\Omega|}}_{=\text{const}}$$

mit $(W_0^n, 1)_h = 0$ zugrundegelegt. Um den Term $(\text{const}, \eta - U^n)_h$ mit der unbekannten Konstante zu handhaben, wird in (Q^h) nur mit $\eta \in K_m^h$ getestet, so dass gilt

$$(\text{const}, \eta - U^n)_h = \text{const}\,((1, \eta)_h - (1, U^n)_h)) = \text{const}(m - m) = 0.$$

Bemerkung 4.3.1 *Jede Lösung von (P^h) ist eine Lösung von (Q^h).* ■

4.4 Existenz und Eindeutigkeit von Lösungen

Im Folgenden wird die Existenz und Eindeutigkeit von Lösungen von (P^h) und (Q^h) jeweils für die Teilprobleme in einem Zeitschritt gezeigt und die entsprechende Eigenschaft für den vorherigen als gegeben angenommen. Dies liefert induktiv die Existenz und Eindeutigkeit der gesamten Lösung. Der Beweis der Existenz und Eindeutigkeit ist in [BE92] für die voll-implizite Diskretisierung ausgeführt. Ferner ist angemerkt, dass er mit kleinen Anpassungen auch für die semi-implizite Diskretisierung verwendbar ist. Diese angepasste Version des Beweises wird hier gegeben.

Satz 4.4.1 *Das Problem (Q^h) besitzt eine eindeutig bestimmte Lösung.*

BEWEIS. Zunächst stellen wir das Teilproblem zu einem Zeitpunkt t^n als Minimierungsproblem dar. Dazu definieren wir für $n = 1, ..., M$ auf S_m^h das Energiefunktional

$$J_n^h(\eta) = \frac{1}{2}\left(\gamma|\eta|_1^2 + \frac{1}{\tau}|\mathcal{G}^h(\eta - U^{n-1})|_1^2\right) - (U^{n-1}, \eta)_h + |\Omega|.$$

Da U^{n-1} entweder der Startwert U^0 oder die Lösung aus dem vorherigen Zeitschritt ist, gilt $U^{n-1} \in K_m^h$ und $(\eta - U^{n-1}, 1)_h = 0$. Also ist J_n^h auf S_m^h wohldefiniert. Nun betrachten wir das restringierte Minimierungsproblem

$$U^n \in K_m^h : \qquad J_n^h(U^n) \leq J_n^h(\eta) \qquad \forall \eta \in K_m^h. \tag{4.4.1}$$

Offensichtlich ist J_n^h konvex und stetig. Wegen $|U^{n-1}(x)| \leq 1$ und $|\eta(x)| \leq 1$ gilt nach Eigenschaft (6) von $(\cdot,\cdot)_h$ $|\Omega| - (U^{n-1}, \eta)_h = (1 - U^{n-1}\eta, 1)_h \geq 0$. Somit gilt ferner $J_n^h(\eta) \geq \frac{\gamma}{2}|\eta|_1^2$. Da für $\eta \in S_m^h$ die Masse $(\eta, 1)_h$ festgelegt ist, ist J_n^h nach der diskreten Poincaré-Ungleichung in S_m^h koerziv, es gilt also

$$|J_n^h(\eta)| \to \infty \qquad \text{für} \qquad \|\eta\|_h \to \infty$$

in S_m^h. Damit besitzt das Minimierungsproblem nach Satz 2.2.3 eine Lösung. Nach Korollar 2.2.1 ist das Minimierungsproblem äquivalent zu

$$0 \leq DJ_n^h(U^n)(\eta - U^n) \qquad \forall \eta \in K_m^h, \tag{4.4.2}$$

und damit zu Gleichung (4.3.4) in (Q^h). Nun zeigen wir, dass die Lösung eindeutig bestimmt ist. Dazu nehmen wir an, dass U_1^n und U_2^n zwei Lösungen sind und setzen $\tilde{U}^n = U_1^n - U_2^n$. Testen von Gleichung (4.3.4) in (Q^h) für U_1^n mit U_2^n und für U_2^n mit U_1^n ergibt dann

$$\gamma(\nabla U_1^n, -\tilde{U}^n) + \frac{1}{\tau}(\mathcal{G}^h(U_1^n - U^{n-1}), -\tilde{U}^n)_h \geq (U^{n-1}, -\tilde{U}^n)_h,$$
$$\gamma(\nabla U_2^n, \tilde{U}^n) + \frac{1}{\tau}(\mathcal{G}^h(U_2^n - U^{n-1}), \tilde{U}^n)_h \geq (U^{n-1}, \tilde{U}^n)_h.$$

Wegen $(U_1^n, 1)_h = (U_2^n, 1)_h = (U^{n-1})_h = m$ gilt nun

$$\mathcal{G}^h(U_2^n - U^{n-1}) - \mathcal{G}^h(U_1^n - U^{n-1}) = \mathcal{G}^h(U_2^n - U_1^n - U^{n-1} + U^{n-1}) = -\mathcal{G}^h(\tilde{U}^n).$$

Damit folgt nach Summieren der beiden Ungleichungen und Multiplikation mit -1

$$\gamma|\tilde{U}^n|_1^2 + \frac{1}{\tau}|\mathcal{G}^h\tilde{U}^n|_1^2 = \gamma|\tilde{U}^n|_1^2 + \frac{1}{\tau}(\mathcal{G}^h(\tilde{U}^n), \tilde{U}^n) \leq (U^{n-1} - U^{n-1}, \tilde{U}^n)_h = 0, \qquad (4.4.3)$$

womit insbesondere $|\tilde{U}^n|_1 = 0$ gegeben ist, was mit $(\tilde{U}^n, 1)_h = 0$ nach der diskreten Poincaré-Ungleichung zu $\tilde{U}^n = 0$ führt. ■

Bemerkung 4.4.1 *Für die hier untersuchte semi-implizite Diskretisierung taucht auf der rechten Seite von Ungleichung* (4.4.3) *der Term* $(U^{n-1} - U^{n-1}, \tilde{U}^n)_h = 0$ *auf. Da induktiv angenommen wird, dass für die Lösung im vorherigen Zeitschritt bereits Eindeutigkeit gezeigt ist, ist dieser Term gleich* 0*. Im Fall der in [BE92] auch untersuchten voll-impliziten Diskretisierung taucht statt dessen der Term* $\|\tilde{U}^n\|_h^2$ *auf. Dafür wird die Abschätzung*

$$\|\tilde{U}^n\|_h^2 = (\nabla\mathcal{G}^h\tilde{U}^n, \nabla\tilde{U}^n) \leq \frac{1}{\tau}\|\tilde{U}^n\|_{-h}^2 + \frac{\tau}{4}|\tilde{U}^n|_1^2$$

verwendet. Zusammen mit dem Analogon von (4.4.3) *ergibt dies*

$$\left(\gamma - \frac{\tau}{4}\right)|\tilde{U}^n|_1^2 \leq 0.$$

In diesem Fall ist die Eindeutigkeit nur für $\tau < 4\gamma$ *gegeben.*

Wie die Konstruktion der Probleme nahelegt, gilt für Lösungen $(U_P^n, W_P^n)^T$ von (P^h) und (U_Q^n) von (Q^h) der Zusammenhang

$$\begin{aligned} U_P^n &= U_Q^n, \\ W_P^n &= -\mathcal{G}^h\left(\frac{U_Q^n - U_Q^{n-1}}{\tau}\right) + \lambda, \end{aligned}$$

mit $\lambda = (W_P^n, 1)_h|\Omega|^{-1}$. Nach Bemerkung 4.3.1 löst U_P^n auch (Q^h). Um nun aus der Lösung von (Q^h) eine Lösung für (P^h) zu konstruieren, „fehlt“ eben diese Konstante λ. Im folgenden Existenzbeweis wird eine Funktion konstruiert, die zu gegebenem λ, als „Vorschlag“ für die Masse von W^n, die Masse von der Lösung U^n des zu λ gehörenden Problems liefert. Aus dem Zwischenwertsatz folgt dann die Existenz eines λ, für das U^n die richtige Masse $m = (U^0, 1)_h$ besitzt.

Satz 4.4.2 *Das Problem (P^h) besitzt eine Lösung.*

BEWEIS. Es sei U^n die Lösung von (Q^h) für den entsprechenden Zeitschritt. Nun betrachten wir das Problem:

(PQ$_\mu^h$) Finde $U_\mu^n \in K^h$ mit

$$\gamma(\nabla U_\mu^n, \nabla\eta - \nabla U_\mu^n) + (U_\mu^n, \eta - U_\mu^n)_h \geq (f + \mu, \eta - U_\mu^n)_h \qquad \forall\eta \in K^h.$$

Dabei sei $\mu \in [\mu_L, \mu_R]$ gegeben, wobei gelte

$$\begin{aligned} f &:= U^n + U^{n-1} - \mathcal{G}^h(\frac{U^n - U^{n-1}}{\tau}), \\ \mu_L &:= -1 - \max_{x\in\Omega} f(x), \\ \mu_R &:= 1 - \min_{x\in\Omega} f(x). \end{aligned}$$

Da die Bilinearform in (PQ$_\mu^h$) symmetrisch, stetig und koerziv ist, existiert nach Satz 2.2.3 für jedes μ eine eindeutig bestimmte Lösung von (PQ$_\mu^h$). Gilt nun $(U_\mu^n, 1)_h = m = (U^n, 1)_h$ für ein μ, so können wir die Ungleichung in (Q^h) mit U_μ^n und jene in (PQ$_\mu^h$) mit U^n testen, beide addieren, mit -1 multiplizieren und erhalten

$$\gamma |U_\mu^n - U^n|_1^2 + \|U_\mu^n - U^n\|_h^2 \leq 0$$

und somit $U_\mu^n = U^n$. In diesem Fall reproduziert Testen von (PQ$_\mu^h$) mit $\eta \in K_m^h$ die Ungleichung aus (Q^h). Setzen wir dann $W^n = \mu - \mathcal{G}^h(\frac{U^n - U^{n-1}}{\tau})$, so reproduziert (PQ$_\mu^h$) die Ungleichung aus (P^h) und $(U^n, W^n)^T$ ist Lösung von (P^h). Ziel ist es also, ein μ zu finden, für dass $(U_\mu^n, 1)_h = m$ gilt. Dazu definieren wir die Abbildung

$$M^h : [\mu_L; \mu_R] \to \mathbb{R}, \quad \mu \mapsto M^h(\mu) = (U_\mu^n, 1)_h,$$

wobei U_μ^n die Lösung von (PQ$_\mu^h$) für dieses μ ist. Aus der Definition von μ_L und μ_R folgt

$$-1 \geq f + \mu_L \qquad \text{und} \qquad 1 \leq f + \mu_R.$$

Für $\eta \in K^h$ folgt nach Multiplikation der ersten Ungleichung mit $(\eta(x) - (-1)) \geq 0$ und Multiplikation der zweiten mit $(\eta(x) - 1) \leq 0$ aus den Eigenschaften (4) und (6) von $(\cdot, \cdot)_h$

$$\begin{aligned} (-1, \eta - (-1))_h &\geq (f + \mu_L, \eta - (-1))_h, \\ (1, \eta - 1)_h &\geq (f + \mu_R, \eta - 1)_h. \end{aligned}$$

Dies ist die Ungleichung von (PQ$_\mu^h$) für $U_{\mu_L}^n = -1$ und $U_{\mu_R}^n = 1$. Somit gilt $M^h(\mu_L) = -|\Omega|$ und $M^h(\mu_R) = |\Omega|$. Wir zeigen nun noch, dass M^h monoton und stetig ist. Dann existiert nach dem Zwischenwertsatz (siehe etwa [Rud64]) ein $\mu \in [\mu_L; \mu_R]$ mit

$$M^h(\mu) = m \in [M^h(\mu_L); M^h(\mu_R)] = [-|\Omega|; |\Omega|]$$

und die Existenz einer Lösung ist gezeigt. Zum Beweis der Stetigkeit und Monotonie seien $\mu_1, \mu_2 \in [\mu_K; \mu_R]$ beliebig. Dann liefert das Testen von (PQ$_{\mu_1}^h$) mit $U_{\mu_2}^n$, von (PQ$_{\mu_2}^h$) mit $U_{\mu_1}^n$, Addition der Ungleichungen und Multiplikation mit -1

$$\begin{aligned} 0 \leq \gamma |U_{\mu_1}^n - U_{\mu_2}^n|_1^2 + \|U_{\mu_1}^n - U_{\mu_2}^n\|_h^2 &\leq (\mu_1 - \mu_2, U_{\mu_1}^n - U_{\mu_2}^n)_h \\ &= (\mu_1 - \mu_2)(M^h(\mu_1) - M^h(\mu_2)). \end{aligned}$$

Daraus folgt zum einen die Monotonie von M^h. Zum anderen liefert die Cauchy-Schwarz-Ungleichung wegen $\|1\|_h = |\Omega|^{\frac{1}{2}}$

$$\begin{aligned}|M^h(\mu_1) - M^h(\mu_2)|^2 = (1, U^n_{\mu_1} - U^n_{\mu_2})^2_h &\leq \|1\|^2_h \|U^n_{\mu_1} - U^n_{\mu_2}\|^2_h \\ &\leq |\Omega|(\mu_1 - \mu_2)(M^h(\mu_1) - M^h(\mu_2)),\end{aligned}$$

was insbesondere $|M^h(\mu_1) - M^h(\mu_2)| \leq |\Omega||\mu_1 - \mu_2|$, also die Lipschitz-Stetigkeit von M^h impliziert. ■

Satz 4.4.3 *Die Lösung von (P^h) ist eindeutig bestimmt.*

BEWEIS. Da für eine Lösung $(U^n, W^n)^T$ von (P^h)

$$W^n = -\mathcal{G}^h(\frac{U^n - U^{n-1}}{\tau}) + \text{const}$$

gilt, zeigt das Testen mit $\eta \in K^h_m$, dass U^n die eindeutige Lösung von (Q^h) ist. Es ist noch Eindeutigkeit von W^n zu zeigen. Seien dazu W^n_1 und W^n_2 zwei Lösungen für W^n. Dann zeigt Subtraktion von Gleichung (4.3.1) in (P^h) für W^n_1 und W^n_2 und Testen mit $W^n_1 - W^n_2$, dass $|W^n_1 - W^n_2|_1 = 0$ gilt. Also sind W^n_1 und W^n_2 bis auf die Addition von Konstanten identisch. Da $|(U^n, 1)_h| = |m| < |\Omega|$ gefordert war, existiert mindestens ein $x_i \in \mathcal{N}^h$ mit $|U^n(x_i)| < 1$. Sei nun $\delta \in \mathbb{R}$ jeweils so gewählt, dass $\eta_\pm = U^n \pm \delta\phi_i \in K^h$ gilt. Testen von (4.3.2) mit η_+ und η_- liefert dann

$$(W^n_1, \phi_i)_h = \gamma(\nabla U^n, \nabla\phi_i) - (U^{n-1}, \phi_i)_h = (W^n_2, \phi_i)_h,$$

also $W^n_1(x_i) = W^n_2(x_i)$, woraus die Eindeutigkeit von W^n folgt. ■

4.5 Konvergenz

In diesem Abschnitt wird ein Überblick über die Konvergenzresultate, die Blowey und Elliot in [BE92] für die semi-implizite Diskretisierung zeigen, gegeben. Zunächst werden dort Stabilitätsabschätzungen bewiesen. Dann wird eine Fehlerfunktion definiert und unter Zuhilfenahme der Stabilitätsabschätzungen eine Abschätzung des Fehlers hergeleitet. Da die Beweise sehr technisch sind, werden nur die wesentlichen Resultate zitiert. Für den Beweis der Fehlerabschätzung selbst wird eine sehr kurze Skizze gegeben.

Eine Folgerung der Stabilitätsabschätzungen ist:

Korollar 4.5.1 *Es gelte $\tau < 4\gamma$. Dann gilt für die Lösung von (P^h) mit von h und τ unabhängigen Konstanten C*

$$\max_{n=0,\ldots,M} \|U^n\|^2_1 + \sum_{n=1}^{M} \tau \left\| \frac{U^n - U^{n-1}}{\tau} \right\|_{-h^2} + \sum_{n=1}^{M} \|U^n - U^{n-1}\|^2_1 \leq C,$$

$$|(W^n,1)_h| \le C + \frac{C}{\tau^{\frac{1}{2}}},$$
$$\sum_{n=1}^{M} \tau \|W^n\|_1^2 \le C.$$

BEWEIS. Siehe [BE92]. ■

Die Lösung des kontinuierlichen Problems (P) liegt in $W(0,T;H^1(\Omega)) \times L^2(0,T;H^1(\Omega))$ und damit zu fast allen Zeitpunkten im Sobolevraum $H^1(\Omega)$. Während die Lösung von (P^h) im Ort in S^h, also einem Unterraum von $H^1(\Omega)$, liegt, ist sie in der Zeit nur zu bestimmten Zeitpunkten definiert. Um die Lösung von (P^h) mit der von (P) vergleichen zu können, wird nun U^n stückweise konstant in der Zeit fortgesetzt. Setze also

$$U_{h,\tau}(\cdot,t) := \begin{cases} U^n & \text{für } t \in J^n = (t^{n-1},t^n] \\ U^0 & \text{für } t = 0 \end{cases} .$$

Die Fehlerfunktion ist definiert durch

$$e(x,t) := u(x,t) - U_{h,\tau}(x,t).$$

Nun ist folgende Fehlerabschätzung gegeben:

Satz 4.5.1 *Es gelte $\tau < 2\gamma$. Sei $(U^n,W^n)^T$ die Lösung von (P^h) und $(u,w)^T$ die Lösung von (P). Dann gilt für $U^0 = P^h u_0 \in K^h$ die Fehlerabschätzung*

$$\|e\|^2_{L^\infty(0,T;(H^1(\Omega)'))} + \|e\|^2_{L^2(0,T;H^1(\Omega))} \le C\left(\frac{h^4}{\tau} + h^2 + \tau\right) =: \sigma(h,\tau) \qquad (4.5.1)$$

mit von h und t unabhängiger Konstante C.

Der vollständige Beweis ist in [BE92] für die voll-implizite Diskretisierung ausgeführt. Für die semi-implizite Diskretisierung sind Hinweise gegeben, wie er anzupassen ist. Hier wird nur eine sehr kurze Beweisskizze gegeben, um zu verdeutlichen, an welchen Stellen sich der Beweis von dem für die voll-implizite Diskretisierung unterscheidet und warum er sich nicht analog zum Eindeutigkeitsbeweis ohne die Beschränkung $\tau < 2\gamma$ führen lässt.

Für den Beweis werden für in der Zeit stetige,integrierbare Funktionen folgende Notationen verwendet:

$$\eta^n(\cdot) := \eta(\cdot,t^n),$$
$$\overline{\eta}^n(\cdot) := \frac{1}{\tau}\int_{J^n} \eta(\cdot,s)ds,$$
$$\overline{\eta}^0(\cdot) := \eta^0(\cdot).$$

Ferner werden Mittelwerte über Ω betrachtet

$$U_m^n := U^n - \frac{m}{|\Omega|},$$
$$u_m^n := u^n - \frac{m}{|\Omega|}.$$

Der Beweis verwendet mehrfach die Regularität $u \in L^2(0,T;H^2(\Omega))$. Dies ist für hinreichend glatten Rand von Ω gegeben. Liegt diese Regularität jedoch nicht vor, ist die Fehlerabschätzung nicht anwendbar.

