

**Numerische Lösung des
parabolischen Hindernisproblems mit einer
inneren Punkte Methode**

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Matthias Max Wick

aus Zuzwil (SG)

Basel 2000

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Frau Professorin Dr. C. Bandle und Herrn Professor Dr. R. Scholz

Basel, den 7. Dezember 1999

Prof. Dr. Andreas D. Zuberbühler
Dekan

Inhaltsverzeichnis

1	Einführung	1
2	Parabolische Variationsungleichung	4
2.1	Problemstellung	4
2.1.1	Bezeichnungen und Definitionen	4
2.1.2	Die abstrakte Variationsungleichung	9
2.2	Der Raum der finiten Elemente	19
2.2.1	Basisfunktionen	19
2.2.2	Interpolationsfehler	21
2.2.3	Eigenwertabschätzung des elliptischen Operators in S_h	24
2.3	Approximation der Variationsungleichung	27
2.3.1	Die semidiskrete Variationsungleichung	27
2.3.2	Eigenschaften der semidiskreten Lösung	28
2.3.3	Diskrete Variationsungleichung	37
2.3.4	Eigenschaften der diskreten Lösung	37
2.4	Fehlerabschätzung	40
2.4.1	Fehler zwischen den Lösungen der starken Variationsungleichung und der semidiskreten Variationsungleichung	40
2.4.2	Fehler zwischen den Lösungen der semidiskreten und der diskreten Variationsungleichung	44
2.4.3	Gesamtfehler	46
3	Lösungsmethode für stark monotone Variationsungleichungen	48
3.1	Numerische Lösungsmethode mit quadratischen Schnitten	49
3.2	Berechnungsbeispiele	64
4	Zusatzbetrachtungen und Ausblick	69
4.1	Kollokationsverfahren bei Variationsungleichungen	69
4.2	Semilineare Variationsungleichungen	71

Kapitel 1

Einführung

Thema

Variationsungleichungen wurden zuerst von Lions und Stampacchia [14] untersucht. Sie treten in verschiedenen Anwendungen der Mechanik, Thermodynamik und der Kontrolltheorie auf. Zum Beispiel kann die Wertbestimmung Amerikanischer Call-Optionen als Lösung einer parabolischen Variationsungleichung aufgefasst werden (vgl. Dewynne et al. [7]). Eine weitere Anwendung findet sich im Bereich der Thermodynamik (vgl. Duvaut und Lions [8]). Wird die Temperaturverteilung u in einem Gebiet Ω während dem Zeitintervall $(0, T)$ mit der Anfangsverteilung u_0 und dem Wärmequellterm f untersucht, so erfüllt $u(x, t)$ die Wärmeleitungsgleichung

$$\begin{aligned} \dot{u} - \Delta u &= f && \text{in } \Omega \times (0, T), \\ u(x, 0) &= u_0(x) && \text{in } \Omega. \end{aligned} \tag{1.1}$$

Soll die Temperatur an jeder Stelle $x \in \Omega$ möglichst in einem vorgegebenen Intervall $[\underline{\psi}(x), \overline{\psi}(x)]$ liegen, so muss die Temperatur beim Überschreiten der Intervallgrenzen durch Kühlung bzw. Heizung korrigiert werden. Dieser Effekt kann mit dem zusätzlichen Quellterm

$$\beta(u(x, t)) := \begin{cases} u(x, t) - \overline{\psi}(x) & \text{für } u(x, t) > \overline{\psi}(x), \\ u(x, t) - \underline{\psi}(x) & \text{für } u(x, t) < \underline{\psi}(x), \\ 0 & \text{sonst,} \end{cases}$$

in die Gleichung (1.1) eingebaut werden. Der Term β wird auch Strafterm genannt. Wird die Kapazität der Heizung und der Kühlung beliebig vergrößert, d.h. $\beta(u)$ durch einen Faktor beliebig verstärkt, so wird die Temperaturfunktion u im Limes für alle x und t im Intervall $[\underline{\psi}(x), \overline{\psi}(x)]$ liegen. Die Funktionen $\underline{\psi}$ bzw. $\overline{\psi}$ werden untere bzw. obere Hindernisfunktion genannt. Im Grenzfall erhält man ein parabolisches Hindernisproblem.

Die Lösung erfüllt die Hindernisbedingung $\underline{\psi}(x) \leq u(x, t) \leq \overline{\psi}(x)$ und zudem gilt

$$\begin{aligned} \dot{u} - \Delta u - f &\leq 0 && \text{für } u(x, t) = \overline{\psi}(x), \\ \dot{u} - \Delta u - f &= 0 && \text{für } \underline{\psi}(x) < u(x, t) < \overline{\psi}(x), \\ \dot{u} - \Delta u - f &\geq 0 && \text{für } u(x, t) = \underline{\psi}(x). \end{aligned} \tag{1.2}$$

Beschränkt man sich im betrachteten Funktionenraum auf die konvexe, abgeschlossene Teilmenge K aller Funktionen v mit $\underline{\psi}(x) \leq u(x, t) \leq \overline{\psi}(x)$, so kann das Problem (1.2) als Variationsungleichungsproblem formuliert werden. Für alle $v \in K$ ist $v(x, t) - \underline{\psi}(x) \geq 0$ und $v(x, t) - \overline{\psi}(x) \leq 0$ fast überall in Ω . Ist D der Definitionsbereich des Operators $\frac{d}{dt}$, so lässt sich das Problem (1.2) in schwacher Form wie folgt als Variationsungleichungsproblem formulieren:

Finde $u \in K \cap D$ mit

$$\int_0^T (\dot{u} - f, v - u) dt + \int_0^T (\nabla u, \nabla(v - u)) dt \geq 0 \quad \forall v \in K. \quad (1.3)$$

Das Hauptziel dieser Arbeit ist es, Variationsungleichungen, die aus Hindernisproblemen entstehen, numerisch zu lösen. Dieses Problem wurde von Scholz im elliptischen [18] und im parabolischen Fall [19] untersucht. Er wandelt (1.3) mit Hilfe der Penaltymethode in eine Gleichung um und löst die erhaltene Gleichung numerisch.