BEWEIS. (Skizze)
Essentiell ist, dass sowohl für u als auch für U^n die Massenerhaltung

$$(U^n,1) = (U^n,1)_h = (U^0,1)_h = (P^h u_0,1)_h = (u_0,1) = (u,1) = m$$

gegeben ist. Der gesamte Beweis basiert auf der Betrachtung des Terms

$$\begin{aligned}(A) := &\int_{J^n} \left\langle \frac{du}{dt} - \frac{U^n - U^{n-1}}{\tau}, \mathcal{G}(u-U^n) \right\rangle dt \\ &+ \int_{J^n} \gamma(\nabla u - \nabla U^n, \nabla u - \nabla U^n) dt - \int_{J^n} (u-U^n, u-U^n) dt.\end{aligned}$$

Zunächst wird unabhängig davon, ob u und U^n Lösungen von (P) und (P^h) sind, die Identität

$$\begin{aligned}\sum_{n=1}^k (A) = &\left[\frac{1}{2}\|e(t^k)\|_{-1}^2 + \gamma \int_0^{t^k} |e(t)|_1^2 dt - \int_0^{t^k} \|e(t)\|_0^2 dt\right] - \frac{1}{2}\|e(o)\|_{-1}^2 \\ &+ \left[\frac{1}{2}\sum_{n=1}^k \|U^n - U^{n-1}\|_{-1}^2 - \sum_{n=1}^k (U^n - U^{n-1}, \mathcal{G}(\overline{u}^n) - u^{n-1}))\right] \\ =:&(I) - (II) + (III)\end{aligned} \tag{4.5.2}$$

gezeigt. Nun wird (Q) mit U^n und die Ungleichung in (P^h) mit $I^h\overline{u}^n$ getestet. Damit die Interpolation I^h auf $\overline{u}^n$ angewandt werden kann, ist $u(t) \in C(\overline{\Omega})$ notwendig. Dies ist normalerweise für $u(t) \in H^1(\Omega)$ nur bei $d = 1$ gegeben. Hier wird also bereits $u(t) \in H^2(\Omega)$ benötigt. Werden nun die erhaltenen Ungleichungen in (A) eingesetzt, so ergibt sich

$$\begin{aligned}(A) \leq &\tau \left[(\nabla U^n, \nabla I^h\overline{u}^n - \nabla\overline{u}^n) - \left(\frac{U^n - U^{n-1}}{\tau}, \mathcal{G}(\overline{u}^n - U^n)\right) + (W^n, U^n - I^h\overline{u}^n)_h\right] \\ &+ \tau\left[-(U^n, I^h\overline{u}^n - \overline{u}^n) + (U^n, I^h\overline{u}^n - U^n) - (B)\right].\end{aligned}$$

Der Term (B) ist dabei im Beweis für die voll-implizite Diskretisierung gegeben durch

$$(B) = (U^n, I^h\overline{u}^n - U^n)^h.$$

Für die semi-implizite Diskretisierung geht hier U^{n-1} statt U^n ein. Der von Blowey und Elliot gegebene Hinweis sieht folgende Anpassung vor:

$$(B) = (U^{n-1}, U^n - I^h\overline{u}^n) = (U^n, I^h\overline{u}^n - U^n)^h + (U^{n-1} - U^n, U^n - I^h\overline{u}^n)^h.$$

Jetzt wird im Beweis noch der Term $(W^n, U^n - I^h\overline{u}^n)$ umgeformt und die Abschätzung für (A) über $n = 1, ..., k$ summiert. Dies führt auf eine Abschätzung der Form

$$\sum_{n=1}^{k}(A) \leq (IV) + (V) + (VI) + (VII) + (VIII) + (IX) + (X) + (IX), \tag{4.5.3}$$

wobei der letzte Term

$$(IX) = \tau \sum_{n=1}^{k}(U^{n-1} - U^n, U^n - I^h\overline{u}^n)^h$$

für die semi-implizite Diskretisierung zusätzlich auftaucht. Die Terme (II), $(IV) - (X)$ werden unter Verwendung von Regularitätsergebnissen, Stabilitätsabschätzungen und Interpolationsabschätzungen und der Abschätzung für die numerische Integration in $(\cdot, \cdot)_h$ nach oben , der Term (III) nach unten abgeschätzt. Dies führt für (II), (III), (V), (VI), $(VIII)$-(X) zu Abschätzungen durch $\sigma(h, \tau)$. Für (XI) kann mit denselben Methoden eine ebensolche Abschätzung gezeigt werden. Des Weiteren werden Abschätzungen

$$(IV) \leq \frac{Ch^2}{\nu} + \frac{\gamma\nu}{2}\int_0^{t^k} |e(t)|_1^2 dt, \qquad (VII) \leq \frac{Ch^4}{\gamma\nu\tau} + \frac{\gamma\nu}{2}\int_0^{t^k} |e(t)|_1^2 dt.$$

für $\nu > 0$ gezeigt. Dies führt zusammen mit (4.5.2) und (4.5.3) zu

$$(I) \leq \sigma(h, \tau) + \nu\gamma \int_0^{t^k} |e(t)|_1^2 dt. \tag{4.5.4}$$

Die Anwendung der Ungleichung

$$\|e(t)\|_0^2 = (\nabla\mathcal{G}(e(t)), \nabla e(t)) \leq \rho\gamma|e(t)|_1^2 + \frac{1}{4\gamma\rho}\|e(t)\|_{-1}^2$$

für $\rho > 0$ auf (4.5.4) ergibt nun

$$\frac{1}{2}\|e(t^t)\|_{-1}^2 + \gamma(1 - \rho - \nu)\int_{J^n} |e(t)|_1^2 dt \leq \frac{1}{4\gamma\nu}\sum_{n=1}^{k}\int_{J^n} \|e(t)\|_{-1}^2 dt + \sigma(h, \tau).$$

Ab dieser Stelle funktioniert der Beweis exakt so wie für die voll-implizite Diskretisierung. Eigentlich ist man an einer Abschätzung von $\|e(t^t)\|_{-1}$ nach oben interessiert. Das Integral auf der rechten Seite steht dieser im Weg. Allerdings hat der Term (IX) scheinbar keinen direkten

Bezug zu den Termen (IV) und (VII). In Bemerkung 4.4.1 ist erörtert, wie im Eindeutigkeitsbeweis für die semi-implizite Diskretisierung ein Term auf der „falschen“ Seite der Abschätzung durch den induktiven Beweisansatz gleich Null wird. Für den voll-impliziten Fall erfordert dies die Bedingung $\tau < 4\gamma$. Da hier kein vergleichbarer induktiver Ansatz gegeben ist, lässt sich ein ähnliches Argument nicht anwenden. Durch die semi-implizite Diskretisierung ist ein zusätzlicher Term zu behandeln. Dieser ist zwar „gutartig“, also durch $\sigma(h, \tau)$ abschätzbar, führt aber nicht zum Verschwinden der störenden Terme. Dies bedeutet insbesondere, dass sich das Resultat des Satzes nicht mit einem analogen Argument wie im Eindeutigkeitsbeweis ohne die Bedingung $\tau < 2\gamma$ zeigen lässt.

Im weiteren Verlauf des Beweises wird zur Abschätzung dieser störenden Terme die Bedingung $\tau < 2\gamma$ notwendig. Eine Anwendung der Gronwall Ungleichung führt dann zu einer Abschätzung

$$\|e(t^k)\|_{-1}^2 \leq \sigma(h, \tau),$$

aus der schließlich die gesuchten Abschätzungen hergeleitet werden. ■

5 Lösung der diskreten Probleme

Nun werden verschiedene Löser für die bei der semi-impliziten Diskretisierung der Cahn-Hilliard Gleichung in jedem Zeitschritt entstehenden diskreten, nichtlinearen Gleichungssysteme betrachtet. Allen Lösern gemein ist, dass sie die Probleme aus einem einzelnen Zeitschritt lösen und für alle Zeitschritte nacheinander angewandt werden. Der Übersicht halber wird deshalb in diesem Kapitel eine neue, etwas einfachere Notation für diese Probleme eingeführt, die sich von der vorherigen unterscheidet. u ist jetzt die Lösung des diskreten Problems in einem Zeitschritt und nicht mehr die des kontinuierlichen Problems. In jedem Zeitschritt ist ein Problem der folgenden Form zu lösen:

(CH) Finde $(u, w)^T \in K^h \times S^h$ mit

$$(u, v)_h + \tau(\nabla w, \nabla v) = (u^{\text{alt}}, u)_h \qquad \forall v \in S^h, \tag{5.0.1}$$

$$\gamma(\nabla u, \nabla v - \nabla u) - (w, v - u) \geq (u^{\text{alt}}, v - u)_h \qquad \forall v \in K^h. \tag{5.0.2}$$

Dabei wird stets $(u^{\text{alt}}, w^{\text{alt}})^T \in K^h \times S^h$ mit $|(u^{alt}, 1)_h| < |\Omega|$ als gegeben vorausgesetzt.

Im Folgenden wird für das approximierte L^2-Skalarprodukt stets das so genannte „gelumpte“ L^2-Skalarprodukt

$$(u, v)_h := \int_\Omega I_h(uv)(x)dx$$

verwendet, wobei $I_h(v) \in S^h$ die Projektion von v in S^h ist, die $I_h(v)(x_i) = v(x_i)$ für alle Knoten x_i der Triangulierung $\mathcal{T}^h$ erfüllt. In diesem Fall gilt $(\phi_i, \phi_j)_h = 0$ für $i \neq j$.

5.1 Die algebraische Darstellung der diskreten Cahn-Hilliard Gleichung

Soll ein diskretes Problem mit Hilfe eines Rechners gelöst werden, so wird im Prinzip immer eine algebraische Formulierung des jeweiligen Algorithmus implementiert, die ein Problem in $\mathbb{R}^N$ und nicht in einem Funktionenraum löst. Zwar sind beide Formulierungen äquivalent, doch ist eine algebraische Darstellung des Algorithmus näher an den Details der Implementierung. Deshalb werden die Algorithmen zunächst in einer jeweils „passenden“ Form präsentiert und dann in ihrer algebraischen Form dargestellt. In diesem Abschnitt wird durch Einsetzen der Knotenbasis $\phi_1, ..., \phi_N$ des Finite-Element-Raums S^h in die Bilinearformen und Funktionale auf S^h die algebraische Formulierung von (CH) gegeben.

Das Einsetzen der Knotenbasis liefert

$$\begin{aligned}
M &:= (M_{ij})_{i,j=1,...,N}, & M_{ij} &:= (\phi_i, \phi_j)_h, & (x, y)_h &= \underline{y}^T M \underline{x}, \\
A &:= (A_{ij})_{i,j=1,...,N}, & A_{ij} &:= (\nabla\phi_i, \nabla\phi_j), & (\nabla x, \nabla y) &= \underline{y}^T A \underline{x}, \\
b &:= (b_i)_{i=1,...,N}, & b_i &:= (u^{\text{alt}}, \phi_i)_h, & (u^{\text{alt}}, x)_h &= \underline{x}^T b
\end{aligned}$$

für $i,j = 1,...,N$ und alle $x,y \in S^h$. Dabei ist $\underline{x} \in \mathbb{R}^n$ der Koeffizientenvektor zu $x \in S^h$, es gilt also $x = \sum_{i=1}^N \underline{x}_i \phi_i$. Um die Eigenschaft $(x,1)_h$ algebraisch darzustellen, wird ferner der Vektor $(1,...,1) = \underline{1}$ als Koeffizientenvektor der konstanten 1-Funktion definiert. Damit gilt offensichtlich $(x,1)_h = \underline{x}^T M \underline{1}$. Nun sind noch die Funktionenräume algebraisch darzustellen. Da die Funktionen in S^h stückweise linear sind, kann man die Zugehörigkeit zur konvexen Menge K^h an den Knoten testen, was zu

$$\begin{aligned}
\underline{S}^h &:= \mathbb{R}^N, \\
\underline{S}_0^h &:= \{\underline{x} \in \underline{S}^h : (x,1)_h = \underline{x}^T M \underline{1} = 0\}, \\
\underline{K}^h &:= \{\underline{x} \in \underline{S}^h : |\underline{x}_i| \leq 1, \text{ für } i = 1,...,N\},
\end{aligned}$$

führt. Für diese Darstellung gilt $x \in S^h \Leftrightarrow \underline{x} \in \underline{S}^h$ und $x \in K^h \Leftrightarrow \underline{x} \in \underline{K}^h$. Damit ergibt sich für (CH) die Darstellung :

(CH) Finde $(\underline{u},\underline{w})^T \in \underline{K}^h \times \underline{S}^h$ mit

$$(\underline{y}-\underline{u})^T \gamma A \underline{u} - (\underline{y}-\underline{u}) M \underline{w} \geq (\underline{y}-\underline{u}) b \qquad \forall \underline{y} \in \underline{K}^h \tag{5.1.1}$$

$$M\underline{u} + \tau A \underline{w} = b. \tag{5.1.2}$$

Dies lässt sich auch als eine große Ungleichung darstellen

$$\begin{pmatrix} \underline{y}-\underline{u} \\ \underline{x}-\underline{w} \end{pmatrix}^T \begin{pmatrix} \gamma A & -M \\ M & \tau A \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{w} \end{pmatrix} \geq \begin{pmatrix} \underline{y}-\underline{u} \\ \underline{x}-\underline{w} \end{pmatrix}^T \begin{pmatrix} b \\ b \end{pmatrix} \qquad \forall (\underline{y},\underline{x})^T \in \underline{K}^h \times \underline{S}^h.$$

Durch Einsetzen der Testfunktionen $(\underline{y},\underline{w})^T, (\underline{u},\underline{w}+\underline{\phi}_i)^T = (\underline{u},\underline{w}+e_i)^T \in \underline{K}^h \times \underline{S}^h$ können die einzelnen Gleichungen reproduziert werden. Im Allgemeinen ist man daran interessiert, derartige Variationsprobleme durch schnelle Mehrgitterverfahren zu lösen. Die Konvergenzgeschwindigkeit ist bei diesen teilweise unabhängig von der Systemgröße. Für Variationsungleichungen, die von Hindernisproblemen stammen, existieren die monotonen Mehrgitterverfahren (siehe [Kor97]). Sie basieren auf der monotonen Verkleinerung eines Energiefunktionals. Ziel ist es, dessen Minimum in einer konvexen Menge zu finden. Allerdings setzt dies eine Formulierung der Variationsungleichung als Minimierungsproblem voraus, die im Allgemeinen nur mit symmetrischer, positiv definiter Systemmatrix gegeben ist. Zunächst ist die vorliegende Systemmatrix nicht einmal symmetrisch, was sich durch Multiplikation der Gleichung mit -1 ändern lässt und zur symmetrischen Systemmatrix

$$\mathbf{A} = \begin{pmatrix} \epsilon A & -M \\ -M & -\tau A \end{pmatrix}$$

führt. Eine Multiplikation der Ungleichung ist natürlich nicht möglich. Allerdings ist $\mathbf{A}$ nicht positiv definit, wie das Beispiel

$$\begin{pmatrix} 0 \\ e_i \end{pmatrix}^T \mathbf{A} \begin{pmatrix} 0 \\ e_i \end{pmatrix} = -\tau |\phi_i|_1^2 < 0$$

zeigt. Bisher sind keine Löser für dieses Problem bekannt, die mit zufriedenstellender Geschwindigkeit arbeiten. Der erste hier präsentierte Algorithmus, das Operator-Splitting von Lions und Mercier, basiert auf einer Formulierung als Differential-Inklusion. Die nichtlineare Block-Gauß-Seidel Iteration nutzt die algebraische Darstellung von (CH). Da sie nicht auf einem Minimierungsproblem beruht, ist eine Erweiterung zu einem monotonen Mehrgitterverfahren nicht ohne weiteres möglich. Um solche Verfahren anwenden zu können, soll die Variationsungleichung zu einem Minimierungsproblem umformuliert werden. Dies hat allerdings den „Preis“ eine zusätzliche Nebenbedingung stellen zu müssen. Auf das dabei entstehende Sattelpunktproblem wird der Uzawa Algorithmus angewandt.

5.2 Der Splitting-Algorithmus von Lions-Mercier

Blowey und Elliot präsentieren in [BE92] einen Algorithmus zur Lösung der diskreten Probleme und weisen darauf hin, dass sie nach Fertigstellung der Arbeit noch einen überlegenen Algorithmus, basierend auf dem Operator-Splitting von Lions-Mercier, entdeckt haben. Der Splitting Algorithmus von Lions und Mercier (siehe [LM79]) dient der Lösung von Operatorgleichungen der Form

$$0 \in C(u) = A(u) + B(u), \tag{5.2.1}$$

wobei C ein mengenwertiger Operator auf einem Hilbertraum H ist, dessen Werte Teilmengen von H sind und der sich als Summe zweier maximal monotoner, mengenwertiger Operatoren A und B darstellen lässt. Der Algorithmus basiert auf folgendem von Peaceman und Rachford eingeführten Iterationsverfahren für lineare Operatoren mit Parameter $\lambda > 0$:

$$u^{n+1} = (I + \lambda B)^{-1}(I - \lambda A)(I + \lambda A)^{-1}(I - \lambda B)u^n. \tag{5.2.2}$$

Die genaue Bedeutung dieser Iteration für mengenwertige Operatoren ist zunächst unklar, da die einzelnen Operationen normalerweise mengenwertig sind. Deshalb wird in [LM79] eine Iteration verwendet, die für einwertige Operatoren äquivalent zu (5.2.2) ist und bei den mengenwertigen Operationen in (5.2.2) jeweils der Auswahl eines Elementes des jeweiligen Bildes entspricht.

5.2.1 Der Lions-Mercier Algorithmus für maximal monotone Operatoren

Zur Formulierung der Iteration wird für maximal monotone Operatoren A zunächst

$$J_A^\lambda := (I + \lambda A)^{-1}$$

definiert. Da Hilberträume strikt konvex sind, ist diese Abbildung nach Satz 2.3.2 einwertig. Falls auch A einwertig ist, gilt offensichtlich $(I + \lambda A)J_A^\lambda = I$. Im mengenwertigen Fall gilt für

$u \in H$ und $v \in (I + \lambda A)u$ noch

$$\begin{aligned}(2J_A^\lambda - I)v &= 2J_A^\lambda v - v \\ &= 2u - v \\ &\in 2u - (I + \lambda A)u \\ &= (I - \lambda A)u = (I - \lambda A)(I + \lambda A)^{-1}v.\end{aligned}$$

Dies legt nahe, zur Spezifizierung von (5.2.2) zu u^n jeweils ein Element $v^n \in (I + \lambda B)u^n$ und dafür folgende Iteration zu betrachten

$$v^{n+1} = (2J_A^\lambda - I)(2J_B^\lambda - I)v^n.$$

v^{n+1} ist nun auch für mengenwertige maximal monotone A und B wohldefiniert. Für den Fall einwertiger Operatoren ist dies äquivalent zu (5.2.2). Für mengenwertige Operatoren ist lediglich zu u^0 ein festes $b^0 \in Bu^0$ zu wählen und damit $v^0 := u^0 + \lambda b^0 \in (I + \lambda B)u^0$ festzulegen. Alle restlichen v^n sind dann durch die Iteration ebenso festgelegt. Da J_B^λ einwertig ist, ist mit v^n auch $u^n = J_B^\lambda v^n$ für alle n festgelegt. Für diese Iteration zeigen Lions und Mercier in [LM79] folgendes Konvergenzresultat :

Satz 5.2.1 *A und B seien maximal monoton und es existiere eine Lösung u von (5.2.1), also ein $u \in H$ und $a \in A(u)$, $b \in B(u)$ mit $a + b = 0$. Setze dann $v := u + \lambda b$ und $b^n := \frac{v^n - u^n}{\lambda}$, dann gilt*

$$\lim_{n\to\infty}(b^n - b, u^n - u) = 0.$$

BEWEIS. Siehe [LM79]. ■

Korollar 5.2.1 *Gilt zusätzlich für $x_1, x_2 \in H$ und $y_i \in B(x_i)$*

$$(y_1 - y_2, x_1 - x_2) \geq C\|x_1 - x_2\|_H^2$$

mit einem $C > 0$, so konvergiert der Algorithmus mit $\lim_{n\to\infty} C\|u^n - u\|_H^2 = 0$. ■

5.2.2 Die Cahn-Hilliard Gleichung als Operator-Inklusion

Nun wird das Problem (CH) als Operator-Gleichung der Form (5.2.1) dargestellt. Nach Kapitel 4 ist klar, dass (CH) äquivalent zu folgender Ungleichung ist:

(CH^1) Finde $u \in K_m^h$ mit

$$\gamma(\nabla u, \nabla v - \nabla u) + \frac{1}{\tau}(\mathcal{G}^h(u - u^{\text{alt}}), v - u)_h \geq (u^{\text{alt}}, v - u)_h \qquad \forall v \in K_m^h.$$

Betrachte nun folgende Abbildung $a : S_m^h \times S^h \to \mathbb{R}$

$$a(u, v) = \gamma(\nabla u, \nabla v) + \frac{1}{\tau}(\mathcal{G}^h(u - u^{\text{alt}}), v)_h - (u^{\text{alt}}, v)_h.$$

Offenkundig ist nun für ein festes $u \in S_m^h$ die Abbildung $a(u,\cdot) : S^h \to \mathbb{R}$ linear und, da S^h endlichdimensional ist, auch stetig. Nach dem Darstellungssatz von Frechét-Riesz existiert also zu jedem $u \in S_m^h$ ein $B_0(u) \in S^h$, so dass für $v \in S^h$

$$(B_0(u), v)_h = \gamma(\nabla u, \nabla v) + \frac{1}{\tau}(\mathcal{G}^h(u - u^{\text{alt}}), v)_h - (u^{\text{alt}}, v)_h$$

gilt. Ob hier das Skalarprodukt $(\cdot,\cdot)$ oder $(\cdot,\cdot)_h$ auf S^h gewählt wird, ist im Prinzip egal. Nur unterscheiden sich natürlich die Operatoren, die der Darstellunssatz liefert. Hier wurde das „gelumpte“ L^2-Skalarprodukt $(\cdot,\cdot)_h$ gewählt. Ferner wird $B_0(u) = 0$ für $u \in S^h \setminus S_m^h$ gesetzt. Nun ist die Ungleichung in (CH) äquivalent zu

$$(-B_0(u), v - u)_h + I_{K_m^h}(u) \le I_{K_m^h}(v) \qquad \forall v \in S^h,$$

und damit nach Definition des Subgradienten zu $-B_0(u) \in \partial I_{K_m^h}(u)$. Also ist (CH^1) äquivalent zu

(CH^2) Finde $u \in S^h$ mit

$$0 \in \partial I_{K_m^h}(u) + B_0(u).$$

Diese Aufspaltung für den Lions-Mercier Algorithmus zu verwenden, ist jedoch sehr ungünstig, da dann der Operator $(I + \lambda \partial I_{K_m^h})$ invertiert werden muss. Deshalb wird dieser Operator zunächst weiter aufgespalten. Offensichtlich gilt $I_{K_m^h} = I_{K^h} + I_{S_m^h}$. Da $|m| < |\Omega|$ vorausgesetzt wird, liegt die Funktion $\eta_m \equiv \frac{m}{|\Omega|} \in K_m^h = \operatorname{dom} I_{K^h} \cap \operatorname{dom} I_{S_m^h}$ nirgendwo auf dem Hindernis ± 1, ist also ein innerer Punkt von K^h. Damit ist I_{K^h} in einer Umgebung von η_m konstant, also auch stetig. Somit liefert Satz 2.2.1 für $v \in S^h$ $\partial I_{K_m^h}(v) = \partial I_{K^h}(v) + \partial I_{S_m^h}(v)$. Damit ist (CH^2) äquivalent zu:

(CH^3) Finde $u \in S^h$ mit

$$0 \in \underbrace{\partial I_{K^h}(u)}_{:=A(u)} + \underbrace{\partial I_{S_m^h}(u) + B_0(u)}_{:=B(u)}.$$

Soll nun auf diese Darstellung der Lions-Mercier Algorithmus angewandt werden, so ist zu zeigen, dass $A = \partial I_{K^h}$ und $B = B_0 + \partial I_{S_m^h}$ maximal monoton sind.

Satz 5.2.2 *Die Operatoren $A = \partial I_{K^h}$ und $B = B_0 + \partial I_{S_m^h}$ sind maximal monoton.*

BEWEIS. Da K^h abgeschlossen und konvex ist, ist I_{K^h} konvex, proper und unterhalbstetig. Ferner ist $B_0(u)$ für $u \in S_m^h$ die Gateaux-Ableitung der stetigen, konvexen Funktion

$$f_B(u) = \frac{1}{2}\left(\gamma |u|_1^2 + \frac{1}{\tau}|\mathcal{G}^h u - u^{\text{alt}}|_1^2\right) - (u^{\text{alt}}, u)_h.$$

Damit gilt nach Korollar 2.2.1 $B(u) = \partial(f_B + I_{S_m^h})(u)$, wobei $f_B + I_{S_m^h}$ konvex, proper und unterhalbstetig ist. Folglich sind nach Satz 2.3.1 A und B als Subdifferentiale konvexer, properer, unterhalbstetiger Funktionen maximal monoton. ■

Korollar 5.2.2 *Der Lions-Mercier Algorithmus konvergiert für das Problem (CH^3).*

BEWEIS. Die Anwendbarkeit des Algorithmus folgt aus Satz 5.2.2. Seien $x_1, x_2 \in S^h$ und $y_i \in B(x_i)$ also $y_i - B_0(x_i) \in \partial I_{S_m^h}(x_i)$ gegeben, dann gilt $x_i \in S_m^h$ und

$$(v - x_i, y_i)_h \leq (v - x_i, B_0(x_i))_h \qquad \forall v \in S_m^h.$$

Testen mit dem jeweils anderen x_i führt nach Addition und Multiplikation mit -1 zu

$$\begin{aligned}(y_1 - y_2, x_1 - x_2)_h &\geq (B_0(x_1) - B_0(x_2), x_1 - x_2)_h \\ &= \gamma |x_1 - x_2|_1^2 + \frac{1}{\tau}(\mathcal{G}^h(x_1 - u^{\text{alt}}) - \mathcal{G}^h(x_2 - u^{\text{alt}}), x_1 - x_2)_h \\ &= \gamma |x_1 - x_2|_1^2 + \frac{1}{\tau}|\mathcal{G}^h(x_1 - x_2)|_1^2 \geq 0.\end{aligned}$$

Da mit $x_1, x_2 \in S_m^h$ auch $(x_1 - x_2, 1)_h = 0$ gilt, folgt aus der diskreten Poincaré-Ungleichung somit weiterhin

$$(y_1 - y_2, x_1 - x_2)_h \geq \gamma |x_1 - x_2|_1^2 \geq C\gamma \|x_1 - x_2\|_h^2.$$

Nun liefert die Anwendung von Satz 5.2.1 und Korollar 5.2.1 die Konvergenz. ■

5.2.3 Die Auswertung der Operatoren

Dass $J_A^\lambda y$ und $J_B^\lambda y$ für jedes $y \in S^h$ existieren und eindeutig bestimmt sind, folgt aus der maximalen Monotonie und Satz 2.3.2. Im Folgenden wird untersucht, wie diese Operatoren ausgewertet werden können.

Sei $y \in S^h$ gegeben. Dann gilt nach Korollar 2.2.1, da $(x - y, \cdot)_h$ die Ableitung der stetigen, konvexen Funktion $\frac{1}{2}\|x - y\|_h^2$ nach x ist:

$$\begin{aligned} & x = J_A^\lambda y \\ \Leftrightarrow \quad & y \in (I + \lambda \partial I_{K^h})x \\ \Leftrightarrow \quad & 0 \in \partial(I_{K^h}(x) + \frac{1}{2}\|x - y\|_h^2) \\ \Leftrightarrow \quad & x \in K^h : \qquad \|x - y\|_h \leq \|z - y\|_h \qquad \forall z \in K^h. \end{aligned}$$

Die Auswertung von $J_A^\lambda y$ entspricht also einer Projektion von y in K^h. Da $x \in K^h$ genau dann gilt, wenn an jedem Knoten $x_i \in \mathcal{N}^h$ der Triangulierung $x(x_i) \in [-1; 1]$ gilt, lässt sich diese Projektion von y in K^h unabhängig an jedem Knoten ausführen. Es gilt also für $x = J_A^\lambda y$ und alle $x_i \in \mathcal{N}^h$

$$x(x_i) = \begin{cases} -1 & \text{, falls } y(x_i) < -1 \\ y(x_i) & \text{, falls } y(x_i) \in [-1; 1] \\ 1 & \text{, falls } y(x_i) > 1 \end{cases}.$$

Damit ist der Operator J_A^λ einfach auszuwerten. Die nun auftauchenden Variablen u und w repräsentieren nicht die Lösung von (CH), werden aber verwendet, damit das Ziel der Umformung klar wird. Analog folgt aus Korollar 2.2.1 für den Operator J_B^λ

$$\begin{aligned}
& u = J_B^\lambda y \\
\Leftrightarrow \quad & y \in (I + \lambda B_0 + \lambda \partial I_{S_m^h}) u \\
\Leftrightarrow \quad & y - u - \lambda B_0 u \in \lambda \partial I_{S_m^h} u \\
\Leftrightarrow \quad & u \in S_m^h : \qquad (B_0(u), v-u)_h + \frac{1}{\lambda}(u-y, v-u)_h \geq 0 \qquad \forall v \in S_m^h. \qquad (5.2.3)
\end{aligned}$$

Da der Algorithmus das direkte Lösen einer Variationsungleichung umgehen soll und der in B_0 enthaltene Operator $\mathcal{G}^h$ sehr aufwendig auszuwerten ist, wird nun eine Darstellung als Variationsgleichung ohne den Operator $\mathcal{G}^h$ entwickelt. Dazu wird $w_0 := -\frac{1}{\tau}\mathcal{G}^h(u - u^{\text{alt}})$ gesetzt. Dann lässt sich (5.2.3) schreiben als:

Finde $u \in S_m^h$ und $w_0 \in S_0^h$ mit

$$\begin{aligned}
\tau(\nabla w_0, \nabla v) + (u, v)_h &= (u^{\text{alt}}, v)_h && \forall v \in S^h, \\
\gamma(\nabla u, \nabla v - \nabla u) + \frac{1}{\lambda}(u, v-u)_h - (w_0, v-u)_h &\geq (u^{\text{alt}}, v-u)_h + \frac{1}{\lambda}(y, v-u)_h && \forall v \in S_m^h.
\end{aligned}$$

Die Forderung $u \in S_m^h$ kann zu $u \in S^h$ gelockert werden, da $(u,1)_h = (u^{\text{alt}}, 1)_h = m$ nach Testen der ersten Gleichung mit $v \equiv 1$ klar ist. Nun sei $v \in S^h$ beliebig, dann gilt $v = v_m + m_v$ mit $m_v = \frac{(v,1)_h - m}{|\Omega|} \in \mathbb{R}$ für ein $v_m \in S_m^h$. Da die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(m_v) = (B_0(u), m_v)_h + \frac{1}{\lambda}(u - y, m_v)_h$$

stetig und linear ist und $(a,b)_h$ für konstante Funktionen a, b ein Skalarprodukt auf $\mathbb{R}$ definiert, existiert immer genau ein $m_w \in \mathbb{R}$, so dass $f(m_v) = (m_w, m_v)_h$ für alle $m_v \in \mathbb{R}$ gilt. Damit gilt ferner

$$f(m_v) = (m_w, m_v)_h + \underbrace{(m_w, v_m - u)_h}_{=0} = (m_w, v-u)_h$$

für alle $v \in S^h$. Wird diese Gleichung zur mit dem jeweiligen v_m getesteten Ungleichung (5.2.3) addiert, so lässt sich diese darstellen als:

Finde $u \in S^h$, $w_0 \in S_0^h$ und $m_w \in \mathbb{R}$ mit

$$\begin{aligned}
\tau(\nabla w_0, \nabla v) + (u, v)_h &= (u^{\text{alt}}, v)_h && \forall v \in S^h, \\
\gamma(\nabla u, \nabla v - \nabla u) + \frac{1}{\lambda}(u, v-u)_h - (w_0, v-u)_h & \\
\geq (u^{\text{alt}}, v-u)_h + \frac{1}{\lambda}(y, v-u)_h + (m_w, v-u)_h & && \forall v \in S^h,
\end{aligned}$$

wobei die Ungleichung nun mit allen $v = v_m + m_v \in S^h$ getestet werden kann. Der für $v \notin S^h_m$ zusätzlich entstehende Term wird dabei durch $(m_w, v-u)_h$ kompensiert. Da in der Ungleichung mit allen Elementen des Raumes S^h getestet werden kann, führt Testen mit $u + v$ und $u - v$ zu einer Variationsgleichung. Wird nun noch $w := w_0 + m_w$ gesetzt, so erfüllt offensichtlich auch w die erste Gleichung. Damit ist $u = J^\lambda_B y$ äquivalent zu:

Finde $u \in S^h$ und $w \in S^h$ mit

$$\begin{aligned} \tau(\nabla w, \nabla v) + (u, v)_h &= (u^{\text{alt}}, v)_h && \forall v \in S^h, \\ \gamma(\nabla u, \nabla v) + \frac{1}{\lambda}(u, v)_h - (w, v)_h &= (u^{\text{alt}}, v)_h + \frac{1}{\lambda}(y, v)_h && \forall v \in S^h. \end{aligned}$$

5.2.4 Implementierung des Algorithmus

Jetzt wird eine algebraische Formulierung des Algorithmus gegeben und auf die Implementierung eingegangen. Zu einem Startwert u^0 sei $v^0 \in (I + \lambda B)u^0$ beliebig gewählt. Dann ist der Algorithmus definiert als

$$v^{n+1} = (2J^\lambda_A - I)(2J^\lambda_B - I)v^n,$$

wobei $u^n = J^\lambda_B v^n$ für alle n gilt. Setze nun

$$z^n := 2u^n - v^n, \qquad u^{n+\frac{1}{2}} := J^\lambda_A z^n,$$

dann lässt sich der Algorithmus für einen Startwert u^0 und $v^0 \in (I + \lambda B)u^0$ schreiben als:

(1) Für $n = 0, ...$ bestimme u^{n+1} durch:
(2) Setze $z^n = 2u^n - v^n$.
(3) Setze $u^{n+\frac{1}{2}} = J^\lambda_A z^n$.
(4) Setze $v^{n+1} = 2u^{n+\frac{1}{2}} - z^n$.
(5) Setze $u^{n+1} = J^\lambda_B v^{n+1}$.

In dieser Formulierung entspricht der Algorithmus mit der angegebenen Zerlegung der Operatoren im Wesentlichen dem, den Barrett, Blowey und Garcke in [BBG99] für die Lösung einer diskreten Cahn-Hilliard Gleichung mit degenerierter Mobilität entwickeln. Auch Copetti und Elliott verwenden in [CE92] einen ähnlichen Splitting-Algorithmus für eine diskrete Cahn-Hilliard Gleichung mit logarithmischer freier Energie.

Zum beliebigen Startwert u^0 muss zu Beginn ein $v^0 \in (I + \lambda B)u^0$ gewählt werden. Die Darstellung des Operators $J^\lambda_B = (I + \lambda B)^{-1}$ aus dem vorigen Abschnitt ergibt dafür in der algebraischen Formulierung

$$\begin{aligned} \tau A\underline{w}^0 + M\underline{u}^0 &= b \\ \gamma A\underline{u}^0 + \frac{1}{\lambda}M\underline{u}^0 - M\underline{w}^0 &= b + \frac{1}{\lambda}\underline{v}^0. \end{aligned}$$

Zunächst ist also ein $\underline{w}^0$ mit $\tau A\underline{w}^0 = b - M\underline{u}^0$ zu finden. Dieses ist offensichtlich nicht eindeutig bestimmt. Da M invertierbar ist, erfüllt genau

$$\underline{v}^0 = \underline{u}^0 + \lambda M^{-1}\left(\gamma A\underline{u}^0 - M\underline{w}^0 - b\right)$$

die Bedingung $v^0 \in (I + \lambda B)u^0$. Durch die Festlegung der Konstante für $\underline{w}^0$ ist die Wahl von $\underline{v}$ getroffen. Wird der Algorithmus in der Basis von S^h dargestellt und werden die Erkenntnisse aus dem vorherigen Abschnitt genutzt, so ergibt dies die folgende algebraische Formulierung zum Startwert $\underline{u}^0$:

(1) Finde ein $\underline{w}^0$ mit $\tau A\underline{w}^0 = b - M\underline{u}^0$.

(2) Setze $\underline{v}^0 = \underline{u}^0 + \lambda M^{-1}\left(\gamma A\underline{u}^0 - M\underline{w}^0 - b\right)$.

(3) Für $n = 0, ...$ bestimme $\underline{u}^{n+1}$ durch:

(4) Setze $\underline{z}^n = 2\underline{u}^n - \underline{v}^n$.

(5) Für $i = 1, ..., N$ bestimme $\underline{u}_i^{n+\frac{1}{2}}$ durch:

(6) Setze $\underline{u}_i^{n+\frac{1}{2}} = \min\{\max\{\underline{z}_i^n, -1\}, 1\}$.

(7) Setze $\underline{v}^{n+1} = 2\underline{u}^{n+\frac{1}{2}} - \underline{z}^n$.

(8) Löse das lineare Gleichungssystem $\underline{u}^{n+1}, \underline{w}^{n+1} \in \underline{S}^h$:

$$\tau A\underline{w}^{n+1} + M\underline{u}^{n+1} = b,$$
$$\left(\gamma A + \frac{1}{\lambda}M\right)\underline{u}^{n+1} - M\underline{w}^{n+1} = b + \frac{1}{\lambda}\underline{v}^{n+1}.$$

Da das lineare Gleichungssystem in (8) ein nichtsymmerisches Ssttelpunktproblem (siehe etwa [Bra97]) darstellt, ist es sehr aufwendig zu lösen.

5.3 Das Gauß-Seidel Verfahren

Das lineare Gauß-Seidel Verfahren ist ein klassisches Iterationsverfahren für lineare Gleichungssysteme. Eine klassische Methode, den Algorithmus zu motivieren, besteht darin, das lineare Gleichungssystem

$$x \in \mathbb{R}^N : \qquad Ax = b \tag{5.3.1}$$

mit $A \in \mathbb{R}^{N,N}$ und $b \in \mathbb{R}^N$ als Fixpunktproblem darzustellen und dessen Lösung mittels einer Fixpunktiteration zu approximieren. Offensichtlich ist die Darstellung als Fixpunktproblem nicht eindeutig bestimmt, ein Beispiel ist etwa $x = \Phi(x)$ mit $\Phi(x) = Ax - b + x$. Für die Gauß-Seidel Iteration wird eine Zerlegung der Matrix A in obere und untere Dreiecksmatrix zugrundegelegt. Dazu sei $A = A_D - A_L - A_U$ mit

$$\begin{aligned}
(A_D)_{ij} &= A_{ij} && \text{für } i = j, \text{ sonst} = 0,\\
(A_L)_{ij} &= -A_{ij} && \text{für } i > j, \text{ sonst} = 0,\\
(A_U)_{ij} &= -A_{ij} && \text{für } i < j, \text{ sonst} = 0.
\end{aligned}$$

Falls A symmetrisch ist, gilt $A_U = A_L^T$. In jedem Fall ist das Gleichungssystem äquivalent zu $(A_D - A_L)x = b + A_U x$, was für invertierbares $A_D - A_L$ zu

$$x = (A_D - A_L)^{-1}(b + A_U x) = \Phi(x)$$

führt. Das Lösen des Gleichungssystems ist also äquivalent zum Finden eines Fixpunktes von Φ. Die Gauß-Seidel Iteration ist die dazu gehörende Fixpunktiteration $x^{n+1} = \Phi(x^n)$. Aus dem Banach'schen Fixpunktsatz (siehe zum Beisiel [Růž04]) folgt, falls Φ eine Kontraktion ist, die Konvergenz der Iteration gegen die Lösung des linearen Gleichungssystems. Für beliebige Matrizen A ist dies allerdings im Allgemeinen nicht gegeben. Auch die Invertierbarkeit von $(A_D - A_L)$ und damit die Durchführbarkeit der Iteration ist im Allgemeinen nicht gegeben. Eine andere Motivation für das Gauß-Seidel Verfahren ist das Prinzip der Energieminimierung. Vor dem Hintergrund dieser Idee lässt sich das Verfahren auch auf allgemeinere Probleme ausweiten.

5.3.1 Das Gauß-Seidel Verfahren für lineare Probleme

Um das Gauß-Seidel Verfahren durch Energieminimierung zu erklären, wird das Gleichungssystem unter der Voraussetzung, dass A symmetrisch und positiv definit ist, als Minimierungsproblem dargestellt. Dazu wird das Energiefunktional $f(x) = \frac{1}{2}x^T Ax - x^T b$ betrachtet. Für symmetrisches, positiv definites A ist es strikt konvex (siehe z.B. [Bra97], [KA00]) und die Lösung des Gleichungssytems ist äquivalent zu

$$x \in \mathbb{R}^N : \qquad f(x) \leq f(y) \qquad \forall y \in \mathbb{R}^N.$$

Zur Approximation der Lösung dieses Minimierungsproblems soll nun die Energie der Iterierten sukzessive verkleinert werden. Ein Minimieren im gesamten Raum $\mathbb{R}^N$ bedeutet das exakte, aufwändige Lösen des Gleichungssystems. Um dies durch einfache Operationen anzunähern, wird in einem Iterationsschritt nacheinander in Richtung jedes Basisvektors $e_i = (0, ..., 0, 1, 0, ..., 0)$ minimiert, also jeweils in eindimensionalen Unterräumen. Nun wird

$$x^{n+1,i} := (x_1^{n+1}, ..., x_{i-1}^{n+1}, x_i^{n+1}, x_{i+1}^n, ..., x_N^n)$$

für $i = 0, ..., N$ gesetzt. Damit ist zu bekanntem $x^{n+1,i-1}$ jeweils $x_i^{n+1,i} = x_i^{n+1} \in \mathbb{R}$ mit

$$f(x^{n+1,i}) \leq f(x^{n+1,i-1} + \mu e_i) \qquad \forall \mu \in \mathbb{R}$$

gesucht. Für die Ableitung von f in Richtung eines e_i an der Stelle $x^{n+1,i}$ gilt

$$\begin{aligned}
\frac{df}{de_i}(x^{n+1,i}) &= \frac{1}{2}\left(\sum_{k \neq i} a_{ki} x_k^{n+1,i} + \sum_{j \neq i} a_{ij} x_j^{n+1,i}\right) + a_{ii} x_i^{n+1,i} - b_i \\
&= \sum_{j=1}^{i} a_{ij} x_j^{n+1} + \sum_{j=i+1}^{N} a_{ij} x_j^n - b_i \\
&= \left((A_D - A_L)x^{n+1} - A_U x^n - b\right)_i .
\end{aligned}$$

Da f differenzierbar und strikt konvex ist, lässt sich das Minimieren in Richtung von e_i durch Setzen von $\frac{df}{de_i}(x^{n+1,i}) = 0$ erreichen, was auf $(A_D - A_L)x^{n+1} = b + A_U x^n$ und damit auf die eingangs genannte Fixpunktiteration führt.

Im Allgemeinen nimmt die Konvergenzgeschwindigkeit des Gauß-Seidel Verfahrens mit der Größe des Systems rapide ab, was es als Löser für die bei der Diskretisierung partieller Differentialgleichungen entstehenden Probleme ungeeignet macht. Jedoch kann es in einer symmetrischen Version als Vorkonditionierer zum Beispiel das CG-Verfahren (siehe zum Beipiel [Bra97],[KA00]) beschleunigen. Der Minimierungseigenschaft kommt bei den Mehrgittermethoden (siehe zum Beispiel [Bra93], [Bra97]) besondere Bedeutung zu. Hier wird eine Hierarchie von Gittern und dazugehörigen Unterräumen von S^h betrachtet. Es wird nicht nur in Richtung der hochfrequenten Basisfunktionen auf dem feinsten Gitter, sondern auch in Richtung der niederfrequenten Basisfunktionen auf gröberen Gittern minimiert. So wird das Spektrum des (Laplace-)Operators besser approximiert und teilweise eine Gitter-(also Systemgrößen-) unabhängige Konvergenzgeschwindigkeit erreicht. Den Gauß-Seidel Schritten kommt hier die Rolle eines so genannten Glätters auf jedem Gitter zu.