In dieser Arbeit wird auf die Penaltymethode verzichtet. Zur Diskretisierung von (1.3) wird die Variationsungleichung in einen endlichdimensionalen Funktionenraum, den Raum der finiten Elemente, übertragen und der zeitliche Ableitungsoperator $\frac{d}{dt}$ durch einen Differenzenquotienten approximiert. Dadurch entsteht für jeden Zeitschritt eine diskrete Variationsungleichung. Sei $K_n \subset \mathbb{R}^n$ eine konvexe, beschränkte und abgeschlossene Menge und $F : K_n \rightarrow \mathbb{R}^n$ eine stark monotone Abbildung. Die diskrete Variationsungleichung in \mathbb{R}^n hat die Form

$$y \in K_n : \quad (F(y), z - y) \geq 0 \quad \forall z \in K_n. \quad (1.4)$$

Diese Variationsungleichung wird direkt, also ohne Umwandlung in eine Gleichung, mit Hilfe einer inneren Punkte Methode numerisch gelöst. Es wird eine Punktfolge in K_n konstruiert, die gegen die Lösung von (1.4) konvergiert. Der Vorteil dieser Methode gegenüber der Penaltymethode besteht darin, dass jede Näherung der Lösung von (1.4) im Innern von K_n liegt und somit die zugehörige Funktion die Hindernisbedingung erfüllt. Bei der Penaltymethode ist dies nicht der Fall. Dort wird das Nichterfülltsein der Hindernisbedingung bestraft. Da die Bestrafung nur für Funktionen ausserhalb von K wirkt, liegen alle berechneten Näherungen der Lösung von (1.3) ausserhalb von K .

Die untersuchte Methode zur Lösung von (1.4) basiert darauf, dass zu jedem inneren Punkt von K_n eine Kugel in \mathbb{R}^n bestimmt werden kann, auf deren Oberfläche dieser innere Punkt liegt und welche die Lösung von (1.4) enthält. Ist ein innerer Punkt von K_n gefunden, so kann der Teil von K_n , in dem die Lösung nicht liegt, weggeschnitten werden. In der so erhaltenen Teilmenge von K_n wird erneut ein innerer Punkt berechnet. Für die Bestimmung der inneren Punkte wird die von Sonnevend [21] eingeführte Methode der analytischen Zentren verwendet.

Verfahren zur Lösung von Variationsungleichungen in \mathbb{R}^n mit linearen Schnitten wurden von Goffin, Marcotte und Zhu [11] und von Magnanti und Perakis [16] untersucht. Neu an der Methode mit quadratischen Schnitten (ACQCM) von Lüthi und Büeler [15] ist, dass die starke Monotonie der Abbildung F für die Konstruktion der Schnitte ausgenutzt wird. Wir führen in der ACQCM n -fach gewichtete Schnitte ein. Dadurch wird eine bessere Konvergenz im Fall der Hindernisprobleme erreicht.

Aufbau des Berichtes

Im Kapitel 2 werden zunächst die Grundbegriffe und die genaue Problemstellung eingeführt. Anschliessend wird die Variationsungleichung in zwei Stufen diskretisiert. Durch die Einführung

des Raumes der finiten Elemente wird die Variationsungleichung in der Ortsvariablen x diskretisiert. Es entsteht die semidiskrete Variationsungleichung. Neben einigen Eigenschaften der semidiskreten Lösung wird die Existenz und Eindeutigkeit einer Lösung konstruktiv bewiesen, indem zu jedem Punkt $y(t_0) \in K_n$ eine eindeutige Richtung $\dot{y}(t_0)$ konstruiert wird. Durch Diskretisierung der Zeitachse erhält man mit dem rückwärts Eulerverfahren die diskrete Variationsungleichung (1.4). Das Hauptresultat dieses Kapitels ist die Fehlerabschätzung zwischen der Lösung der Variationsungleichung (1.3) und der Lösung der diskreten Variationsungleichung (1.4).

Das Kapitel 3 befasst sich mit der numerischen Methode zur Lösung der diskreten Variationsungleichung. Zuerst wird die geometrische Idee der Methode der quadratischen Schnitte mit analytischen Zentren (ACQCM) erläutert und anschliessend die Konvergenz bewiesen. Zur Illustration werden einige Beispiele gezeigt.

Im Kapitel 4 weisen wir auf zwei weitere Aspekte hin. Zum einen wird gezeigt, dass bei Variationsungleichungen die rückwärts Eulermethode nicht ohne weiteres durch die allgemeinere Kollokationsmethode ersetzt werden kann. Zum andern ist die vorgeschlagene Methode so flexibel, dass sie auch zur Lösung von semilinearen parabolischen Variationsungleichungen verwendet werden kann. Ist die Semilinearität monoton, so kann sie sogar in die Konstruktion der quadratischen Schnitte integriert werden. Dadurch entsteht ein vollständig implizites Verfahren. Bei einer nicht monotonen Semilinearität muss das Verfahren in diesem Term explizit sein, d.h. man setzt den Wert des vorangehenden Zeitschrittes ein und betrachtet den Term als konstant. Es wird der nicht monotone und der monotone Fall einer Semilinearität anhand eines Beispiels aufgezeigt.