Es besteht auch die Möglichkeit jeweils in Unterräumen der Dimension $m \geq 1$ zu minimieren. Dies führt bei geeigneter Nummerierung zu einer Zerlegung in „Blockmatrizen“, die aus $m \times m$-Blöcken besteht. Es sind dann jeweils $m-$dimensionale Teilprobleme zu lösen. In diesem Fall spricht man vom Block-Gauß-Seidel Verfahren.

5.3.2 Das Gauß-Seidel Verfahren für Variationsungleichungen

Ausgehend vom Minimierungsansatz lässt sich das Gauß-Seidel Verfahren auch für Minimierungsprobleme der Art

$$x \in K: \qquad f(x) \leq f(y) \qquad \forall y \in K$$

erweitern. Dabei ist f strikt konvex und $K = K_1 \times ... \times K_N$ das Produkt abgeschlossener, nichtleerer, konvexer Mengen $K_i \subset \mathbb{R}$, also selbst konvex. Es ist klar, dass solche Mengen genau von der Form $K_i = [a_i; b_i] \cap \mathbb{R}$ mit $-\infty \leq a_i \leq b_i \leq \infty$ sind, also abgeschlossene beschränkte oder unbeschränkte Intervalle. Um sicherzustellen, dass ein solches Minimum existiert, wird zusätzlich $J(x) \to \infty$ für $|x| \to \infty$ gefordert. Für derartige Probleme wird im projizierten Gauß-Seidel Verfahren in jedem Schritt nicht mehr in einem eindimensionalen Unterraum, sondern in einer konvexen Teilmenge K_i eines solchen minimiert. Gesucht ist nun also zu bekanntem $x^{n+1,i-1}$ jeweils $x_i^{n+1,i} = x_i^{n+1} \in K_i$ mit

$$f(x^{n+1,i}) \leq f(x^{n+1,i-1} + \mu e_i) \qquad \forall \mu \in K_i.$$

Da f strikt konvex ist, ist dies mit $x_k^{n+1,i-\frac{1}{2}} = x_k^{n+1,i} = x_k^{n+1,i-1}$ für $k \neq i$ äquivalent zu

$$x_i^{n+1,i-\frac{1}{2}} \in \mathbb{R}: \qquad f(x^{n+1,i-\frac{1}{2}}) \leq f(x^{n+1,i-1} + \mu e_i) \qquad \forall \mu \in \mathbb{R},$$
$$x_i^{n+1} = P_{K_i} x_i^{n+1,i-\frac{1}{2}} := \min\{\max\{x_i^{n+1,i-\frac{1}{2}}, a_i\}, b_i\},$$

also zum Minimieren in $\mathbb{R}$ und Projizieren der Lösung in K_i. Es ist klar, dass das projizierte Gauß-Seidel Verfahren im unrestringierten Fall $K = \mathbb{R}^N$ in das lineare Gauß-Seidel Verfahren übergeht. Ist f von der Form $f(x) = \frac{1}{2}x^T Ax - x^T b$ mit symmetrischer, positiv definiter Matrix A, so ist das unrestringierte Minimierungsproblem äquivalent zum linearen Gleichungssystem $Ax = b$, also zur Variationsgleichung

$$x \in \mathbb{R}^N : \qquad y^T Ax = y^T b \qquad \forall y \in \mathbb{R}^N.$$

Das lineare Gauß-Seidel Verfahren entspricht dann einer Zerlegung der Matrix in dieser Variationsgleichung. Im restringierten Fall $K \neq \mathbb{R}^N$ nimmt f nach Korollar 2.2.1 genau dann sein Minimum an einer Stelle x in der abgeschlossenen, konvexen, nichtleeren Menge $K \subset \mathbb{R}^N$ an, wenn $x \in K$ und $(Df(x), y - x) \geq 0$ für alle $y \in K$, also

$$x \in K : \qquad (y - x)^T Ax \geq (y - x)^T b \qquad \forall y \in K$$

gilt. Diese Variationsungleichung für das restringierte Problem ist ein Analogon zur Variationsgleichung für das unrestringierte Problem. Anwendung von Korollar 2.2.1 auf die Minimierungsprobleme im projizierten Gauß-Seidel Verfahren ergibt, dass diese äquivalent zum Suchen von $x_i^{n+1} \in K_i$ mit

$$(y_i - x_i^{n+1,i}) \left(\frac{df}{de_i}(x^{n+1,i}) \right) = (y_i - x_i^{n+1}) \left((A_D - A_L)x^{n+1} - A_U x^n - b \right)_i \geq 0$$

für alle $y_i \in K_i$ sind. Summieren dieser Gleichungen für $i = 1, ..., N$ ergibt

$$x^{n+1} \in K : \qquad (y - x^{n+1})^T (A_D - A_L)x^{n+1} \geq (y - x^{n+1})^T (b + A_U x^n) \qquad \forall y \in K.$$

Die einzelnen Gleichungen für x_i^{n+1} können reproduziert werden, indem mit solchen $y \in K$ getestet wird, für die $y_i \in K_i$ beliebig ist und $y_j = x_j^{n+1}$ für $j \neq i$ gilt. Damit entspricht auch das projizierte Gauß-Seidel Verfahren einer Zerlegung der Systemmatrix im zum Minimierungsproblem gehörenden Variationsproblem, das hier eine Variationsungleichung ist. Das folgende Konvergenzresultat gilt auch für allgemeinere Probleme mit nichtquadratischem f.

Satz 5.3.1 *Ist $f : \mathbb{R}^N \to \mathbb{R}$ stetig differenzierbar, strikt konvex und gilt ferner $f(x) \to \infty$ für $|x| \to \infty$, so konvergiert das projizierte Gauß-Seidel Verfahren gegen die Lösung des restringierten Minimierungsproblems in $K = K_1 \times ... \times K_N$.*

BEWEIS. Ein Beweis ist in [GLT81] gegeben. ■

Dies lässt sich nun weiter zu einer Blockvariante verallgemeinern, wobei sich die Anforderungen an die konvexe Menge ändern. Die allgemeine Formulierung sieht wie folgt aus: Es soll das Minimierungsproblem in $K \subset V = \mathbb{R}^M$ gelöst werden. Es sei $V = V_1 \times ... \times V_N$ eine Zerlegung von V in Unterräume der Dimension $\dim(V_i) = m_i$ und $K = K_1 \times ... \times K_N$ eine Zerlegung von K in abgeschlossene, nichtleere, konvexe Mengen $K_i \subset V_i$. Dann ist in jedem Iterationsschritt nacheinander in jeder konvexen Menge K_i zu minimieren. Dabei ist jeweils ein Problem der Dimension m_i zu lösen. Die Konvergenz lässt sich hierfür unter denselben Voraussetzungen an

f genauso zeigen wie für eine Zerlegung in eindimensionale Teilräume. Offensichtlich umfasst diese Formulierung alle vorherigen als Spezialfälle.

Allen Formulierungen gemein ist, dass ihnen die Minimierung des Funktionals f zugrundeliegt. Für Variationsprobleme mit nichtsymmetrischer oder nicht positiv definiter Systemmatrix ist die Äquivalenz zu einem Minimierungsproblem jedoch nicht in der benötigten Form gegeben. In diesem Fall lässt sich zwar durch eine analoge Zerlegung der Matrix eine Iteration „konstruieren“, doch ist dann Satz 5.3.1 nicht anwendbar und es ist im Allgemeinen keine Konvergenz gegen die Lösung des Variationsproblems zu erwarten.

5.3.3 Ein Gauß-Seidel Verfahren für die diskrete Cahn-Hilliard Gleichung

Obwohl das Problem (CH) keine symmetrische, positiv definite Systemmatrix besitzt, wird eine Block-Gauß-Seidel Iteration darauf angewandt. Diese basiert nicht auf dem Minimierungsansatz, sondern auf einer blockweisen Zerlegung der Systemmatrix in untere und obere Dreiecksmatrizen. Da Satz 5.3.1 auf diese Iteration nicht anwendbar ist, wird die Konvergenz speziell hierfür gezeigt. Der Konvergenzbeweis für diese Iteration wurde von Barrett für eine nichtlineare diskrete Cahn-Hilliard Gleichung ohne Hindernispotential geführt, allerdings nicht veröffentlicht. In [Sen03] ist dieser Beweis für die hier behandelte Gleichung mit Hindernispotential angepasst. Da er weder den Banach'schen Fixpunktsatz noch Minimierungsargumente verwendet, wie dies beim Gauß-Seidel Verfahren im Allgemeinen der Fall ist, wird er hier kurz ausgeführt.

Die Iterationsfolge $(u^n, w^n)^T$ wird durch die für das Gauß-Seidel Verfahren typische Zerlegung der Steifigkeitsmatrizen $A = A_D - A_L - A_U$ in der algebraische Formulierung von (CH) gewonnen. Da es sich bei der Massenmatrix M wegen der Verwendung des „gelumpten“ L^2-Skalarproduktes um eine Diagonalmatrix handelt, muss dafür keine Zerlegung betrachtet werden. Die Iteration sieht dann wie folgt aus:

Finde $(\underline{u}^{n+1}, \underline{w}^{n+1})^T \in \underline{K}^h \times \underline{S}^h$ mit

$$\begin{aligned}(\underline{y} - \underline{u}^{n+1})^T \gamma (A_D - A_L)\underline{u}^{n+1} - (\underline{y} - \underline{u}^{n+1})^T M \underline{w}^{n+1} \\ \geq (\underline{y} - \underline{u}^{n+1})^T b + (\underline{y} - \underline{u}^{n+1})^T \gamma (A_U)\underline{u}^n \qquad \forall \underline{y} \in \underline{K}^h, \quad (5.3.2)\end{aligned}$$

$$M\underline{u}^{n+1} + \tau(A_D - A_L)\underline{w}^{n+1} = b + \tau(A_U)\underline{w}^{n+1}. \quad (5.3.3)$$

Die Wohldefiniertheit der Iterationsfolge wird im nächsten Abschnitt zur Implementierung des Algorithmus gezeigt.

Satz 5.3.2 *Sei $\mathcal{T}^h$ schwach spitzwinklig, dann konvergiert die Folge $(u^n, w^n)^T$ gegen die Lösung $(u, w)^T$ von (CH).*

BEWEIS. Wir definieren zunächst für $n \geq 1$ die Fehlerfunktionen $\underline{e}_u^n = \underline{u} - \underline{u}^n$ und $\underline{e}_w^n = \underline{w} - \underline{w}^n$, subtrahieren die Gleichung (5.3.3) für die Iterationsfolge von der Gleichung (5.1.2) in (CH) für

die Lösung und erhalten

$$M\underline{e}_u^{n+1} + \tau(A_D - A_L)\underline{e}_w^{n+1} = \tau(A_U)\underline{e}_w^n. \tag{5.3.4}$$

Nun multiplizieren wir die Differenz mit $(\underline{e}_w^{n+1})^T$

$$(\underline{e}_w^{n+1})^T M\underline{e}_u^{n+1} + \tau(\underline{e}_w^{n+1})^T(A_D - A_L)\underline{e}_w^{n+1} = \tau(\underline{e}_w^{n+1})^T(A_U)\underline{e}_w^n. \tag{5.3.5}$$

Eine solche Subtraktion der Ungleichungen ist nicht möglich. Lemma 5.3.1 zeigt jedoch, dass für diese die ähnliche Relation

$$\gamma(\underline{e}_u^{n+1})^T(A_D - A_L)\underline{e}_u^{n+1} - (\underline{e}_u^{n+1})^T M\underline{e}_w^{n+1} \leq \gamma(\underline{e}_u^{n+1})^T(A_U)\underline{e}_u^n \tag{5.3.6}$$

gilt. Addition von (5.3.5) und (5.3.6) liefert nun

$$\begin{aligned}\gamma(\underline{e}_u^{n+1})^T(A_D - A_L)\underline{e}_u^{n+1} + \tau(\underline{e}_w^{n+1})^T(A_D - A_L)\underline{e}_w^{n+1} \\ \leq \gamma(\underline{e}_u^{n+1})^T(A_U)\underline{e}_u^n + \tau(\underline{e}_w^{n+1})^T(A_U)\underline{e}_w^n.\end{aligned} \tag{5.3.7}$$

Um daraus nun auf die Konvergenz der Folge zu schließen, benötigen wir folgende Eigenschaften der Matrix A:

(1) Da A symmetrisch ist, gilt für $\underline{x} \in \mathbb{R}^N$

$$\begin{aligned}\underline{x}^T(A_D - A_L)\underline{x} &= \sum_{i=1}^N A_{ii}\underline{x}_i^2 + \sum_{i=1}^N \sum_{j=1}^{i-1} A_{ij}\underline{x}_i\underline{x}_j \\ &= \frac{1}{2}\sum_{i=1}^N \left(A_{ii}\underline{x}_i^2 + \sum_{j=1}^{i-1} A_{ij}\underline{x}_i\underline{x}_j + \sum_{j=1}^{i-1} A_{ij}\underline{x}_i\underline{x}_j \right) + \frac{1}{2}\underline{x}^T A_D\underline{x} \\ &= \frac{1}{2}\underline{x}^T A\underline{x} + \frac{1}{2}\underline{x}^T A_D\underline{x}.\end{aligned}$$

(2) Wir definieren die Diagonalmatrizen A_{D_L} und A_{D_U} durch

$$(A_{D_L})_{jj} = \sum_{i=1}^{j-1} -A_{ij}, \qquad (A_{D_U})_{ii} = \sum_{j=i+1}^{N} -A_{ij}.$$

Da $\mathcal{T}^h$ schwach spitzwinklig ist, gilt nach Bemerkung 4.1.1 $-A_{ij} = -(\nabla\phi_i, \nabla\phi_j) \geq 0$ für $i \neq j$ und $(A_{D_L})_{ii}, (A_{D_U})_{ii} \geq 0$. Wegen $a^2 + b^2 \geq 2ab$ gilt damit für $\underline{x}, \underline{y} \in \mathbb{R}^N$

$$\begin{aligned}\underline{y}^T A_U\underline{x} = \sum_{i=1}^N \sum_{j=i+1}^N -A_{ij}\underline{y}_i\underline{x}_j &\leq \sum_{i=1}^N \sum_{j=i+1}^N -A_{ij}\frac{1}{2}(\underline{y}_i^2 + \underline{x}_j^2) \\ &= \frac{1}{2}\sum_{i=1}^N \underbrace{\sum_{j=i+1}^N -A_{ij}}_{=(A_{D_U})_{ii}}\underline{y}_i^2 + \frac{1}{2}\sum_{j=1}^N \underbrace{\sum_{i=1}^{j-1} -A_{ij}}_{=(A_{D_L})_{jj}}\underline{x}_j^2 \\ &= \frac{1}{2}\underline{y}^T A_{D_U}\underline{y} + \frac{1}{2}\underline{x}^T A_{D_L}\underline{x}.\end{aligned}$$

(3) Es gilt $A_{ii} - (A_{D_L})_{ii} - (A_{D_U})_{ii} = \sum_{j=1}^{N} A_{ij} = (\nabla\phi_i, \nabla 1) = 0$ und somit ferner $A_D = A_{D_L} + A_{D_U}$.

Wenden wir nun diese Eigenschaften auf (5.3.7) an, so erhalten wir

$$\underbrace{\gamma(\underline{e}_u^{n+1})^T A(\underline{e}_u^{n+1}) + \tau(\underline{e}_w^{n+1})^T A(\underline{e}_w^{n+1})}_{:=x^{n+1}} + \underbrace{\gamma(\underline{e}_u^{n+1})^T A_{D_L}(\underline{e}_u^{n+1}) + \tau(\underline{e}_w^{n+1})^T A_{D_L}(\underline{e}_w^{n+1})}_{:=y^{n+1}}$$
$$\leq \underbrace{\gamma(\underline{e}_u^{n})^T A_{D_L}(\underline{e}_u^{n}) + \tau(\underline{e}_w^{n})^T A_{D_L}(\underline{e}_w^{n})}_{:=y^{n}},$$

also $x^{n+1} + y^{n+1} \leq y^n$. Damit ist die Folge y^n monoton fallend. Wegen $(A_{D_L})_{ii} \geq 0$ ist sie nach unten beschränkt und somit konvergent. Damit konvergiert die positive Folge x^n gegen Null, woraus $\gamma|e_u^n|_1^2 + \tau|e_w^n|_1^2 \to 0$ folgt. Multiplizieren wir nun (5.3.4) mit $e_N = (0, ..., 0, 1)$, so erhalten wir

$$0 = M_{NN}(\underline{e}_u^{n+1})_N + \tau \sum_{j=1}^{N} A_{Nj}\underline{e}_w^{n+1}.$$

Da e_w^n in der $|\cdot|_1$-Halbnorm gegen 0 geht, konvergiert der rechte und somit auch der linke Term gegen 0, woraus $e_u^n(x_N) \to 0$ am Knoten x_N folgt. Da ferner die $|\cdot|_1$-Halbnorm von e_u^n gegen 0 geht, folgt daraus die Konvergenz von $\underline{e}_u^n$ gegen $0 \in \mathbb{R}^N$. Da $|(u,1)_h| < |\Omega|$ vorausgesetzt war, existiert ein Knoten x_i mit $|u(x_i)| = |\underline{u}_i| < 1$. Da $\underline{u}_i^n$ gegen $\underline{u}_i$ konvergiert, gilt ab einem $n_0 > 1$ auch $|\underline{u}_i^n| < 1$. Nun liefert das Testen von Ungleichung (5.1.2) für die Lösung mit geeignetem $\underline{y}$ (setze: $\underline{y}_j = \underline{u}_j$ für $j \neq i$ und $\underline{y}_i = \underline{u}_i \pm \epsilon$) und von Ungleichung (5.3.2) für die Iterationsfolge mit ebenfalls geeignetem $\underline{y}$ für $n > n_0$

$$\gamma \sum_{j=1}^{N} (A_D - A_L - A_U)_{ij}\underline{u}_j - M_{ii}\underline{w}_i = b_i,$$
$$\gamma \sum_{j=1}^{N} (A_D - A_L)_{ij}\underline{u}_j^{n+1} - \gamma \sum_{j=1}^{N} (A_U)_{ij}\underline{u}_j^{n} - M_{ii}\underline{w}_i^{n+1} = b_i.$$

Subtraktion der Gleichungen liefert

$$\gamma \sum_{j=1}^{N} (A_D - A_L)_{ij}(\underline{e}_u^{n+1})_j - \gamma \sum_{j=1}^{N} (A_U)_{ij}(\underline{e}_u^{n})_j = M_{ii}(\underline{e}_w^{n+1})_i.$$

Da die linke Seite gegen Null konvergiert, gilt auch $(\underline{e}_w^n)_i = e_w^n(x_i) \to 0$ und da ferner die $|\cdot|_1$-Halbnorm von e_w^n gegen Null geht, folgt daraus die Konvergenz von $\underline{e}_w^n$ gegen $0 \in \mathbb{R}^N$. ■

Lemma 5.3.1 *Sei $(u, w)^T$ die Lösung von (CH) und $(u^{n+1}, w^{n+1})^T$ eine Iterierte der Gauß-Seidel Iteration, dann gilt*

$$\gamma(\underline{e}_u^{n+1})^T(A_D - A_L)\underline{e}_u^{n+1} - (\underline{e}_u^{n+1})^T M \underline{e}_w^{n+1} \leq \gamma(\underline{e}_u^{n+1})^T(A_U)\underline{e}_u^{n}.$$

BEWEIS. Setzen wir

$$\underline{r} := (A_D - A_L - A_U)\underline{u} - M\underline{u} - b,$$
$$\underline{r}^{n+1} = (A - A_D)\underline{u}^{n+1} - A_U\underline{u}^n - M\underline{u}^{n+1} - b,$$

dann gilt $(\underline{y} - \underline{u})^T\underline{r} \geq 0$ und $(\underline{y} - \underline{u}^{n+1})^T\underline{r}^{n+1} \geq 0$ für $\underline{y} \in \underline{K}^h$. Setzen wir $\underline{y}_j = \underline{u}_j$ für $j \neq i$, so erhalten wir durch Einsetzen von $\underline{y}$ mit $\underline{y}_i \in [-1, 1]$ geeignet $\underline{r}_i = 0$ für $\underline{u}_i \in (-1, 1)$, $\underline{r}_i \leq 0$ für $\underline{u}_i = 1$ und $\underline{r}_i \geq 0$ für $\underline{u}_i = -1$. Für $\underline{u}_i^{n+1}$ und $\underline{r}_i^{n+1}$ erhalten wir dieselben Aussagen. Zusammen liefert dies:

$$\begin{aligned}
&\underline{u}_i = \underline{u}_i^{n+1} && \Rightarrow && (\underline{u} - \underline{u}^{n+1})_i = 0,\\
&\underline{u}_i, \underline{u}_i^{n+1} \in (-1, 1) && \Rightarrow && (\underline{r} - \underline{r}^{n+1})_i = 0,\\
&\underline{u}_i = 1, \underline{u}_i^{n+1} < 1 && \Rightarrow && (\underline{u} - \underline{u}^{n+1})_i > 0, (\underline{r} - \underline{r}^{n+1})_i \leq 0,\\
&\underline{u}_i = -1, \underline{u}_i^{n+1} > -1 && \Rightarrow && (\underline{u} - \underline{u}^{n+1})_i < 0, (\underline{r} - \underline{r}^{n+1})_i \geq 0,\\
&\underline{u}_i < 1, \underline{u}_i^{n+1} = 1 && \Rightarrow && (\underline{u} - \underline{u}^{n+1})_i < 0, (\underline{r} - \underline{r}^{n+1})_i \geq 0,\\
&\underline{u}_i > -1, \underline{u}_i^{n+1} = -1 && \Rightarrow && (\underline{u} - \underline{u}^{n+1})_i > 0, (\underline{r} - \underline{r}^{n+1})_i \leq 0.
\end{aligned}$$

Damit gilt $(\underline{u} - \underline{u}^{n+1})_i(\underline{r} - \underline{r}^{n+1})_i \leq 0$, also auch $(\underline{u} - \underline{u}^{n+1})^T(\underline{r} - \underline{r}^{n+1}) \leq 0$. ■

5.3.4 Implementierung des Algorithmus

Jetzt soll für den Algorithmus eine Formulierung gegeben werden, die es ermöglicht, die einzelnen $\underline{u}^{n+1}$ zu berechnen. Im symmetrischen Fall geschieht dies durch das Lösen lokaler Minimierungsprobleme. Um eine ähnliche Darstellung zu erhalten, wird die i-te Zeile der Gleichung (5.3.3) betrachtet und die Ungleichung (5.3.2) mit $\underline{y}$ mit $\underline{y}_j = \underline{u}_j^{n+1}$ für $i \neq j$ und $\underline{y}_i = \xi \in [-1; 1]$ beliebig getestet. Da M eine Diagonalmatrix ist, ergibt dies für $\underline{u}_i^{n+1}$ die Beziehung

$$M_{ii}\underline{u}_i^{n+1} + \tau A_{ii}\underline{w}_i^{n+1} = \alpha_i^{n+1}, \tag{5.3.8}$$
$$(\xi - \underline{u}_i^{n+1})\left(\gamma A_{ii}\underline{u}_i^{n+1} - M_{ii}\underline{w}_i^{n+1}\right) \geq (\xi - \underline{u}_i^{n+1})\beta_i^{n+1} \qquad \forall \xi \in [-1; 1]. \tag{5.3.9}$$

mit $\underline{u}_i^{n+1} \in [-1; 1]$ und

$$\alpha_i^{n+1} := b - \tau \sum_{j=1}^{i-1} A_{ij}\underline{w}_j^{n+1} - \tau \sum_{j=i+1}^{N} A_{ij}\underline{w}_j^n,$$
$$\beta_i^{n+1} := b - \gamma \sum_{j=1}^{i-1} A_{ij}\underline{u}_j^{n+1} - \gamma \sum_{j=i+1}^{N} A_{ij}\underline{u}_j^n.$$

Dabei hängen α_i^{n+1} und β_i^{n+1} nur von $\underline{u}^n$ und $\underline{u}_j^{n+1}$ mit $j < i$ ab. Also können die $(\underline{u}_i^{n+1}, \underline{w}_i^{n+1})^T$ nacheinander berechnet werden. Es ist noch zu klären, ob solche $(\underline{u}_i^{n+1}, \underline{w}_i^{n+1})^T$ immer existieren und wie sie zu berechnen sind. Da $A_{ii} > 0$ gilt, lässt sich Gleichung (5.3.8) darstellen als

$$\underline{w}_i^{n+1} = \frac{1}{\tau A_{ii}}(\alpha_i^{n+1} - M_{ii}\underline{u}_i^{n+1}).$$

Einsetzen in die Ungleichung (5.3.9) ergibt $\underline{u}_i^{n+1} \in [-1;1]$ mit

$$(\xi - \underline{u}_i^{n+1}) \underbrace{\left(\gamma A_{ii} + \frac{M_{ii}^2}{\tau A_{ii}}\right)}_{=Q} \underline{u}_i^{n+1} \geq (\xi - \underline{u}_i^{n+1}) \underbrace{\left(\beta_i^{n+1} + \frac{M_{ii}}{\tau A_{ii}} \alpha_i^{n+1}\right)}_{=P} \qquad \forall \xi \in [-1;1].$$

Nach Korollar 2.2.1 ist diese Ungleichung äquivalent zum Minimieren des Energiefunktionals $J(\xi) := \frac{1}{2} Q \xi^2 - P\xi$ in $[-1;1]$. Wegen $Q > 0$ ist J strikt konvex in $\mathbb{R}$. Also existieren nach Satz 2.2.3 eindeutige Minima von J in $\mathbb{R}$ und $[-1;1]$. Folglich gilt:

Bemerkung 5.3.1 *Die Iterationsfolge $(\underline{u}^n, \underline{w}^n)^T$ ist wohldefiniert.* ■

Sei $\hat{\underline{u}}_i^{n+1}$ das Minimum in $\mathbb{R}$, also $DJ(\hat{\underline{u}}_i^{n+1}) = 0$. Für $\hat{\underline{u}}_i^{n+1} \in [-1;1]$ gilt $\underline{u}_i^{n+1} = \hat{\underline{u}}_i^{n+1}$. Für $\hat{\underline{u}}_i^{n+1} > 1$ ist J in $[-1;1]$ streng monoton fallend, für $\hat{\underline{u}}_i^{n+1} < 1$ streng monoton wachsend. Somit ist $\underline{u}_i^{n+1}$ genau die Projektion von $\hat{\underline{u}}_i^{n+1}$ in $[-1;1]$. Zu gegebenem Startwert $\underline{u}^0$ sieht der Algorithmus damit wie folgt aus:

(1) Für $n = 0, ...$ bestimme $(\underline{u}^{n+1}, \underline{w}^{n+1})^T$ durch:

(2) Für $i = 1, ..., N$ bestimme $(\underline{u}_i^{n+1}, \underline{w}_i^{n+1})^T$ durch:

(3) Berechne α_i^{n+1}, β_i^{n+1}.

(4) Löse $\left(\gamma A_{ii} + \frac{M_{ii}^2}{\tau A_{ii}}\right) \hat{\underline{u}}_i^{n+1} = \left(\beta_i^{n+1} + \frac{M_{ii}}{\tau A_{ii}} \alpha_i^{n+1}\right)$.

(5) Setze $\underline{u}_i^{n+1} = \min\{\max\{\hat{\underline{u}}_i^{n+1}, -1\}, 1\}$.

(6) Setze $\underline{w}_i^{n+1} = \frac{1}{\tau A_{ii}} (\alpha_i^{n+1} - M_{ii} \underline{u}_i^{n+1})$.

Es ist noch anzumerken, dass die Berechnung von $(\underline{u}^{n+1}, \underline{w}^{n+1})^T$ nur $O(N)$ Operationen erfordert, da die Matrix A dünn besetzt ist. Ferner hat der Algorithmus den Vorteil, dass keine Parameter zu wählen sind.

5.4 Die Cahn-Hilliard Gleichung als Minimierungsproblem

Für lineare Variationsprobleme ist eine Vielzahl effizienter, also schneller, Lösungsalgorithmen bekannt. Beispiele sind etwa das vorkonditionierte CG-Verfahren für symmetrische Probleme und Varianten wie BICGSTAB für nicht symmetrische Probleme. Insbesondere Mehrgittermethoden ermöglichen eine schnelle Lösung linearer Probleme. Auch können die Methoden auf verschiedene Arten kombiniert werden. So ist etwa ein durch Mehrgittermethoden vorkonditioniertes CG-Verfahren möglich. Glatte nichtlineare Probleme können beispielsweise mit dem Newton-Verfahren gelöst werden, wobei für die linearen Teilprobleme schnelle lineare Löser verwendet werden. Auch ein direkter Mehrgitteransatz für glatte nichtlineare Probleme ist möglich.

Für nichtglatte nichtlineare Probleme ist die Auswahl an effizienten Lösungsverfahren nicht so groß. Ein wichtige Klasse nichtglatter nichtlinearer Probleme bilden Hindernisprobleme wie

die Cahn-Hilliard Gleichung. Operator-Splitting Methoden oder Relaxationsmethoden wie das Gauß-Seidel Verfahren zur Lösung solcher Probleme sind typischerweise sehr langsam. Auch die Ankopplung des Hindernisses mit Lagrange-Multiplikatoren (siehe [Glo84], [GLT81]) ist oft ineffizient. Die monotonen Mehrgitterverfahren (siehe [Kor97]) bieten eine Möglichkeit, nichtglatte nichtlineare Probleme sehr schnell zu lösen. Sie sind zum Beispiel auf Hindernisprobleme anwendbar, die sich als Minimierungsproblem darstellen lassen. Dabei wird ein konvexes Funktional in einer konvexen Menge minimiert. Wesentlich ist hierfür die Symmetrie und positive Definitheit des Problems.

Für nichtsymmetrische Hindernisprobleme ist kein ähnlicher Algorithmus bekannt. In diesem Abschnitt soll (CH) zu einem Minimierungsproblem umformuliert werden, um Mehrgittermethoden zur Lösung anwenden zu können. Dies führt allerdings auf eine zusätzliche Nebenbedingung. Das Problem stellt sich dann als Minimierungsproblem mit Box-Constraints, das heißt es ist $u(x) \in [a;b]$ gefordert, mit einer partiellen Differentialgleichung als Nebenbedingung. Diese lässt sich durch die Einführung eines Lagrange-Multiplikators an das Problem koppeln, was zu einem Sattelpunktproblem führt. Eine Einführung in die Theorie solcher restringierter Minimierungsaufgaben ist zum Beispiel in [Spe93] und [GK02] gegeben.

5.4.1 Aufstellen eines Energiefunktionals

Wie in Kapitel 5.1 gezeigt wurde, lässt sich (CH) symmetrisch formulieren. Allerdings ist die entstehende Systemmatrix dann nicht positiv definit. In diesem Abschnitt wird eine Art Energiefunktional für (CH) aufgestellt. Um das Vorgehen eleganter zu gestalten, wird eine Notation eingeführt, die u und w als eine Variable betrachtet. Im Folgenden sei $U = (u,w)^T$ und $V = (v_u, v_w)^T$. Die Indizierung $v_u, v_w, l_u, l_w, a_u, a_w$ der Testfunktionen, Funktionale und Bilinearformen soll keine Abhängigkeit von u und w ausdrücken, sondern dient lediglich der Orientierung. Mit dieser Notation lässt sich (CH) schreiben als:

(CH) Finde $U \in K^h \times S^h$ mit

$$\mathbf{a}(U, V-U) \geq \mathbf{l}(V-U) \qquad \forall V \in K^h \times S^h,$$

wobei gilt:

$$\begin{aligned}
\mathbf{a}(U,V) &:= \underbrace{\gamma(\nabla u, \nabla v_u) - (w, v_u)_h}_{=:a_u(U,v_u)} + \underbrace{(-(u, v_w)_h - \tau(\nabla w, \nabla v_w))}_{=:a_w(U,v_w)} \\
\mathbf{l}(V) &:= \underbrace{(u^{\text{alt}}, v_u)_h}_{=:l_u(v_u)} + \underbrace{(-(u^{\text{alt}}, v_w)_h)}_{=:l_w(v_w)}.
\end{aligned}$$

Die einzelnen Gleichungen in (CH) lassen sich durch das Testen mit $(v,w)^T$ und $(u, v+w)^T$ reproduzieren. Da die Bilinearform $\mathbf{a}(\cdot,\cdot)$ symmetrisch, aber nicht positiv-definit ist, kann nicht erwartet werden, dass das Lösen von (CH) äquivalent zum Minimieren eines Energiefunktionals

ist. Trotzdem wird nun das in „üblicher“ Weise für $V \in S^h \times S^h$ aufgestellte Energiefunktional

$$\begin{aligned}\widetilde{J}(V) :&= \frac{1}{2}\mathbf{a}(V,V) - \mathbf{l}(V)\\ &= \frac{1}{2}a_u(V,v_u) + \frac{1}{2}a_w(V,v_w) - l_u(v_u) - l_w(v_w)\\ &= \frac{1}{2}\gamma|v_u|_1^2 - \frac{1}{2}\tau|v_w|_1^2 - (v_u,v_w)_h - l_u(v_u) - l_w(v_w).\end{aligned}$$

betrachtet. Durch den negativen Term $-\frac{1}{2}\tau|v_w|_1^2$ ist dieses Funktional nicht nach unten beschränkt, besitzt also kein endliches, globales Minimum. Das Lösen von (CH) ist aber äquivalent zum Lösen von (CH) unter der Nebenbedingung:

(NB) Für $V = (v_u, v_w)^T \in S^h \times S^h$ gelte:

$$a_w(V,v) = l_w(v) \qquad \forall v \in S^h.$$

Definition 5.4.1 Nun wird der affine Unterraum von $S^h \times S^h$ und die Teilmenge von $K^h \times S^h$, auf denen (NB) gilt, definiert als

$$\begin{aligned}S_{\text{(NB)}} :&= \{V \in S^h \times S^h : a_w(V,v) = l_w(v) \forall v \in S^h\},\\ K_{\text{(NB)}} :&= \{V \in K^h \times S^h : a_w(V,v) = l_w(v) \forall v \in S^h\}.\end{aligned}$$

Wird $\widetilde{J}$ nur auf $S_{\text{(NB)}}$ betrachtet, so liefert das Testen von (NB) mit v_w und das Einsetzen in $\widetilde{J}$ für $V \in S_{\text{(NB)}}$

$$\widetilde{J}(V) = \frac{1}{2}a_u(V,v_u) - \frac{1}{2}a_w(V,v_w) - l_u(v_u) = \frac{1}{2}\gamma|v_u|_1^2 + \frac{1}{2}\tau|v_w|_1^2 - l_u(v_u).$$

Dieses Funktional wird nun auf dem ganzen Raum $S^h \times S^h$ betrachtet.

Definition 5.4.2 Für $V \in S^h \times S^h$ definiere

$$\begin{aligned}J(V) &:= \frac{1}{2}a_u(V,v_u) - \frac{1}{2}a_w(V,v_w) - l_u(v_u)\\ &= \frac{1}{2}\gamma|v_u|_1^2 + \frac{1}{2}\tau|v_w|_1^2 - (u^{\text{alt}},v_u)_h.\end{aligned}$$

5.4.2 Das Minimierungsproblem

Da die Lösung von (CH) in $K_{\text{(NB)}}$ liegt, hoffen wir, dass das Lösen von (CH) äquivalent zum Minimieren von $\widetilde{J}$ auf $K_{\text{(NB)}}$ und damit zu folgendem Problem ist:

($\widetilde{\text{M}}$) Finde $U \in K_{\text{(NB)}}$ mit

$$J(U) \le J(V) \qquad \forall V \in K_{\text{(NB)}}.$$

Es fällt jedoch auf, dass sowohl J als auch (NB) nur von ∇v_w und nicht von v_w abhängen. Ist also $U = (u, w)^T$ eine Lösung von ($\widetilde{\text{M}}$), so ist auch $(u, w + \text{const})^T$ eine Lösung von ($\widetilde{\text{M}}$). Folglich hat ($\widetilde{\text{M}}$) im Gegensatz zu (CH) keine eindeutig bestimmte Lösung und kann somit nicht äquivalent zu (CH) sein.

Um die Lösung von ($\widetilde{\text{M}}$) eindeutig zu machen, wird nun die Masse von w festgelegt. Da (CH) nur Massenerhaltung für u, nicht aber für w liefert, wird die Masse willkürlich auf 0 festgelegt. Dieses normierte w wird im Folgenden w_0 genannt. Die zusätzliche Nebenbedingung wird nun direkt in den Lösungsraum eingebaut.

Definition 5.4.3 Definiere den Lösungsraum als

$$K^0_{\text{(NB)}} := \{V \in K^h \times S^h_0 : a_w(V, v) = l_w(v) \forall v \in S^h\}.$$

Jetzt wird folgendes Minimierungsproblem mit normierter Masse für w_0 betrachtet:

(M) Finde $U_0 = (u, w_0)^T \in K^0_{\text{(NB)}}$ mit

$$J(U_0) \leq J(V) \qquad \forall V \in K^0_{\text{(NB)}}. \tag{5.4.1}$$

Es ist dabei zu beachten, dass nur w_0 in S^h_0 gesucht, jedoch die Nebenbedingung weiterhin mit v aus S^h getestet wird. Diese Modifikation führt zur eindeutigen Lösbarkeit von (M), wie folgender Satz zeigt:

Satz 5.4.1 *Die Lösung von (M) ist, falls existent, eindeutig bestimmt.*

BEWEIS. Sei $U_0 = (u, w_0)^T \in K^0_{\text{(NB)}}$ Lösung von (M) und $V = (v_u, v_w)^T \in K^0_{\text{(NB)}}$ beliebig. Nun gilt nach Korollar 2.2.1, da U_0 die konvexe, stetige, Gateaux-differenzierbare Funktion J minimiert und $K^0_{\text{(NB)}}$ konvex und wegen $\dim(S^h_0) < \infty$ abgeschlossen und ferner nicht leer ist:

$$\begin{aligned} 0 &\leq (DJ(U_0))(V - U_0) \\ &= \gamma(\nabla u, \nabla(v_u - u)) + \tau(\nabla w_0, \nabla(v_w - w_0)) - l_u(v_u - u). \end{aligned}$$

Also gilt

$$\begin{aligned} J(V) &= J(U_0) + \gamma|v_u - u|_1^2 + \tau|v_w - w_0|_1^2 \\ &\quad + \underbrace{\gamma(\nabla u, \nabla(v_u - u)) + \tau(\nabla w_0, \nabla(v_w - w_0)) - l_u(v_u - u)}_{\geq 0, \text{ da } U_0 \quad J \text{ minimiert}} \\ &\geq J(U_0) + \gamma|v_u - u|_1^2 + \tau|v_w - w_0|_1^2. \end{aligned}$$

Nun gilt, da U_0 und V (NB) erfüllen,

$$(u - v_u, v)_h + \tau(\nabla(w_0 - v_w), \nabla v) = 0 \qquad \forall v \in S^h.$$

Für $U_0 \neq V$ gibt es zwei Fälle zu untersuchen:
1. Gilt $u \neq v_u$, so folgt aus dieser Gleichung $\nabla w_0 \neq \nabla v$.

2. Gilt $w_0 \neq v_w$, so folgt, da die Masse von w_0 und v_w gleich ist, ebenfalls $\nabla w_0 \neq \nabla v$. Damit gilt immer $\tau|v_w - w_0|_1^2 > 0$, also $J(V) > J(U_0)$. ■

Zunächst ist unklar, in welchem Zusammenhang (M) und (CH) stehen. Im Allgemeinen werden jedoch die Lösungen von (M) und (CH) nicht übereinstimmen, da die Bedingung $(w, 1)_h = 0$ für Lösungen von (CH) nicht erfüllt sein muss. Bevor der genaue Zusammenhang von (M) und (CH) untersucht wird, wird das Minimierungsproblem mit Nebenbedingung als Sattelpunktproblem dargestellt.

5.4.3 Das Sattelpunktproblem

Nun wird das Minimierungsproblem (M) mit seiner Nebenbedingung (NB) in kanonischer Weise als Sattelpunktproblem dargestellt. Dazu wird die Nebenbedingung mittels eines neu eingeführten Lagrange-Multiplikators in ein geeignetes Lagrange-Funktional eingebracht. Die Grundidee eines solchen Vorgehens besteht darin, anzunehmen, dass die Lösung des restringierten Minimierungsproblems durch das unrestringierte Minimum dieses Lagrange-Funktionals für festen, richtig gewählten Lagrange-Multiplikator gegeben ist. Ob ein solcher „richtiger“ Lagrange-Multiplikator existiert, ist im Allgemeinen nicht klar. In diesem Abschnitt wird gezeigt, wie er sich für die diskrete Cahn-Hilliard Gleichung natürlich ergibt.

Definition 5.4.4 Für $V \in K^h \times S_0^h$ und $\mu \in S^h$ definiere

$$\begin{aligned} L(V, \mu) &:= J(V) + (a_w(V, \mu) - l_w(\mu)) \\ &= \frac{1}{2}\gamma|v_u|_1^2 + \frac{1}{2}\tau|v_w|_1^2 - (u^{\text{alt}}, v_u)_h - (v_u, \mu)_h - \tau(\nabla v_w, \nabla\mu) - (u^{\text{alt}}, \mu)_h. \end{aligned}$$

Das Sattelpunktproblem für L lautet dann:

(S) Finde $U_0 \in K^h \times S_0^h$ und $\lambda \in S^h$ mit

$$L(U_0, \mu) \leq L(U_0, \lambda) \leq L(V, \lambda) \qquad \forall V \in K^h \times S_0^h, \forall \mu \in S^h.$$

Nun wird der Zusammenhang der Probleme (CH), (S) und (M) untersucht.