Dank

Die vorliegende Arbeit entstand am Mathematischen Institut der Universität Basel unter der Anleitung von Frau Prof. Dr. C. Bandle. An dieser Stelle möchte ich ihr für die intensive Betreuung herzlich danken. Sie hat sich immer die Zeit genommen, meine Arbeit durch Ratschläge und wissenschaftliche Gespräche zu unterstützen.

Für die finanzielle Unterstützung während dieser Arbeit bin ich dem Schweizerischen Nationalfond zu Dank verpflichtet.

Ich danke meinen Eltern, die mir diese Ausbildung ermöglicht haben und mich jederzeit unterstützten. Ein besonderen Dank gilt meiner Frau Claudia für ihre Geduld während der letzten Jahre. Sie hat dadurch einen wichtigen Teil zum Gelingen dieser Arbeit beigetragen.

Kapitel 2

Parabolische Variationsungleichung

2.1 Problemstellung

2.1.1 Bezeichnungen und Definitionen

Für die Beschreibung des Problems benötigen wir die folgenden Bezeichnungen:

$\Omega \subset \mathbb{R}^N$ beschränktes, offenes Gebiet mit genügend glattem Rand, z.B. $\partial\Omega$ stückweise C^2 .

$Q = \Omega \times (0, T]$ mit $T > 0$ fest.

\mathcal{U} reeller, separabler und reflexiver Banachraum mit Norm $\|\cdot\|$.

\mathcal{U}' Dualraum zu \mathcal{U} , $\langle \cdot, \cdot \rangle$ Paarungsprodukt zwischen \mathcal{U}' und \mathcal{U} .

\mathcal{H} reeller, separabler Hilbertraum mit Norm $|\cdot|$ und Skalarprodukt (\cdot, \cdot) .

\mathcal{U} liege bezüglich der Norm von \mathcal{H} dicht in \mathcal{H} . Es gilt für alle $v \in \mathcal{U}$ $|v| \leq \text{const}\|v\|$. Damit erhalten wir das Evolutionstripel $\mathcal{U} \subseteq \mathcal{H} \subseteq \mathcal{U}'$.

In dieser Arbeit werden die Funktionenräume

$$\mathcal{U} = H_0^1(\Omega) = W_0^{1,2}(\Omega) \quad \text{und} \quad \mathcal{H} = L^2(\Omega)$$

verwendet. Wenn nichts anderes vermerkt ist, gilt

$$|v|^2 = \int_{\Omega} |v(x)|^2 dx$$
$$\|v\|^2 = |v|^2 + \sum_{i=1}^N \left| \frac{\partial}{\partial x_i} v(x) \right|^2.$$

Für alle $w \in \mathcal{H}$ und $v \in \mathcal{U}$ kann das Paarungsprodukt geschrieben werden als

$$\langle w, v \rangle = \int_{\Omega} w(x)v(x) dx = (w, v). \quad (2.1)$$

Sei $u : (0, T) \rightarrow \mathcal{U}$ die Funktion definiert durch $[u(t)](x) := u(x, t)$. Den Raum all dieser Funktionen mit

$$\left(\int_0^T \|u(t)\|^2 dt \right)^{\frac{1}{2}} < \infty$$

bezeichnen wir mit $L^2(0, T; \mathcal{U})$. Analog ist $L^2(0, T; \mathcal{H})$ definiert. $\|\cdot\|_T$ bzw. $|\cdot|_T$ bezeichnen die Normen dieser Räume. Funktionen aus diesen Räumen werden auch Bochner integrierbare Funktionen genannt. Für $u \in L^2(0, T; \mathcal{U})$ und $\nu \in \mathcal{U}'$ gilt (vgl. [1] S. 185)

$$\left\| \int_0^T u(t) dt \right\| \leq \int_0^T \|u(t)\| dt \quad \text{und} \quad \left\langle \nu, \int_0^T u(t) dt \right\rangle = \int_0^T \langle \nu, u(t) \rangle dt. \quad (2.2)$$

Im Raum $L^2(0, T; \mathcal{U})$ ist die zeitliche Ableitung wie folgt definiert.

Definition 1 $u \in L^2(0, T; \mathcal{U})$ besitzt auf $(0, T)$ eine verallgemeinerte Ableitung genau dann, wenn ein $w \in L^2(0, T; \mathcal{U}')$ existiert mit

$$\int_0^T \dot{\varphi}(t) u(t) dt = - \int_0^T \varphi(t) w(t) dt, \quad \forall \varphi \in C_0^\infty(0, T).$$

Wir setzen dann $\dot{u} := w$.

Eigenschaften der Ableitung

Zunächst wird der Glättungsoperator J_ϵ eingeführt. Es sei $\eta \in C^\infty(\mathbb{R})$ mit

$$\eta(s) = \begin{cases} C e^{\frac{1}{|s|^2-1}} & \text{für } |s| < 1, \\ 0 & \text{für } |s| \geq 1. \end{cases}$$

C sei so gewählt, dass $\int_{\mathbb{R}} \eta(s) ds = 1$. Für jedes $\epsilon > 0$ sei $\eta_\epsilon(s) := \frac{1}{\epsilon} \eta\left(\frac{s}{\epsilon}\right)$. Der Glättungsoperator J_ϵ ist durch

$$J_\epsilon(u(t)) = \eta_\epsilon(t) * u(t) = \int_0^T \eta_\epsilon(t-s) u(s) ds =: u^\epsilon(t)$$

definiert.