Bemerkung 5.4.1 *Sei $(U_0, \lambda)^T = (u, w_0, \lambda)^T \in K^h \times S_0^h \times S^h$ eine Lösung von (S), dann gilt*

$$\begin{aligned} &\forall v_u \in K^h, \qquad \forall v_w \in S_0^h, \qquad \forall v \in S^h : \\ &0 \leq (D_u L(U_0, \lambda))\,(v_u - u) = \gamma(\nabla u, \nabla(v_u - u) - l_u(v_u - u) - (\lambda, v_u - u)_h \qquad (5.4.2) \\ &0 = (D_{w_0} L(U_0, \lambda))\,(v_w) = \tau(\nabla w_0, \nabla v_w) - \tau(\nabla\lambda, \nabla v_w) \qquad (5.4.3) \\ &0 = (D_\lambda L(U_0, \lambda))\,(v) = a_w(U_0, v) - l_w(v) = -(u, v)_h - \tau(\nabla w_0, \nabla v) - l_w(v) \qquad (5.4.4) \end{aligned}$$

BEWEIS. Wende Korollar 2.2.1 an. ■

Satz 5.4.2 *Sei $(U_0, \lambda)^T = (u, w_0, \lambda)^T \in K^h \times S_0^h \times S^h$ eine Lösung von (S), dann löst auch $U = (u, \lambda)^T \in K^h \times S^h$ (CH).*

BEWEIS. Da (5.4.3) nur mit ∇v_w getestet wird, gilt (5.4.3) auch für alle $v_w \in S^h$. Damit sind w und λ bis auf Konstanten identisch. Insbesondere gilt nach Einsetzen dieser Beziehung in (5.4.4)

$$a_w(U_0, v_w) = -(u, v_w)_h - \tau(\nabla\lambda, \nabla v_w) = l_w(v_w) \qquad \forall v_w \in S^h.$$

Ferner gilt mit (5.4.2)

$$a_u(U_0, v_u - u) \geq l_u(v_u - u) \qquad \forall v_u \in K^h.$$

Damit ist $U = (u, \lambda)^T$ Lösung von (CH). ■

Satz 5.4.3 *Sei $U = (u, w)^T \in K^h \times S^h$ Lösung von (CH), definiere dann $w_0 := w - \frac{(w,1)_h}{|\Omega|}$, $U_0 = (u, w_0)^T$ und $\lambda := w$. Dann löst $(U_0, \lambda)^T \in K^h \times S_0^h \times S^h$ das Sattelpunktproblem (S).*

BEWEIS. Wir weisen zunächst die Maximalität von L in λ nach. Dafür nutzen wir aus, dass U und somit auch U_0 die Nebenbedingung erfüllen. Für $\mu \in S^h$ gilt

$$\begin{aligned} L(U_0, \lambda) &= J(U_0) + a_w(U_0, \lambda) - l_w(\lambda) = J(U_0) \\ &= J(U_0) + a_w(U_0, \mu) - l_w(\mu) = L(U_0, \mu). \end{aligned}$$

Nun zeigen wir die Minimalität von L bezüglich U_0. Sei $V = (v_u, v_w)^T \in K^h \times S_0^h$, dann gilt $V = U_0 + V - U_0 = (u + v_u - u, w_0 + v_w - w_0)^T$ und

$$\begin{aligned} L(V, \lambda) &= L(U_0, \lambda) + J(V - U_0) + \gamma(\nabla u, \nabla(v_u - u)) \\ &\quad + \tau(\nabla w_0, \nabla(v_w - w_0)) + a_w(V - U_0, \lambda) \\ &= L(U_0, \lambda) + J(V - U_0) + \gamma(\nabla u, \nabla(v_u - u)) \\ &\quad + \underbrace{\tau(\nabla w_0, \nabla(v_w - w_0)) - \tau(\nabla(v_w - w_0), \nabla\lambda)}_{=0, \text{ da } \nabla w_0 = \nabla\lambda} - (v_u - u, \lambda)_h \\ &= L(U_0, \lambda) + \frac{1}{2}\gamma|v_u - u|_1^2 + \frac{1}{2}\tau|v_w - w_0|_1^2 \\ &\quad \underbrace{-l_u(v_u - u) + \gamma(\nabla u, \nabla(v_u - u)) - (v_u - u, \lambda)_h}_{\geq 0, \text{ da } \nabla w = \nabla\lambda \text{ und wegen 5.0.2 in (CH)}} \geq L(U_0, \lambda). \end{aligned} \tag{5.4.5}$$

Also ist $(U_0, \lambda)^T$ Sattelpunkt von (S). ■

Korollar 5.4.1 *Sei $U = (u, w)^T \in K^h \times S^h$ Lösung von (CH), definiere nun $w_0 := w - \frac{(w,1)_h}{|\Omega|}$, $U_0 = (u, w_0)^T$. Dann löst $U_0 \in K^0_{\text{(NB)}}$ das Minimierungsproblem (M).*

BEWEIS. Nach Satz 5.4.3 ist für $\lambda = w$ auch $(U_0, \lambda)^T$ Lösung von (S) und erfüllt insbesondere (NB). Sei nun $V = (v_u, v_w)^T \in K^0_{\text{(NB)}}$ beliebig, dann erfüllt auch V (NB). Damit gilt

$$\begin{aligned} J(U_0) &= J(U_0) + a_w(U_0, \lambda) - l_w(\lambda) = L(U_0, \lambda) \leq L(V, \lambda) \\ &= J(V) + a_w(V, \lambda) - l_w(\lambda) = J(V). \end{aligned} \tag{5.4.6}$$

■

Korollar 5.4.2 *Sei $U_0 = (u, w_0)^T \in K^0_{(\mathrm{NB})}$ Lösung von (M), dann existiert $m_w \in \mathbb{R}$, so dass $U = (u, w_0 + m_w)^T \in K^h \times S^h$ Lösung von (CH) ist.*

BEWEIS. Da die Lösung von (M) eindeutig bestimmt ist, gilt nach Korollar 5.4.1 für die Lösung von (M) $U_0 = (u, w_0)^T = (u, w - 1/|\Omega|(w, 1)_h)^T$, wobei $U = (u, w)^T$ die Lösung von (CH) ist. ■

5.5 Der Uzawa Algorithmus

Zur Lösung von Minimierungsproblemen unter Nebenbedingungen, beziehungsweise der Lösung von Sattelpunktproblemen, existieren viele Varianten des Uzawa-Algorithmus.

5.5.1 Der Uzawa Algorithmus für Minimierungsprobleme

Eine recht allgemeine Formulierung des Uzawa Algorithmus, zu finden bei [GLT81], sieht wie folgt aus:

Es seien zwei Hilberträume H und L, abgeschlossene, konvexe Mengen $M \subset H$ und $\Lambda \subset L$ und folgende Abbildungen gegeben:

(1) Eine stetige, koerzive Bilinearform $a : H \times H \to \mathbb{R}$,

(2) ein stetiges, lineares Funktional $l : H \to \mathbb{R}$,

(3) ein Energiefunktional $J : H \to \mathbb{R}$ mit $J(V) = \frac{1}{2}a(V, V) - l(V)$,

(4) eine lipschitz-stetige Abbildung $\Phi : M \to L$, für welche $V \mapsto (\mu, \Phi(V))_L$ für jedes $\mu \in L$ unterhalb-stetig und konvex auf H ist,

(5) das Lagrange-Funktional $L : M \times L \to \mathbb{R}$ mit $L(V, \mu) = J(V) + (\mu, \Phi(V))_L$.

Gesucht ist nun eine Lösung des Problems:

(Min) Finde $u \in M$ mit

$$J(U) + \sup_{\mu\in\Lambda}(\mu, \Phi(U))_L \leq J(V) + \sup_{\mu\in\Lambda}(\mu, \Phi(V))_L \qquad \forall V \in M.$$

Das dazugehörige Sattelpunktproblem lautet:

(Sat) Finde $U \in M$ und $\lambda \in \Lambda$ mit

$$L(U, \mu) \leq L(U, \lambda) \leq L(V, \lambda) \qquad \forall \mu \in \Lambda \quad \forall V \in M.$$

Sei $P_\Lambda : L \to \Lambda$ die Projektion von L nach Λ, dann ist der Uzawa Algorithmus zur iterativen Lösung von (Min) zu gegebenem Startwert $\lambda_0 \in \Lambda$ wie folgt definiert:

(1) Für $n = 0, ...$ bestimme U^n und λ^{n+1} durch:

(2) Löse für festes λ^n das Minimierungsproblem

$$U^n \in M: \qquad L(U^n, \lambda^n) \leq L(V, \lambda^n) \qquad \forall V \in M.$$

(3) Setze $\lambda_{n+1} = P_\Lambda(\lambda^n + \rho_n \Phi(U^n))$.

Satz 5.5.1 *Wenn das Sattelpunktproblem (Sat) eine Lösung besitzt und die obigen Bedingungen (1) - (5) erfüllt sind, existieren Konstanten $0 < \alpha_0 \leq \alpha_1$, so dass für $0 < \alpha_0 \leq \rho_n \leq \alpha_1$ die Folge U^n stark gegen die Lösung U von (Min) konvergiert.*

BEWEIS. Der Beweis ist in [GLT81] gegeben. ■

Meist ist der Uzawa Algorithmus so formuliert, dass für festes λ unrestringierte Minimierungsprobleme im ganzen Raum H zu lösen sind. Dabei wird davon ausgegangen, dass etwaige Nebenbedingungen durch den Langrange-Multiplikator an das Problem gekoppelt werden. Die hier dargestellte Version besitzt für die Anwendung auf die Cahn-Hilliard Gleichung den Vorteil, die Nebenbedingung $u \in K^h$ getrennt von (NB) zu behandelt. Da Minimierungsprobleme mit Box-Constraints ähnlich effektiv durch Mehrgittermethoden gelöst werden können wie lineare Probleme, ist es nicht notwendig, derartige Nebenbedingungen mittels eines Lagrange-Multiplikators an das Problem zu koppeln.

5.5.2 Die Teilprobleme im Uzawa Algorithmus

Der Uzawa Algorithmus besteht aus einer Iteration für den Lagrange-Multiplikator, bei der in jedem Schritt das Lagrange-Funktional ohne Nebenbedingung für festen Lagrange-Multiplikator minimiert wird. Somit ist eine wesentliche Voraussetzung zur Anwendung des Algorithmus die eindeutige Lösbarkeit dieser Probleme. Diese ist durch die Koerzivität der Bilinearform gegeben. Im Fall des Minimierungsproblems (M) und seiner Sattelpunktformulierung (S) führt der Algorithmus auf Teilprobleme der Art:

Finde $U^k \in K^h \times S_0^h$ mit

$$L(U^k, \mu) \leq L(V, \mu) \qquad \forall V \in K^h \times S_0^h. \tag{5.5.1}$$

Ist nun etwa $\mu = u^{n-1}$, so gilt

$$L(V, \mu) = \frac{1}{2}\gamma |v_u|_1^2 + \frac{1}{2}\tau |v_w|_1^2 - \tau(\nabla v_w, \nabla \mu) - l_w(\mu).$$

Da L dann nur von ∇v_u abhängt und v_u nicht in einem mit $(v_u, 1)_h = 0$ normierten Raum betrachtet wird, muss die Lösung von (5.5.1) nicht eindeutig bestimmt sein. Das ist auch insofern klar, als die Bilinearform, die J und L erzeugt, nicht koerziv ist.

5.5.3 Erzwingen von Massenerhaltung

Nun werden (M) und (S) abermals modifiziert, um die eindeutige Lösbarkeit dieser Teilprobleme zu gewährleisten. Lösungen von (M) und (S) erfüllen (NB), was durch Testen mit $v \equiv 1 \in S^h$ insbesondere zu

$$\int_\Omega u(x)dx = (u,1)_h = (u^{\text{alt}},1)_h =: m$$

führt. Die Masse m von u und u^{alt} ist dabei bekannt. Somit kann der Term $\frac{1}{2}((u,1)_h - m)^2$ zum Energiefunktional addiert werden, ohne die Lösung des Minimierungsproblems zu verändern, da er bei Einhalten der Massenerhaltung zu Null wird. Liegt keine Massenerhaltung vor, wird dieser Term echt positiv und „bestraft“ damit die „falsche“ Masse.

Definition 5.5.1 Für $V = (v_u, v_w)^T \in S^h \times S^h$ und $\mu \in S^h$ definiere mit $\theta > 0$

$$\begin{aligned} h_\theta(v_u) &= \frac{1}{2}\theta((v_u,1)_h - m)^2 = \frac{1}{2}\theta(v_u,1)_h^2 - \theta m(v_u,1)_h + \frac{1}{2}\theta m^2 \\ J_\theta(V) &= J(V) + h_\theta(v_u) - \frac{1}{2}\theta m^2, \\ L_\theta(V,\mu) &= J_\theta(V) + (a_w(V,\mu) - l_w(\mu)) = L(V,\mu) + h_\theta(v_u) - \frac{1}{2}\theta m^2. \end{aligned}$$

Minimierungs und Sattelpunktproblem lauten für diese Funktionale:

(M_θ) Finde $U_0 = (u, w_0)^T \in K^0_{(\text{NB})}$ mit

$$J_\theta(U_0) \le J_\theta(V) \qquad \forall V \in K^0_{(\text{NB})}.$$

(S_θ) Finde $U_0 = (u, w_0)^T \in K^h \times S^h_0$ und $\lambda \in S^h$ mit

$$L_\theta(U_0,\mu) \le L_\theta(U_0,\lambda) \le L_\theta(V,\lambda) \qquad \forall V \in K^h \times S^h_0, \forall \mu \in S^h.$$

Im Wesentlichen wird durch dieses Vorgehen die Bilinearform

$$(v,1)_h(u,1)_h = \int_\Omega v(x)dx \int_\Omega u(x)dx$$

zur Bilinearform von J addiert. Der lineare und konstante Teil sorgt dafür, dass der Term bei Massenerhaltung verschwindet. Da dieser Term von $(v,1)_m$ abhängt, ist zu hoffen, dass er zur Koerzivität von J_θ führt. Dies wird im nächsten Abschnitt gezeigt.

Satz 5.5.2 *Es gilt für* $(U_0,\lambda)^T \in K^h \times S^h_0 \times S^h$

$$(U_0,\lambda)^T \text{ löst (S)} \Leftrightarrow (U_0,\lambda)^T \text{ löst } (\text{S}_\theta).$$

BEWEIS. '$\Leftarrow$': Sei $(U_0, \lambda)^T$ Lösung von (S_θ), dann erhält man nach dem Ableiten von L_θ nach λ und dem Einsetzen von $v \equiv 1$ die Massenerhaltung $m = (u, 1)_h$. Damit gilt

$$L(U_0, \mu) = L_\theta(U_0, \mu) + \frac{1}{2}\theta m^2 \leq L_\theta(U_0, \lambda) + \frac{1}{2}\theta m^2 = L(U_0, \lambda). \tag{5.5.2}$$

Ferner gilt

$$\left(D_u h_\theta(u)\right)(v) = \theta(u, 1)_h(v_u - u, 1)_h - \theta m(v_u - u, 1)_h = 0.$$

Damit gelten (5.4.2) - (5.4.4) aus Bemerkung 5.4.1 auch für L_θ, und wir erhalten ähnlich dem Vorgehen (5.4.5) in Satz 5.4.3

$$L(V, \lambda) \geq L(U_0, \lambda).$$

'$\Rightarrow$': Sei $(U_0, \lambda)^T$ Lösung von (S), dann erhält man aus (5.4.4) in Bemerkung 5.4.1 ebenfalls die Massenerhaltung $m = (u, 1)_h$ und damit analog zu (5.5.2) die Maximalität bezüglich λ. Wegen $h_\theta(v_u) \geq 0$ folgt auch die Minimalität bezüglich U sofort:

$$\begin{aligned} L_\theta(U_0, \lambda) &= L(U_0, \lambda) - \frac{1}{2}\theta m^2 \leq L(V, \lambda) - \frac{1}{2}\theta m^2 \\ &\leq L(V, \lambda) - \frac{1}{2}\theta m^2 + h_\theta(v_u) = L_\theta(V, \lambda). \end{aligned}$$

■

Satz 5.5.3 *Es gilt für $U_0 \in K^0_{(\mathrm{NB})}$:*

$$U_0 \text{ löst (M)} \Leftrightarrow U_0 \text{ löst } (\mathrm{M}_\theta).$$

BEWEIS. Sei $V \in K^0_{(\mathrm{NB})}$ beliebig, dann erfüllt V (NB). Somit gilt $(v, 1)_h = m$ und

$$J(V) = J(V) + h_\theta(v_u) = J_\theta(V) + \frac{1}{2}\theta m^2.$$

Also sind J und J_θ auf $K^0_{(\mathrm{NB})}$ bis auf eine Konstante identisch und die Äquivalenz ist trivial erfüllt. ■

Die Beziehungen der verschiedenen Formulierungen der Gleichung werden in folgendem Satz zusammengefasst.

Satz 5.5.4 *Folgende Aussagen sind für $U = (u, w)^T \in K^h \times S^h$, $U_0 = (u, w_0)^T \in K^h \times S^h_0$ und $\lambda \in S^h$ äquivalent:*

(1) *U ist Lösung von (CH).*

(2) *$(U_0, \lambda)^T$ ist Lösung von (S).*

(3) *$(U_0, \lambda)^T$ ist Lösung von (S_θ).*

(4) *U_0 ist Lösung von (M).*

(5) *U_0 ist Lösung von (M_θ).*

Dabei gilt $w_0 = w - 1/|\Omega|(w, 1)_h$ und $\lambda = w$.

5.5.4 Der Uzawa Algorithmus für die diskrete Cahn-Hilliard Gleichung

Nun werden die Eigenschaften von (M_θ) und (S_θ) gezeigt, die es ermöglichen, den Uzawa Algorithmus auf die zu (CH) äquivalenten Probleme anzuwenden.

Definition 5.5.2 Es sei $H := S^h \times S_0^h$ und für $U = (u, w)^T, V = (v_u, v_w)^T \in H$ die Bilinearform a_θ und das Funktional l_θ definiert durch

$$\begin{aligned} a_\theta(U, V) &:= \gamma(\nabla u, \nabla v_u) + \tau(\nabla w, \nabla v_w) + \theta(u, 1)_h(v_u, 1)_h, \\ l_\theta(V) &:= l_u(v_u) + \theta m(v_u, 1)_h. \end{aligned}$$

Bemerkung 5.5.1 *H ist ein Hilbertraum und $K^h \times S_0^h \subset H$ ist konvex und abgeschlossen.*

BEWEIS. H ist als Produkt der Hilberträume S^h und S_0^h selbst Hilbertraum mit dem Skalarprodukt $(U, V)_H = (u, v_u)_h + (\nabla u, \nabla v_u) + (\nabla w, \nabla v_w)$ und der Norm $\|V\|_H = \|v_u\|_h + |v_u|_1 + |v_w|_1$. Konvexität und Abgeschlossenheit von $K^h \times S_0^h$ folgen aus der Konvexität und Abgeschlossenheit von $K^h \subset S^h$ und $S_0^h \subset S_0^h$. ■

Lemma 5.5.1 *J_θ lässt sich darstellen als*

$$J_\theta(V) = \frac{1}{2} a_\theta(V, V) - l_\theta(V) \qquad \forall V \in H$$

und es gilt

(1) *l_θ ist stetig auf H,*

(2) *a_θ ist stetig auf $H \times H$,*

(3) *a_θ ist koerziv, d.h. es existiert $\alpha > 0$ mit $a_\theta(V, V) \geq \alpha \|V\|_H^2 \qquad \forall V \in H$.*

BEWEIS. Die Darstellung von J_θ ergibt sich direkt aus der Definition.

(1) Die Stetigkeit folgt aus der Linearität auf dem endlichdimensionalen Raum H.

(2) a_θ ist bilinear auf dem endlichdimensionalen Raum H, also auch stetig.

(3) Nach der diskreten Poincaré-Ungleichung gilt $\|v\|_h \leq C\left(|(v, 1)_h| + |v|_1\right)$ für $v \in S^h$, also auch

$$\|v\|_h^2 \leq C^2 \left(|(v, 1)_h| + |v|_1\right)^2 \leq 3C^2 \left((v, 1)_h^2 + |v|_1^2\right).$$

Damit gilt

$$\begin{aligned} a_\theta(V, V) &\geq \min\{\gamma/2, \theta\} \left(2|v_u|_1^2 + (v_u, 1)_h^2\right) + \tau |v_w|_1^2 \\ &\geq \frac{\min\{\gamma/2, \theta\}}{3C^2} \|v_u\|_h^2 + \min\{\gamma/2, \theta\} |v_u|_1^2 + \tau |v_w|_1^2 \\ &\geq \underbrace{\min\left\{\frac{\gamma}{6C^2}, \frac{\theta}{3C^2}, \frac{\gamma}{2}, \theta, \tau\right\}}_{=\alpha} \left(\|v_u\|_h^2 + |v_u|_1^2 + |v_w|_1^2\right) = \alpha \|V\|_H^2. \end{aligned}$$

■

Lemma 5.5.2 *Es existiert ein lipschitz-stetiges $\Phi : H \to S^h$, so dass für jedes $V \in H$*

$$(\mu, \Phi(V))_{S^h} = a_w(V, \mu) - l_w(\mu) \qquad \forall \mu \in S^h \tag{5.5.3}$$

gilt. Für jedes $\mu \in S^h$ ist die Abbildung $V \mapsto (\mu, \Phi(V))_{S^h}$ konvex und stetig auf H. Ferner gilt für Φ

$$L_\theta(V, \mu) = J_\theta(V) + (\mu, \Phi(V))_{S^h} \qquad \forall \mu \in S^h \quad \forall V \in H.$$

BEWEIS. Sei $V \in H$ beliebig und fest gewählt, dann ist

$$a_w(V, \cdot) - l_w(\cdot) : S^h \to \mathbb{R}$$

linear und stetig auf dem Hilbertraum S^h. Also existiert nach dem Darstellungssatz von Fréchet-Riesz genau ein $\Phi(V) \in S^h$ mit (5.5.3). Da ein solches $\Phi(V)$ für jedes $V \in H$ existiert und eindeutig ist, ist $\Phi : H \to S^h$ eine wohldefinierte Abbildung. Insbesondere ist sie affin-linear und stetig, also auch lipschitz-stetig. Damit ist auch die Abbildung

$$V \mapsto \Phi(V) \mapsto (\mu, \Phi(V))_{S^h}$$

für jedes μ als Komposition einer affin-linearen, stetigen und einer linearen, stetigen Abbildung konvex und stetig. Die Darstellung von L_θ folgt sofort aus der Definition von Φ. Für die Folge U^n ist dabei die Wahl des Skalarprodukt $(\cdot, \cdot)_{S^h}$ auf S^h unerheblich, sie führt lediglich zu unterschiedlichen Folgen λ^n. ■

Nun wird $M := K^h \times S_0^h \subset H$ und $\Lambda := S^h \subset S^h = L$ gesetzt und der Uzawa Algorithmus angewandt, um (M_θ) zu lösen.

Satz 5.5.5 *Die vom Uzawa Algorithmus definierte Folge U^n konvergiert in H stark gegen die Lösung U von (M_θ).*

BEWEIS. Wegen der Äquivalenz zu (CH) hat (S_θ) eine Lösung. Mit Bemerkung 5.5.1, Lemma 5.5.1 und Lemma 5.5.2 sind nun alle Voraussetzungen für Satz 5.5.1 erfüllt. Also konvergiert U^n gegen ein U_{Uzawa} mit

$$J_\theta(U_{\text{Uzawa}}) + \sup_{\mu \in \Lambda} (\mu, \Phi(U_{\text{Uzawa}}))_{S^h} \leq J_\theta(V) + \sup_{\mu \in \Lambda} (\mu, \Phi(V))_{S^h} \qquad \forall V \in M. \tag{5.5.4}$$

Da wir Φ auf dem ganzen Raum $S^h = \Lambda$ betrachten, gilt für jedes $V \in H$

$$\sup_{\mu \in \Lambda} (\mu, \Phi(V))_{S^h} = \begin{cases} 0 & \text{falls } V \text{ (NB) erfüllt,} \\ \infty & \text{sonst} \end{cases}$$

also ist (5.5.4) zu (M_θ) äquivalent und es gilt $U_{\text{Uzawa}} = U$. ■

5.5.5 Implementierung des Algorithmus

In diesem Abschnitt wird der Uzawa Algorithmus für die diskrete Cahn-Hilliard Gleichung algebraisch formuliert. Ferner wird auf einige Besonderheiten bei der Implementierung eingegangen.

Im Schritt (2) des Uzawa Algorithmus ist $L(u^n, w_0^n, \lambda^n)$ für festes λ^n in $K^h \times S_0^h$ zu minimieren. Dies führt nach Korollar 2.2.1 auf eine Variationsungleichung der Art

$$(D_{u^n} L_\theta(u^n, w_0^n, \lambda^n), v_u - u^n)_{S^h} \geq 0 \qquad \forall v_u \in K^h, \tag{5.5.5}$$

$$\left(D_{w_0^n} L_\theta(u^n, w_0^n, \lambda^n), v_w\right)_{S^h} = 0 \qquad \forall v_w \in S_0^h. \tag{5.5.6}$$

Da in L_θ keine Terme vorkommen, die sowohl u als auch w enthalten, können beide Probleme unabhängig voneinander gelöst werden. Das Skalarprodukt $(\cdot,\cdot)_{S^h}$ sei hier so gewählt, dass es in der Koeffizientendarstellung dem Euklidischen Skalarprodukt auf $\mathbb{R}^N$ entspricht. Damit ergibt die Definition für die Funktion Φ aus Lemma 5.5.2 folgende Darstellung $\underline{\Phi}$ in der Basis von S^h:

$$\underline{\Phi}(\underline{v}_u, \underline{v}_w) = b - M\underline{v}_u - \tau A \underline{v}_w.$$

Für die Variationsprobleme (5.5.5) und (5.5.6) ergibt die Basisdarstellung

$$(\underline{v} - \underline{u}^n)^T(\gamma A + \theta\widehat{M})\underline{u}^n \geq (\underline{v} - \underline{u}^n)^T(b + M\underline{\lambda}^n + \theta m M\underline{1}) \qquad \forall \underline{v} \in \underline{K}^h, \tag{5.5.7}$$

$$\tau A \underline{w}_0^n = \tau A \underline{\lambda}^n. \tag{5.5.8}$$

Dabei ist $\widehat{M} := M\underline{1}\underline{1}^T M$ die Matrix, die die Bilinearform $(x,1)_h(y,1)_h$ darstellt. Sie ist mit $\widehat{M}_{ij} = M_{ii}M_{jj}$ vollbesetzt. Die Aktualisierung des Lagrange-Multiplikators in Schritt (3) lässt sich darstellen als

$$\underline{\lambda}^{n+1} = \underline{\lambda}^n + \rho_n(b - M\underline{u}^n - \tau A \underline{w}_0^n).$$

Da λ^n und w_0^n nach (5.5.8) bis auf eine Konstante übereinstimmen und in Schritt (3) nur ∇w_0^n benötigt wird, kann nach dem Einsetzen von (5.5.8) auf die Berechnung von w_0^n verzichtet werden. Dies entspricht der Tatsache, dass für die Lösung des Sattelpunktproblems $\lambda = w$ gilt und w_0 nur bis auf eine Konstante mit w übereinstimmt. Damit lässt sich der Uzawa Algorithmus zu gegebenem Startwert $\underline{\lambda}^0$ darstellen als:

(1) Für n=0,... bestimme $\underline{u}^n$ und $\underline{\lambda}^{n+1}$ durch:

(2) Löse die Variationsungleichung $\underline{u}^n \in \underline{K}^h$:

$$(\underline{v} - \underline{u}^n)^T(\gamma A + \theta\widehat{M})\underline{u}^n \geq (\underline{v} - \underline{u}^n)^T(b + M\underline{\lambda}^n + \theta m M\underline{1}) \qquad \forall \underline{v} \in \underline{K}^h.$$

(3) Setze $\underline{\lambda}^{n+1} = \underline{\lambda}^n + \rho_n(b - M\underline{u}^n - \tau A \underline{w}_0^n)$.

Dabei lässt sich für die durch $(\gamma A + \theta\widehat{M})$ gegebene symmetrische Bilinearform analog zu Lemma 5.5.1 Koerzivität zeigen. Somit kann die Variationsungleichung in Schritt (2) zum Beispiel mit dem monotonen Mehrgitterverfahren gelöst werden. Die voll besetzte Matrix $\widehat{M}$ stellt dabei kein Problem dar. Matrixmultiplikationen mit voll besetzten Matrizen benötigen normalerweise

$O(N^2)$ Operationen, lassen sich hier aber durch $\widehat{M}x = (M\underline{1})((\underline{1}^T M)x)$ auswerten, also im Wesentlichen durch das Skalarprodukt von $\underline{1}^T M = (M_{11}, ..., M_{NN})$ und x. Soll das Gauß-Seidel Verfahren als Glätter im monotonen Mehrgitterverfahren angewandt werden, so sind Ausdrücke der Form

$$\alpha_i = \sum_{j=1}^{i-1} \widehat{M}_{ij} x_j^{n+1} + \sum_{j=i+1}^{N} \widehat{M}_{ij} x_j^n$$

nacheinander für $i = 1, ..., N$ auszuwerten. Wegen $\widehat{M}_{ij} = M_{ii} M_{jj}$ erhält man durch Ausklammern von M_{ii}

$$\alpha_{i+1} = M_{i+1,i+1} \left(\frac{\alpha_i}{M_{ii}} + M_{ii} x_i^{n+1} - M_{i+1,i+1} x_{i+1}^n \right).$$

Somit ist die Auswertung mit $O(1)$ Operationen und ein Gauß-Seidel Schritt mit $O(N)$ Operationen durchführbar. Es ist jedoch darauf hinzuweisen, dass für den Algorithmus ein Parameter ρ_n gewählt werden muss, der nicht nur quantitativ, sondern auch qualitativ die Konvergenz beeinflusst. Aus dem Konvergenzbeweis in [GLT81] geht hervor, dass

$$2\alpha\rho_n + C_1^2 \rho_n^2 > 0$$

also $\rho_n \in (0; 2\alpha/C_1^2)$ gelten muss, um Konvergenz zu gewährleisten. Dabei ist C_1 die Lipschitz-Konstante von Φ und α die Koerzivitätskonstante der Bilinearform. Da α hier von der Konstanten in der Poincaré-Ungleichung abhängt (siehe Beweis von Lemma 5.5.1), sind die genauen Konstanten im Allgemeinen unzugänglich. Es lässt sich aber immerhin feststellen, dass der Algorithmus für ρ_n klein genug konvergiert.

6 Numerische Experimente

In diesem Abschnitt werden der Gauß-Seidel Algorithmus und der Uzawa Algorithmus auf einige Beispiele angewandt und miteinander verglichen. Es wird sich zeigen, dass die Algorithmen ein höchst unterschiedliches Konvergenzverhalten zeigen.

Der Splitting-Algorithmus von Lions-Mercier wird dem nicht gegenüber gestellt. Es ist dort in jedem Iterationsschritt ein gestörtes lineares Sattelpunktproblem zu lösen, zum Beipiel mit dem Uzawa Algorithmus. Die Nebenbedingung ist dabei die gleiche wie im Sattelpunktproblem (S). Nur die Minimierungsprobleme sind in diesem Fall unrestringiert. Da sich symmetrische Hindernisprobleme mit dem monotonen Mehrgitterverfahren jedoch ähnlich schnell lösen lassen wie lineare, ist der Aufwand für einen Iterationsschritt vergleichbar mit dem, den das Lösen von (CH) mit dem Uzawa Algorithmus erfordert. In der Arbeit [BBG99] von Barrett, Blowey und Garcke wird, um dies zu umgehen, die diskrete Cosinus Transformation zur Lösung der Gleichung verwendet. Sie ist für derartige Gleichungssysteme in der Arbeit [BE92] von Blowey und Elliott beschrieben. Jedoch ist ihre Anwendung auf quadratische, uniforme Gitter beschränkt, weshalb die Verwendung als allgemeiner Löser ausscheidet.

6.1 Implementierung

Es wurde bereits bei der Vorstellung der Algorithmen darauf eingegangen, wie sich die anfallenden Operationen sinnvoll algebraisch durchführen lassen. Um die Algorithmen testen zu können, wurden sie in einem selbst verfassten Programm in C implementiert.

Für die Lösung der Variationsungleichung im Uzawa Algorithmus wurde ein monotones Mehrgitterverfahren angewandt. Den Kern des Programms bildet die *grid*-Struktur. In ihr sind die Knoten und Elemente der Triangulierung als verkettete Liste enthalten. Jeder Knoten speichert dabei in seiner Struktur die zu ihm gehörigen Werte der einzelnen Vektoren und in einem Array die zu ihm gehörigen Matrixeinträge. Alle Matrix- und Vektoroperationen werden nun mittels eines Durchlaufens dieser Listen gestaltet. Die Datenstruktur besitzt den Vorteil, leicht auf die Speicherung einer Gitterhierarchie erweiterbar zu sein. Die *multigrid*-Struktur beinhaltet eine solche Gitterhierarchie. In ihr sind die Gitter der verschiedenen Ebenen ebenfalls durch eine verkette Liste gespeichert.

6.2 Resultate

Um das Verhalten der Algorithmen zu untersuchen, wurden zum einen Folgen von Problemen in der Zeit gelöst. Da jedoch alle Algorithmen die Probleme in jedem Zeitschritt einzeln lösen, wurde besonderes Augenmerk auf die Lösung dieser Teilprobleme zu einem Zeitpunkt gelegt.

Alle Experimente wurden auf dem Einheitsintervall oder dem Einheitsquadrat mit uniformen Gittern durchgeführt. Als Startwert wurde jeweils eine durch zufällige Werte leicht gestörte Nullfunktion ($u_0(x) < \frac{1}{50}$) verwendet. Dies entspricht einer gleichmäßigen Durchmischung der

beiden Spezies mit kleinen Unregelmäßigkeiten, an denen sich die Separation entwickelt. In Abbildung 6.4.1 sind die Zeitpunkte $n = 0$, 10, 20, 30, 50 und 500 der Lösung für das eindimensionale Problem mit $\gamma = 0.001$, $\tau = 0.001$ und mit $h = \frac{1}{64}$ abgebildet. Es ist zu beobachten, dass in der Anfangsphase sehr schnell eine Separation eintritt und die Evolution mit der Zeit immer langsamer wird. In Abbildung 6.4.2 sind die Lösungen für das zweidimensionale Problem mit $\gamma = 0.001$, $\tau = 0.001$ und mit $h = \frac{1}{32}$ zu den Zeitpunkten $n = 0$, 20, 40, 80, 160 und 500 dargestellt. Auch hier ist nur zu Beginn ein Separationsprozess festzustellen. Im späteren Verlauf ist die Evolution dagegen weitestgehend von einer Verringerung der Krümmung des Interfaces geprägt. Ähnliche Ergebnisse erbrachten auch Rechnungen mit anderem Parameter γ und anderen Startwerten.

Als Startwert für die Iteration wurde jeweils die Lösung aus dem vorherigen Zeitschritt verwendet. Da sich zu Beginn die Lösungen aufeinanderfolgender Zeitschritte stärker unterscheiden, benötigten die Algorithmen dort deutlich mehr Iterationsschritte. Deshalb wurde zum Vergleich der Geschwindigkeit das Teilproblem im ersten Zeitschritt genauer untersucht. Zur Kontrolle der Konvergenzgeschwindigkeit der Algorithmen wurde eine Referenzlösung u bestimmt. Dazu wurde die Iteration mit dem Gauß-Seidel Verfahren ausgeführt, bis eine Iterierte erreicht wurde, für die beide Verfahren nur noch Korrekturen, deren Norm kleiner als 10^{-15} ist, ausführten. Da beide Verfahren konvergieren, wurde diese Iterierte als hinreichend genaue Näherung der exakten Lösung betrachtet. Während der Iteration wurden drei Größen verfolgt:

(1) Die $\|\cdot\|_h$-Norm der Differenz zur Referenzlösung: $\|u - u^n\|_h$,

(2) Die $\|\cdot\|_h$-Norm der Korrekturen: $\|u^{n+1} - u^n\|_h$,

(3) Die $\|\cdot\|_h$-Norm des Defektes in Gleichung (4.3.1), also der Nebenbedingung in (S).

Die Iteration wurde jeweils beendet, wenn $\|u-u^n\|_h < tol = 10^{-10}$ erreicht wurde. Der Verlauf dieser Größen ist in Abbildung 6.4.3 - 6.4.5 für 1-D und in 6.4.6 - 6.4.8 für 2-D dargestellt. In allen folgenden Abbildungen stellt die durchgehende Linie den Gauß-Seidel und die gestrichelte Linie den Uzawa Algorithmus dar. Alle drei Größen verhalten sich sehr ähnlich. Es fällt auf, dass sie für großes h beim Gauß-Seidel Verfahren deutlich schneller abfallen. Für kleines h wird die Gauß-Seidel Iteration jedoch deutlich schlechter und die Korrekturen beginnen zu oszillieren.

Stellt man die Anzahl der Iterationsschritte gegenüber, die die Algorithmen benötigen, um unter die vorgegebene Toleranzschranke zu fallen, so erkennt man in Abbildung 6.4.9, dass diese mit kleiner werdendem h beim Gauß-Seidel Algorithmus rapide wächst und beim Uzawa Algorithmus deutlich leichter ansteigt.

Es ist allerdings anzumerken, dass für den Uzawa Algorithmus der Parameter ρ gewählt werden muss und ein zu groß gewähltes ρ Konvergenz verhindert, wohingegen ein zu kleines sie sehr verlangsamt. Der Parameter θ hat jedoch keinen qualitativen Einfluss auf die Konvergenz und wurde immer gleich 1 gesetzt. In den präsentierten Beispielen wurde ρ empirisch durch Ausprobieren möglichst groß und während der Iteration konstant gewählt. Auch der Aufwand in einem Iterationsschritt ist nur schlecht vergleichbar. Während im Gauß-Seidel Verfahren $O(n)$ Operationen anfallen, ist im Uzawa Algorithmus jeweils ein symmetrisches Hindernisproblem zu lösen. Die Experimente haben gezeigt, dass die Anzahl der notwendigen Mehrgitterschritte

mit kleinem h kaum wächst. Um den Aufwand zu verringern, wurde auch eine inexakte Variante des Uzawa Algorithmus, bei der jeweils höchstens drei Mehrgitterschritte ausgeführt wurden, betrachtet. Abbildung 6.4.10 stellt die Anzahl der Iterationsschritte, die dieses Verfahren benötigte, dem Gauß-Seidel Verfahren gegenüber. Es zeigt sich, dass die Geschwindigkeit des Uzawa Algorithmus unter dieser Vereinfachung kaum leidet.

6.3 Fazit

Die Experimente haben gezeigt, dass beide Verfahren nur sehr langsam konvergieren. Beim Uzawa Algorithmus sind dazu auch noch Parameter zu wählen, damit überhaupt Konvergenz vorliegt. Er besitzt gegenüber dem Gauß-Seidel Verfahren jedoch den großen Vorteil, mit kleiner werdender Schrittweite nicht signifikant langsamer zu werden. Auch kann der Aufwand drastisch reduziert werden, wenn die Probleme nicht exakt gelöst werden. Dieses Ergebnis ist allerdings rein empirisch. Zwar wurden in verschiedenen Arbeiten (siehe zum Beispiel [CH03]) inexakte Varianten des Algorithmus untersucht, jedoch nur für den Fall, dass die Teilprobleme unrestringierte Minmierungsprobleme sind. Übliche Techniken, um den Uzawa Algorithmus zu beschleunigen, wie etwa die Augmented-Lagrange Verfahren (siehe [FG83]), lassen sich aufgrund der besonderen Form der Nebenbedingung hier nicht direkt anwenden. Es ist bisher offen, ob und wie sie sich an das hier untersuchte Problem anpassen lassen. Es ist ferner zu untersuchen, ob und wie sich andere Verfahren zur Lösung linearer Sattelpunktprobleme anpassen lassen.

6.4 Ergebnisse der numerischen Experimente

6.4.1 Die Lösung in 1-D

Die Lösung im eindimensionalen Fall mit $\gamma = 0.001$, $\tau = 0.001$ und $h = \frac{1}{64}$

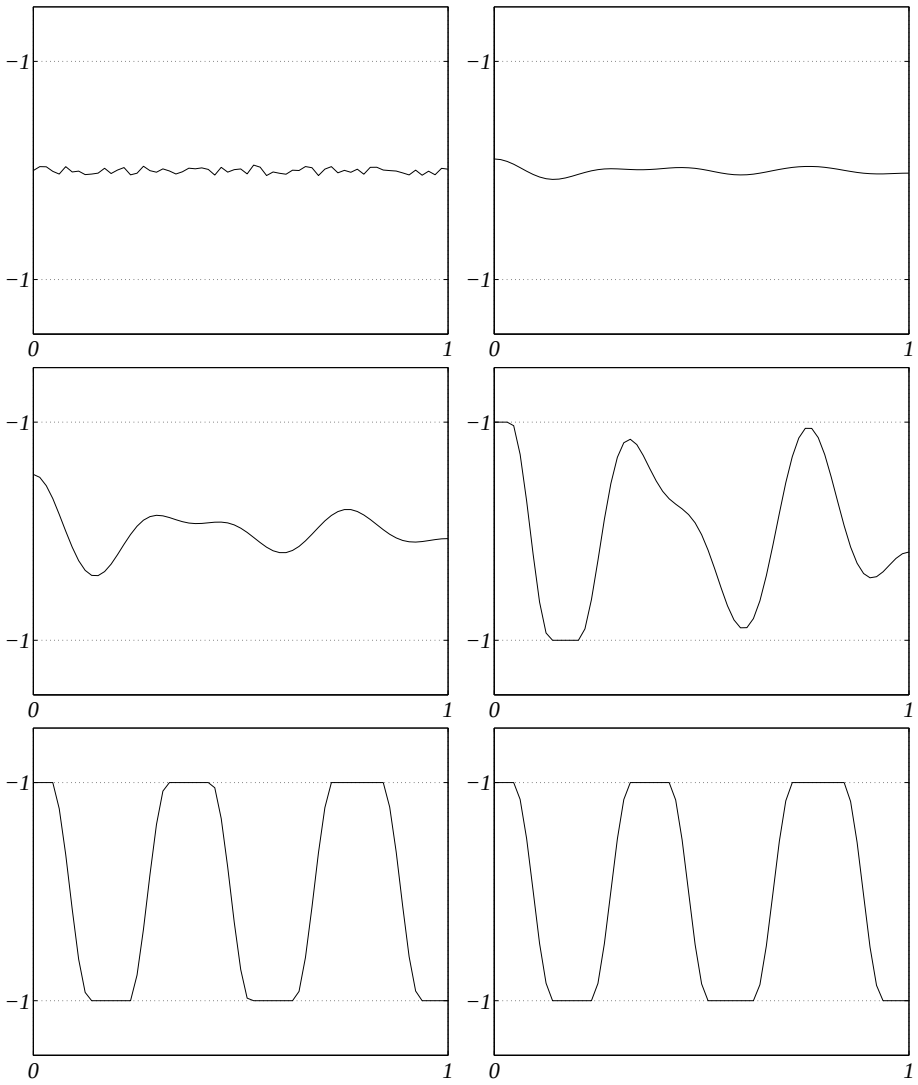

Abbildung 6.4.1: Zeitschritte t^n mit $n = 0, 10, 20, 30, 50, 500$

6.4.2 Die Lösung in 2-D

Die Lösung im zweidimensionalen Fall mit $\gamma = 0.001$, $\tau = 0.001$ und $h = \frac{1}{64}$