Lemma 2.1 Existiert die verallgemeinerte Ableitung für ein $u \in L^2(0, T; \mathcal{U})$, so gilt für $t \in (\epsilon, T - \epsilon)$

$$\frac{d}{dt} J_\epsilon(u(t)) = J_\epsilon(\dot{u}(t)).$$

Beweis: Es gilt

$$\frac{d}{dt} J_\epsilon(u(t)) = \int_0^T \frac{d}{dt} \eta_\epsilon(t-s) u(s) ds = - \int_0^T \frac{d\eta_\epsilon}{ds}(t-s) u(s) ds.$$

Für ein festes $t \in (\epsilon, T - \epsilon)$ gehört $\eta_\epsilon(t-s)$ zu $C_0^\infty(0, T)$. Somit erhält man

$$\frac{d}{dt} J_\epsilon(u(t)) = \int_0^T \eta_\epsilon(t-s) \dot{u}(s) ds = J_\epsilon(\dot{u}(t)).$$

□

Lemma 2.2 Für $u \in L^2(0, T; \mathcal{U})$ mit $\dot{u} \in L^2(0, T; \mathcal{U}')$ existiert eine Folge $u^\epsilon \in C^\infty(0, T; \mathcal{U})$ mit

$$\begin{aligned} u^\epsilon &\rightarrow u && \text{in } L^2(0, T; \mathcal{U}), \\ \dot{u}^\epsilon &\rightarrow \dot{u} && \text{in } L^2(0, T; \mathcal{U}'). \end{aligned}$$

Beweis:

1. Sei $\epsilon > 0$. $u(t)$ werde ausserhalb von $[0, T]$ durch 0 fortgesetzt. Aus der Eigenschaft des Glättungsoperators folgt, dass $u^\epsilon \in C^\infty(0, T; \mathcal{U})$ (vgl. [1] S. 30).

Die Norm der geglätteten Funktion lässt sich wie folgt abschätzen. Es gilt mit (2.2) und der Schwarzschen Ungleichung

$$\begin{aligned} \|u^\epsilon(t)\| &= \left\| \int_0^T \eta_\epsilon(t-s)u(s)ds \right\| \\ &\leq \int_0^T \eta_\epsilon(t-s)\|u(s)\|ds \\ &\leq \underbrace{\left(\int_0^T \eta_\epsilon(t-s)ds \right)^{\frac{1}{2}}}_{\leq 1} \left(\int_0^T \eta_\epsilon(t-s)\|u(s)\|^2 ds \right)^{\frac{1}{2}}. \end{aligned} \quad (2.3)$$

Zusammen mit dem Satz von Fubini folgt

$$\begin{aligned} \int_0^T \|u^\epsilon(t)\|^2 dt &\stackrel{(2.3)}{\leq} \int_0^T \int_0^T \eta_\epsilon(t-s)\|u(s)\|^2 ds dt \\ &\stackrel{\text{Fubini}}{=} \int_0^T \|u(s)\|^2 \underbrace{\int_0^T \eta_\epsilon(t-s)dt}_{\leq 1} ds \leq \|u\|_T^2. \end{aligned} \quad (2.4)$$

Ist $\tilde{u} \in C_0(0, T; \mathcal{U})$, so gilt für $t \in [0, T]$

$$\begin{aligned} \|\tilde{u}^\epsilon(t) - \tilde{u}(t)\| &\leq \int_0^T \eta_\epsilon(t-s)\|\tilde{u}(s) - \tilde{u}(t)\|ds \\ &\leq \sup_{|t-s|<\epsilon} \|\tilde{u}(s) - \tilde{u}(t)\|. \end{aligned}$$

Da $\|\tilde{u}(\cdot)\|$ gleichmässig stetig auf $[0, T]$ ist, konvergiert \tilde{u}^ϵ auf $[0, T]$ gleichmässig gegen \tilde{u} in \mathcal{U} .

Die Dichtheit von $C_0(0, T; \mathcal{U})$ in $L^2(0, T; \mathcal{U})$ lässt sich wie im Fall $L^p(\Omega)$ zeigen. Nach der Definition der Bochner integrierbaren Funktionen ([1], S. 179) existiert zu $u \in L^2(0, T; \mathcal{U})$ eine Folge von „simplen Funktionen“ s_i , die in $L^2(0, T; \mathcal{U})$ gegen u konvergiert. „Simple Funktionen“ sind die Verallgemeinerung der Treppenfunktionen in $L^2(0, T)$. Sie nehmen in jedem Teilintervall von $[0, T]$ einen Wert aus \mathcal{U} an.

Für ein $\delta > 0$ sei $s \in \{s_i\}$ mit

$$\|s - u\|_T < \frac{\delta}{2}. \quad (2.5)$$

Wird $s(t)$ ausserhalb von $[0, T]$ durch 0 fortgesetzt und ist $T < \infty$, so existiert nach dem Satz von Lusin (vgl. [1] S. 14) ein $g \in C_0(0, T; \mathcal{U})$ mit

$$\|g(t)\| \leq \|s\|_{L^\infty(0, T; \mathcal{U})} \quad \text{für alle } t \in [0, T] \quad (2.6)$$

und für ein $\tilde{\delta} > 0$

$$\mu(\{t \in [0, T] : g(t) \neq s(t) \text{ in } \mathcal{U}\}) < \tilde{\delta}. \quad (2.7)$$

Wird $\tilde{\delta} = \frac{\delta}{4} \|s\|_{L^\infty(0, T; \mathcal{U})}^{-1}$ gewählt, so folgt mit der Schwarzschen Ungleichung, (2.7) und (2.6), dass

$$\|s - g\|_T < \frac{\delta}{2}.$$

Zusammen mit (2.5) erhält man die Dichtheit von $C_0(0, T; \mathcal{U})$ in $L^2(0, T; \mathcal{U})$.