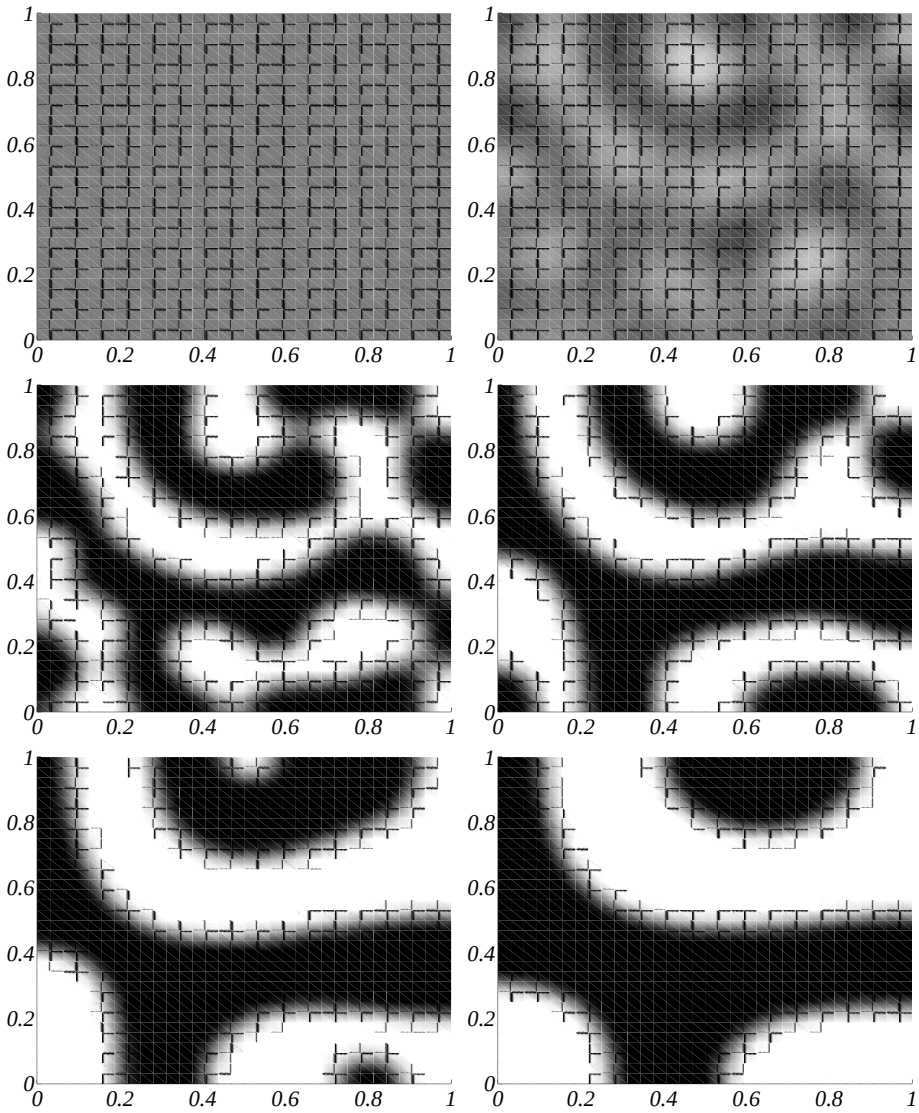

Abbildung 6.4.2: Zeitschritte t^n mit $n = 0, 20, 40, 80, 160, 500$

6.4.3 Verlauf der Iteration in 1-D

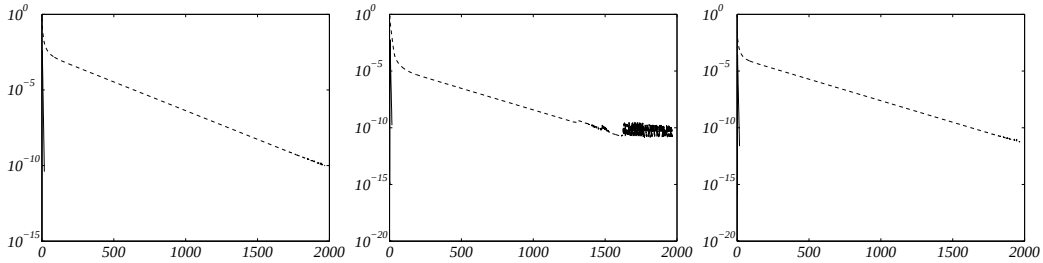

Abbildung 6.4.3: $\|u - u^n\|_h$, $\|u^{n+1} - u^n\|_h$, Defektnorm für $h = \frac{1}{16}$

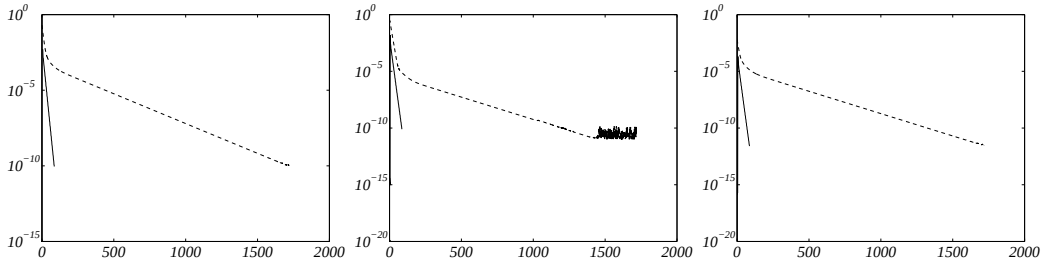

Abbildung 6.4.4: $\|u - u^n\|_h$, $\|u^{n+1} - u^n\|_h$, Defektnorm für $h = \frac{1}{32}$

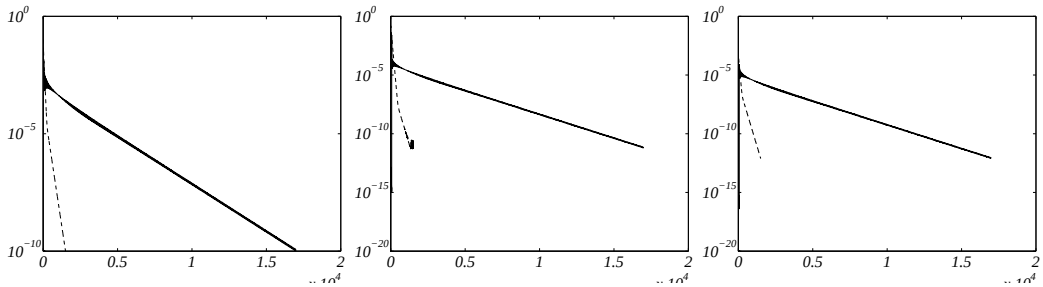

Abbildung 6.4.5: $\|u - u^n\|_h$, $\|u^{n+1} - u^n\|_h$, Defektnorm für $h = \frac{1}{128}$

6.4.4 Verlauf der Iteration in 2-D

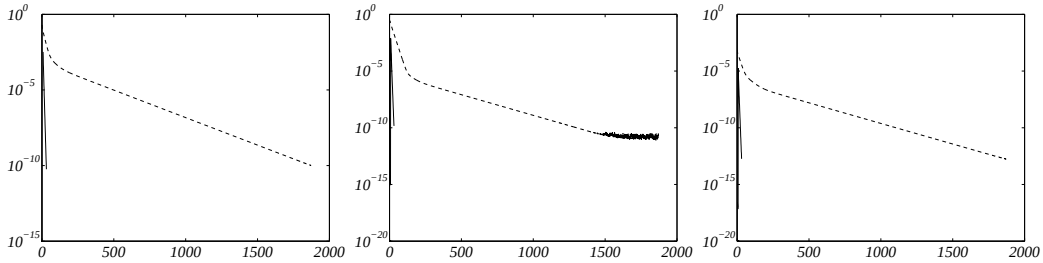

Abbildung 6.4.6: $\|u - u^n\|_h$, $\|u^{n+1} - u^n\|_h$, Defektnorm für $h = \frac{1}{16}$

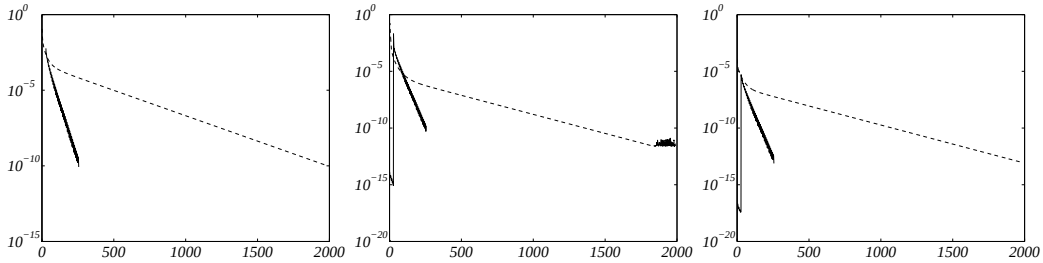

Abbildung 6.4.7: $\|u - u^n\|_h$, $\|u^{n+1} - u^n\|_h$, Defektnorm für $h = \frac{1}{32}$

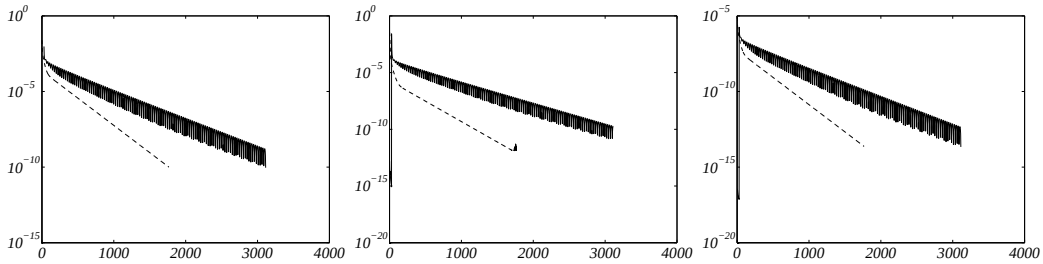

Abbildung 6.4.8: $\|u - u^n\|_h$, $\|u^{n+1} - u^n\|_h$, Defektnorm für $h = \frac{1}{64}$

6.4.5 Anzahl der Iterationsschritte

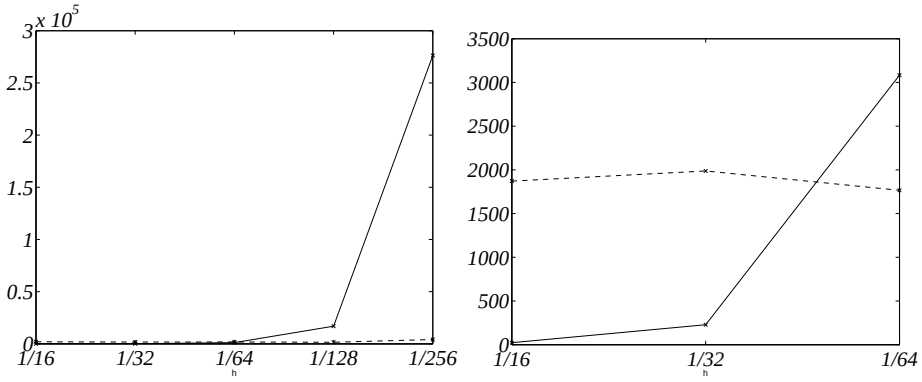

Abbildung 6.4.9: Iterationschritte, um $\|u - u^n\|_h < 10^{-10}$ zu erreichen,1-D und 2-D

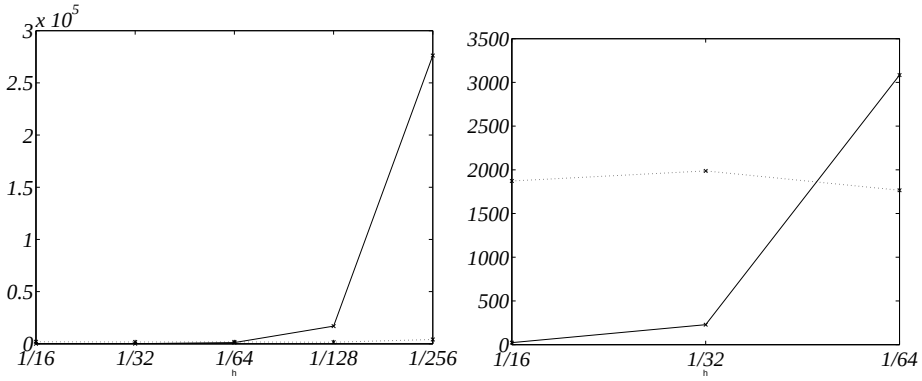

Abbildung 6.4.10: Gauß-Seidel und inexakter (gepunktet) Uzawa Algorithmus im Vergleich

A Notation

Es werden folgende Notationen verwendet:

Ω		beschränktes Gebiet in $\mathbb{R}^d$, $d \in \{1,2,3\}$, das der gleichmäßigen Kegelbedingung genügt
$\|f\|_V$		Norm auf dem Raum V
$(f,g)_V$		Skalarprodukt auf dem Raum V
$\langle f', g\rangle$	$= f'(g)$	duale Paarung von V' und V
T'		adjungierter Operator T
$x \cdot y$	$= x^T y = \sum_{i=1}^{d} x_i y_i$	euklidisches Skalarprodukt in $\mathbb{R}^d$
$\|x\|$	$= (x \cdot x)^{\frac{1}{2}}$	euklidische Norm in $\mathbb{R}^d$ bzw. Betrag in $\mathbb{R}$
(f,g)	$= \int_\Omega f(x)g(x)dx$	L^2-Skalarprodukt
$\|f\|_0$	$= (f,f)^{\frac{1}{2}}$	L^2-Norm
$(\nabla f, \nabla g)$	$= \int_\Omega \nabla f(x) \cdot \nabla g(x)dx$	„L^2-Skalarprodukt“ von ∇f und ∇g
$(f,g)_1$	$= (f,g) + (\nabla f, \nabla g)$	H^1-Skalarprodukt
$\|f\|_1$	$= (\nabla f, \nabla g)^{\frac{1}{2}}$	H^1-Halbnorm
$\|f\|_1$	$= \|f\|_0 + \|f\|_1$	H^1-Norm
$\|f\|_{-1}$	$= \|\mathcal{G}f\|_1$	zu $\|\cdot\|_{(H^1(\Omega))'}$ äquivalente Norm auf $\mathcal{F}$

Literatur

[Ada75] ADAMS, R. A.: *Sobolev Spaces*. Nummer 65 in *Pure and Applied Mathematics*. Academic Press, New York, 1975.

[BBG99] BARRETT, J.W., J.F. BLOWEY und H. GARCKE: *Finite Element Approximation of the Cahn-Hilliard Equation with Degenerate Mobility*. SIAM Jnl Numerical Analysis, 37(1), 1999.

[BE91] BLOWEY, J.F. und C.M. ELLIOT: *The Cahn-Hilliard Gradient Theory for Phase Separation with non-smooth free Energy Part I: Mathematical Analysis*. Euro. Jnl of Applied Mathematics, 2, 1991.

[BE92] BLOWEY, J.F. und C.M. ELLIOT: *The Cahn-Hilliard Gradient Theory for Phase Separation with non-smooth free Energy Part II: Numerical Analysis*. Euro. Jnl of Applied Mathematics, 3, 1992.

[BO91] BABUŠKA, I. und J. OSBORN: *Eigenvalue Problems*. In: CIARLET, P.G. und LIONS J.L. (Herausgeber): *Finite Elements (Part 1)*, Band 2 der Reihe *Handbook of Numerical Analysis*. North Holland, Amsterdam, 1991.

[Bra93] BRAMBLE, J. H.: *Multigrid Methods*. Nummer 294 in *Pitman Research Notes in Mathematics*. Longman Scientific & Technical, Essex, 1993.

[Bra97] BRAESS, D.: *Finite Elemente*. Springer, Berlin Heidelberg New York, 2. Auflage, 1997.

[BS96] BROKATE, M. und J. SPREKELS: *Hysteresis and Phase Transition*. Nummer 121 in *Applied mathematical sciences*. Springer, Berlin Heidelberg New York, 1996.

[CE92] COPETTI, M.I.M und C.M. ELLIOTT: *Numerical Analysis of the Cahn-Hilliard Equation with a Logarithmic free Energy*. Numerische Mathematik, 63, 1992.

[CH58] CAHN, J.W. und J.E. HILLIARD: *Free Energy of a non-uniform System I. Interfacial free Energy*. Jnl. of Chemical Physic, 28, 1958.

[CH03] CHENG, X.-I. und W. HAN: *Inexact Uzawa Algorithems for Variational Inequalities of the second Kind*. Computer Methods in Applied Mechanics and Enngeneering, 192, 2003.

[DB02] DEUFELHARD, P. und F. BORNEMANN: *Numerische Mathematik 2*. Walter de Gruyter, Berlin, 4. Auflage, 2002.

[DL92] DAUTRAY, R. und J.L. LIONS: *Mathematical Analysis and Numerical Methods for Science and Technology. 5: Evolution Problems*. Springer, Berlin Heidelberg New York, 1992.

[Ell89] ELLIOTT, C.M.: *The Cahn-Hilliard Model for the Kinetics of Phase Separation*. In: RODRIGUES, J.F. (Herausgeber): *Mathematical Models for Phase Change Problems*, Band 88 der Reihe *International Series of Numerical Mathematics*. Birkhäuser, Basel, 1989.

[ET76] EKELAND, I. und R. TEMAM: *Convex Analysis*. Nummer 1 in *Studies in Mathematics and its Applications*. North-Holland Publishing Company, Amsterdam New York Oxford, 1976.

[FG83] FORTIN, M. und R. GLOWINSKI: *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*. North-Holland, New York, 1983.

[GK02] GEIGER, C. und C. KANZOW: *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer, Berlin Heidelberg New York, 2002.

[Glo84] GLOWINSKI, R.: *Numerical Methods for Nonlinear Variational Problems*. Springer Series in Computational Physics. Springer, Berlin Heidelberg New York, 3. Auflage, 1984.

[GLT81] GLOWINSKI, R., J.L. LIONS und R. TRÉMOLIÈRES: *Numerical Analysis of Variational Inequalities*. Nummer 8 in *Studies in Mathematics and its Applications*. North-Holland Publishing Company, Amsterdam New York Oxford, 1981.

[Gri85] GRISVARD, P.: *Elliptic Problems in Nonsmooth Domains*. Nummer 24 in *Monographs and studies in mathematics*. Pitman, Boston London Melbourne, 1985.

[KA00] KNABNER, P. und L. ANGERMANN: *Numerik partieller Differentialgleichungen*. Springer, Berlin Heidelberg New York, 2000.

[Kor97] KORNHUBER, R.: *Adaptive Monoton Multigrid Methods for Nonlinear Variational Problems*. Teubner, Stuttgart, 1. Auflage, 1997.

[Lio69] LIONS, J.L.: *Quelques méthodes de résolution des problémes aux limites non linéaires*. Dunod, Paris, 1969.

[LM79] LIONS, P.L. und B. MERCIER: *Splitting Algorithms for the Sum of two Nonlinear Operators*. SIAM Jnl Numerical Analysis, 16, 1979.

[Mei99] MEISTER, A.: *Numerik linearer Gleichungssysteme*. Vieweg, Braunschweig Wiesbaden, 1999.

[Roc70] ROCKAFELLAR, R.T.: *Convex Analysis*. Nummer 28 in *Princeton Mathematical Series*. Princeton University Press, Princeton, 1970.

[Rud64] RUDIN, W.: *Principles of Mathematical Analysis*. McGraw-Hill, New York, 2. Auflage, 1964.

[Růž04] RŮŽIČKA, M.: *Nichtlineare Funktionalanalysis*. Springer, Berlin Heidelberg New York, 2004.

[RW98] ROCKAFELLAR, R.T. und R. J-B. WETS: *Variational Analysis*. Nummer 317 in *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin Heidelberg New York, 1998.

[Sen03] SENF, S. M.: *A Globally Convergent Gauß-Seidel Iteration for the Cahn-Hilliard Equation with Obstacle Potential.Diplomarbeit*. Freie Universität Berlin, 2003.

[Spe93] SPELLUCCI, P.: *Numerische Verfahren der nichtlinearen Optimierung*. Birkhäuser, Basel Berlin, 1993.

[Wer00] WERNER, D.: *Funktionalanalysis*. Springer, Berlin Heidelberg New York, 3. Auflage, 2000.

[Wlo82] WLOKA, J.: *Partielle Differentialgleichungen*. B.G.Teubner, Stuttgart, 1982.

[Zei85a] ZEIDLER, E.: *Nonlinear Functional Analysis and its Applications. IIb: Nonlinear Monotone Operators*. Springer, Berlin Heidelberg New York, 1985.

[Zei85b] ZEIDLER, E.: *Nonlinear Functional Analysis and its Applications. III: Variational Methods and Optimization*. Springer, Berlin Heidelberg New York, 1985.

[Zei90] ZEIDLER, E.: *Nonlinear Functional Analysis and its Applications. IIa: Linear Monotone Operators*. Springer, Berlin Heidelberg New York, 1990.