Sei $\tilde{u} \in C_0(0, T; \mathcal{U})$ mit $\|\tilde{u} - u\|_T < \delta$. Es folgt mit der Dreiecksungleichung und (2.4)

$$\begin{aligned} \|u - u^\epsilon\|_T &\leq \|u - \tilde{u}\|_T + \|\tilde{u} - \tilde{u}^\epsilon\|_T + \|\tilde{u}^\epsilon - u^\epsilon\|_T \\ &\leq 2 \underbrace{\|u - \tilde{u}\|_T}_{< \delta} + \|\tilde{u} - \tilde{u}^\epsilon\|_T. \end{aligned}$$

Der zweite Summand kann wegen der gleichmässigen Konvergenz von \tilde{u}^ϵ beliebig klein gemacht werden. Damit folgt die erste Behauptung.

2. Die duale Norm von \dot{u}^ϵ wird wie im ersten Teil abgeschätzt. Für ein beliebiges $v \in \mathcal{U}$ mit $\|v\| = 1$ folgt mit (2.2) und der Schwarzschen Ungleichung

$$\sup_{v \in \mathcal{U}, \|v\|=1} |\langle \dot{u}^\epsilon(t), v \rangle| \leq \sup_v \left(\int_0^T \eta_\epsilon(t-s) |\langle \dot{u}(s), v \rangle|^2 ds \right)^{\frac{1}{2}} \leq \left(\int_0^T \eta_\epsilon(t-s) \|\dot{u}(s)\|_{\mathcal{U}'}^2 ds \right)^{\frac{1}{2}}.$$

Für die L^2 -Norm bezüglich t erhält man mit dem Satz von Fubini

$$\|\dot{u}^\epsilon\|_{L^2(0, T; \mathcal{U}')} \leq \|\dot{u}\|_{L^2(0, T; \mathcal{U}')} \quad (2.8)$$

Sei $\dot{\tilde{u}} \in C_0(0, T; \mathcal{U}')$. Auch hier folgt die gleichmässige Konvergenz von $\dot{\tilde{u}}^\epsilon$ gegen $\dot{\tilde{u}}$. Es gilt für $t \in [0, T]$ und $v \in \mathcal{U}$ mit $\|v\| = 1$

$$|\langle \dot{\tilde{u}}^\epsilon(t) - \dot{\tilde{u}}(t), v \rangle| \leq \int_0^T \eta_\epsilon(t-s) |\langle \dot{\tilde{u}}(s) - \dot{\tilde{u}}(t), v \rangle| ds \leq \sup_{|s-t| < \epsilon} \|\dot{\tilde{u}}(s) - \dot{\tilde{u}}(t)\|_{\mathcal{U}'}$$

Wie im ersten Teil folgt die Dichtheit von $C_0(0, T; \mathcal{U}')$ in $L^2(0, T; \mathcal{U}')$. Sei $\dot{\tilde{u}} \in C_0(0, T; \mathcal{U}')$ mit $\|\dot{\tilde{u}} - \dot{u}\|_{L^2(0, T; \mathcal{U}')} < \delta$. Wird die Dreiecksungleichung auf $\|\dot{u} - \dot{\tilde{u}} + \dot{\tilde{u}} - \dot{\tilde{u}}^\epsilon + \dot{\tilde{u}}^\epsilon - \dot{u}^\epsilon\|_{L^2(0, T; \mathcal{U}')}$ angewendet, so erhält man zusammen mit (2.8) und der gleichmässigen Konvergenz von $\dot{\tilde{u}}^\epsilon \rightarrow \dot{\tilde{u}}$ die 2. Behauptung.

□

Lemma 2.3 *Ist $u \in L^2(0, T; \mathcal{U})$ mit $\dot{u} \in L^2(0, T; \mathcal{U}')$, so gilt*

1. $u(t)$ ist stetig in t , d.h. $u \in C(0, T; \mathcal{H})$. (Damit ist gemeint, dass in der Äquivalenzklasse $u \in L^2(0, T; \mathcal{H})$ ein stetiger Representant existiert).
2. die Abbildung $t \rightarrow |u(t)|^2$ ist absolut stetig, mit

$$\frac{d}{dt}|u(t)|^2 = 2\langle \dot{u}(t), u(t) \rangle \quad \text{f.f.a. } t \in [0, T].$$

Beweis:

1. Sei $u^\epsilon \in C^\infty(0, T; \mathcal{U})$ die Folge glatter Funktionen, die durch Glättung bezüglich t entsteht. Für $\epsilon, \delta > 0$ gilt

$$\frac{d}{dt}|u^\epsilon(t) - u^\delta(t)|^2 = 2 \left(\frac{d}{dt}u^\epsilon(t) - \frac{d}{dt}u^\delta(t), u^\epsilon(t) - u^\delta(t) \right).$$

Das Lemma 2.1 besagt, dass $\frac{d}{dt}u^\epsilon = \dot{u}^\epsilon$. Damit und wegen (2.1) folgt für jedes $s, t \in [0, T]$

$$|u^\epsilon(t) - u^\delta(t)|^2 = |u^\epsilon(s) - u^\delta(s)|^2 + 2 \int_s^t \langle \dot{u}^\epsilon(r) - \dot{u}^\delta(r), u^\epsilon(r) - u^\delta(r) \rangle dr. \quad (2.9)$$

Da $u^\epsilon(s)$ in L^2 gegen $u(s)$ konvergiert, existiert mindestens ein $s \in (0, T)$, so dass

$$u^\epsilon(s) \rightarrow u(s) \quad \text{in } \mathcal{H}.$$

Aus (2.9) und dem Lemma 2.2 folgt

$$\begin{aligned} \limsup_{\epsilon, \delta \rightarrow 0} \sup_{t \in [0, T]} |u^\epsilon(t) - u^\delta(t)|^2 &\leq \lim_{\epsilon, \delta \rightarrow \infty} \int_0^T \|\dot{u}^\epsilon(r) - \dot{u}^\delta(r)\|_{\mathcal{U}'}^2 + \|u^\epsilon(r) - u^\delta(r)\|^2 dr \\ &= 0. \end{aligned}$$

Somit konvergiert u^ϵ in $C(0, T; \mathcal{H})$ gleichmässig gegen eine Limesfunktion $v \in C(0, T; \mathcal{H})$. Da nach dem Lemma 2.2 $u^\epsilon \rightarrow u$ in $L^2(0, T; \mathcal{U})$, folgt $u(t) = v(t)$ für fast alle $t \in [0, T]$.

2. Für die geglättete Funktion $u^\epsilon(t)$ gilt

$$|u^\epsilon(t)|^2 = |u^\epsilon(s)|^2 + 2 \int_s^t \langle \dot{u}^\epsilon(r), u^\epsilon(r) \rangle dr.$$

Strebt $\epsilon \rightarrow 0$ und wird u mit dem v aus dem ersten Teil des Beweises identifiziert, so folgt

$$|u(t)|^2 = |u(s)|^2 + 2 \int_s^t \langle \dot{u}(r), u(r) \rangle dr \quad \text{für alle } s, t \in [0, T],$$

was die 2. Behauptung beweist.

□

2.1.2 Die abstrakte Variationsungleichung

Es sei (a^{ij}) eine positiv definite, symmetrische $N \times N$ -Matrix. Die Bilinearform $a : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$ habe die Form

$$a(u, v) := \int_{\Omega} \sum_{i,j=1}^N a^{ij} \frac{\partial}{\partial x_i} u(x) \frac{\partial}{\partial x_j} v(x) dx.$$

$a(\cdot, \cdot)$ sei in beiden Argumenten stetig, d.h. $\exists c > 0$ so, dass

$$|a(u, v)| \leq c \|u\| \cdot \|v\| \quad \forall u, v \in \mathcal{U}. \quad (2.10)$$

Zudem soll für ein $\alpha > 0$

$$a(u, u) \geq \alpha \|u\|^2 \quad \forall u \in \mathcal{U} \quad (2.11)$$

gelten. Eine Bilinearform mit letzterer Eigenschaft heisst koerzitiv. Für festes $u \in \mathcal{U}$ ist $a(u, \cdot)$ ein beschränktes Funktional auf \mathcal{U} . Sei $A : \mathcal{U} \rightarrow \mathcal{U}'$ die Abbildung, die jedem $u \in \mathcal{U}$ das entsprechende Funktional in \mathcal{U}' zuordnet. Es gilt für ein $u \in \mathcal{U}$

$$a(u, v) = \langle Au, v \rangle \quad \forall v \in \mathcal{U}.$$

Für $u \in H^2$ ist A ein Differentialoperator zweiter Ordnung.

K sei eine abgeschlossene, konvexe Teilmenge von $L^2(0, T; \mathcal{U})$. Ist $u_0(x) \in K$ die Anfangsbedingung für die Lösung, so wird die Definitionsmenge des zeitlichen Ableitungsoperators definiert durch

$$D := \{v \in L^2(0, T; \mathcal{U}) : \dot{v} \in L^2(0, T; \mathcal{U}'), v(x, 0) = u_0(x)\}.$$

Wegen des Lemmas 2.3 macht die Bedingung $v(x, 0) = u_0(x)$ Sinn. Für $f \in \mathcal{U}'$ können wir die starke parabolische Variationsungleichung im Zeitintervall $(0, T)$ wie folgt formulieren.

Starke Variationsungleichung: Finde ein $u \in K \cap D$ so, dass

$$\int_0^T \langle \dot{u}, v - u \rangle dt + \int_0^T a(u, v - u) dt \geq \int_0^T \langle f, v - u \rangle dt \quad \forall v \in K. \quad (V)$$

Bemerkung: Wie in der Einleitung beschrieben, kann ein parabolisches Hindernisproblem in eine parabolische Variationsungleichung umgeschrieben werden. In diesem Fall ist die Zulässigkeitsmenge K durch die Hindernisfunktionen bestimmt. $\underline{\psi}, \bar{\psi} \in H^2$ seien unabhängig von der Zeit und es gelte $\underline{\psi} < \bar{\psi}$, $\underline{\psi} \leq 0$ und $\bar{\psi} \geq 0$ auf $\partial\Omega$. Nun ist

$$K := \{v \in L^2(0, T; \mathcal{U}) : \underline{\psi}(x) \leq v(x, t) \leq \bar{\psi}(x) \text{ f.ü. in } Q\}.$$

Wegen der Randbedingungen an die Hindernisfunktionen ist $K \neq \emptyset$.

Eigenschaften der Lösung

Bensoussan und Lions haben in [3] bewiesen, dass das Problem (V) eine eindeutige Lösung besitzt. Sie zeigen zudem einige Regularitätseigenschaften der Lösung. Im folgenden Satz sind ihre Ergebnisse zusammengefasst.

Satz 2.4 (Bensoussan und Lions [3]) *Die parabolische Variationsungleichung (V) besitzt eine eindeutige Lösung u mit*

$$\begin{aligned} u &\in L^\infty(0, T; H_0^1(\Omega)) \cap L^2(0, T; H^2(\Omega)), \\ \dot{u} &\in L^\infty(0, T; L^2(\Omega)) \cap L^2(0, T; H_0^1(\Omega)). \end{aligned}$$

\dot{u} liegt nach diesem Satz auch in $L^2(0, T; \mathcal{H})$. Diese Eigenschaft wird im Folgenden bereits in die Voraussetzungen eingebaut. Die Menge D wird im Rest der Arbeit ersetzt durch eine kleinere Menge

$$\mathcal{D} := \{v \in L^2(0, T; \mathcal{U}) : \dot{v} \in L^2(0, T; \mathcal{H}), v(x, 0) = u_0(x)\}.$$

Für alle $w \in \mathcal{D}$ und f.a. $t \in [0, T]$ kann nach (2.1) das Paarungsprodukt durch das Skalarprodukt in \mathcal{H} ersetzt werden,

$$\langle \dot{w}(t), v(t) \rangle = (\dot{w}(t), v(t)) \quad \forall v \in L^2(0, T; \mathcal{U}).$$

Zudem wird im Rest der Arbeit $f \in \mathcal{H}$ und $u_0 \in H^2(\Omega) \cap K$ vorausgesetzt. Brézis ([5], S. 100) hat gezeigt, dass unter diesen Voraussetzungen für die Lösung u von (V) zusätzlich gilt

$$u \in L^\infty(0, T; H^2(\Omega)), \quad u(t) \in H^2(\Omega) \quad \forall t > 0. \quad (2.12)$$

Damit gilt $Au(t) \in \mathcal{H}$ und mit 2.1 folgt $a(u, w) = (Au, w)$ für alle $w \in \mathcal{H}$. Es soll nun Au und \dot{u} in verschiedenen Normen durch die Daten des Problems abgeschätzt werden. Für den Beweis benötigen wir folgendes Hilfsproblem. Die Bedingung $u \in K$ wird abgeschwächt. Die Nichterfüllung dieser Bedingung wird mit dem Penaltyterm $\beta(v) = (v - \underline{\psi})^- + (v - \bar{\psi})^+$, dabei ist $v^+ := \max(v, 0)$ und $v^- := \min(v, 0)$, bestraft. Indem man die Bestrafung verstärkt, rückt die Lösung des Penaltyproblems immer näher an K heran. Dadurch wird die Lösung von (V) beliebig gut approximiert.

Finde $u_\epsilon \in L^2(0, T; H_0^1(\Omega)) \cap \mathcal{D}$ mit

$$\int_0^T (\dot{u}_\epsilon, v) dt + \int_0^T a(u_\epsilon, v) dt + \underbrace{\frac{1}{\epsilon} \int_0^T (\beta(u_\epsilon), v) dt}_{\text{Bestrafungsterm}} = \int_0^T (f, v) dt \quad \forall v \in L^2(0, T; H_0^1(\Omega)). \quad (2.13)$$

Dieses Problem besitzt für jedes $\epsilon > 0$ nach der klassischen Theorie eine eindeutige Lösung u_ϵ . u_ϵ und \dot{u}_ϵ konvergieren schwach gegen u bzw. \dot{u} in $L^2(0, T; H_0^1(\Omega))$ (vgl. [3]).

Lemma 2.5 *Sei $A\underline{\psi}, A\bar{\psi} \in \mathcal{H}$. Ist u die Lösung der Variationsungleichung (V), so gibt es ein c (vgl. 2.10), das von t und u unabhängig ist, mit*

$$|Au|_T \leq 2|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T + \sqrt{\frac{c}{2}} \|u_0\|.$$

Beweis: Wir betrachten das Penaltyproblem 2.13. Der Beweis ist in zwei Teile gegliedert. Zunächst wird eine Abschätzung für den Penaltyterm bewiesen.

1. Nach Voraussetzung gilt $\underline{\psi} \leq 0 \leq \bar{\psi}$ auf $\partial\Omega$. Für alle $v \in \mathcal{U}$ liegt $\beta(v) = (v - \underline{\psi})^- + (v - \bar{\psi})^+$ in \mathcal{U} . Da $\underline{\psi} \leq \bar{\psi}$ in Ω , folgt für alle $v \in L^2(0, T; \mathcal{U})$

$$\begin{aligned} ((v(t) - \underline{\psi})^-, (v(t) - \bar{\psi})^+) &= a((v(t) - \underline{\psi})^-, (v(t) - \bar{\psi})^+) = 0, \\ a(\beta(v), \beta(v)) &= a((v(t) - \underline{\psi})^-, (v(t) - \underline{\psi})^-) + a((v(t) - \bar{\psi})^+, (v(t) - \bar{\psi})^+). \end{aligned} \quad (2.14)$$

Für die verallgemeinerte zeitliche Ableitung gemäss Definition 1 ergibt sich

$$\frac{d}{dt}\beta(v(t)) = \frac{d}{dt}(v - \underline{\psi})^- + \frac{d}{dt}(v - \bar{\psi})^+ = \begin{cases} \dot{v}(t) & \text{falls } \beta(v(t)) \neq 0, \\ 0 & \text{falls } \beta(v(t)) = 0. \end{cases}$$

Damit gilt

$$\frac{1}{2} \frac{d}{dt} |\beta(v(t))|^2 = \left(\frac{d}{dt} \beta(v(t)), \beta(v(t)) \right) = (\dot{v}(t), \beta(v(t))). \quad (2.15)$$

u_ϵ sei die Lösung der Penaltygleichung (2.13). Somit gilt für f.a. $t \in [0, T]$

$$(\dot{u}_\epsilon, v) + a(u_\epsilon, v) + \frac{1}{\epsilon} (\beta(u_\epsilon), v) = (f, v) \quad \forall v \in \mathcal{U}. \quad (2.16)$$

Testen wir (2.16) mit $\beta(u_\epsilon)$, so folgt mit der Koerzivität von $a(\cdot, \cdot)$, (2.14) und (2.15) dass

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |\beta(u_\epsilon(t))|^2 + \alpha \|\beta(u_\epsilon(t))\|^2 + \frac{1}{\epsilon} |\beta(u_\epsilon(t))|^2 \\ \stackrel{(2.15)}{\leq} (\dot{u}_\epsilon(t), \beta(u_\epsilon(t))) + a(\beta(u_\epsilon(t)), \beta(u_\epsilon(t))) + \frac{1}{\epsilon} (\beta(u_\epsilon(t)), \beta(u_\epsilon(t))) \\ \stackrel{(2.14)}{=} (\dot{u}_\epsilon(t), \beta(u_\epsilon(t))) + a((u_\epsilon(t) - \underline{\psi})^-, (u_\epsilon(t) - \underline{\psi})^-) \\ + a((u_\epsilon(t) - \bar{\psi})^+, (u_\epsilon(t) - \bar{\psi})^+) + \frac{1}{\epsilon} (\beta(u_\epsilon(t)), \beta(u_\epsilon(t))) \\ = (\dot{u}_\epsilon(t), \beta(u_\epsilon(t))) + a(u_\epsilon(t), \beta(u_\epsilon(t))) + (-A\underline{\psi}, (u_\epsilon(t) - \underline{\psi})^-) \\ + (-A\bar{\psi}, (u_\epsilon(t) - \bar{\psi})^+) + \frac{1}{\epsilon} (\beta(u_\epsilon(t)), \beta(u_\epsilon(t))) \\ \stackrel{(2.16)}{=} (f, \beta(u_\epsilon(t))) + (-A\underline{\psi}, (u_\epsilon(t) - \underline{\psi})^-) + (-A\bar{\psi}, (u_\epsilon(t) - \bar{\psi})^+). \end{aligned}$$

Weiter erhält man mit $|\beta(u_\epsilon(t))| = |(u_\epsilon(t) - \underline{\psi})^-| + |(u_\epsilon(t) - \bar{\psi})^+|$

$$\frac{1}{2} \frac{d}{dt} |\beta(u_\epsilon(t))|^2 + \alpha \|\beta(u_\epsilon(t))\|^2 + \frac{1}{\epsilon} |\beta(u_\epsilon(t))|^2 \leq (|f| + |A\underline{\psi}| + |A\bar{\psi}|) |\beta(u_\epsilon(t))|. \quad (2.17)$$

Die Integration von (2.17) über die Zeit führt zu

$$\frac{1}{2} |\beta(u_\epsilon(T))|^2 - \frac{1}{2} \underbrace{|\beta(u_\epsilon(0))|^2}_{=0} + \alpha \|\beta(u_\epsilon)\|_T^2 + \frac{1}{\epsilon} |\beta(u_\epsilon)|_T^2 \leq (|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T) |\beta(u_\epsilon)|_T.$$

Damit gilt

$$|\beta(u_\epsilon)|_T \leq \epsilon (|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T). \quad (2.18)$$

2. Multiplikation der Penaltygleichung $\dot{u}_\epsilon + Au_\epsilon + \frac{1}{\epsilon}\beta(u_\epsilon) = f$ mit Au_ϵ führt zu

$$\begin{aligned} & \underbrace{(\dot{u}_\epsilon(t), Au_\epsilon(t))}_{= \frac{1}{2} \frac{d}{dt} a(u_\epsilon(t), u_\epsilon(t))} + (Au_\epsilon(t), Au_\epsilon(t)) + \frac{1}{\epsilon} (\beta(u_\epsilon(t)), Au_\epsilon(t)) = (f, Au_\epsilon(t)). \end{aligned}$$

Nach Integration über die Zeit erhält man

$$\begin{aligned} 0 & \leq -\frac{1}{2} a(u_\epsilon(T), u_\epsilon(T)) + \frac{1}{2} a(u_\epsilon(0), u_\epsilon(0)) - |Au_\epsilon|_T^2 + \int_0^T \left(f - \frac{1}{\epsilon} \beta(u_\epsilon(t)), Au_\epsilon(t) \right) dt \\ & \leq \frac{1}{2} a(u_\epsilon(0), u_\epsilon(0)) - |Au_\epsilon|_T^2 + \left| f - \frac{1}{\epsilon} \beta(u_\epsilon) \right|_T |Au_\epsilon|_T. \end{aligned}$$

Damit folgt

$$\begin{aligned} |Au_\epsilon|_T & \leq \frac{1}{2} \left(\left| f - \frac{1}{\epsilon} \beta(u_\epsilon) \right|_T + \sqrt{\left| f - \frac{1}{\epsilon} \beta(u_\epsilon) \right|_T^2 + 2a(u_\epsilon(0), u_\epsilon(0))} \right) \\ & \leq |f|_T + \frac{1}{\epsilon} |\beta(u_\epsilon)|_T + \sqrt{\frac{1}{2} a(u_\epsilon(0), u_\epsilon(0))}. \end{aligned}$$

Mit (2.18) und der Stetigkeit von $a(.,.)$ ergibt sich

$$|Au_\epsilon|_T \leq 2|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T + \sqrt{\frac{c}{2}} \|u_0\|.$$

Im Limes $\epsilon \rightarrow 0$ folgt die Behauptung. □

Wird in die Penaltygleichung (2.16) $v = \dot{u}_\epsilon + Au_\epsilon - f$ eingesetzt, so kann eine Abschätzung für $|\dot{u} + Au - f|_T$ hergeleitet werden. Es gilt

$$\begin{aligned} |\dot{u}_\epsilon(t) + Au_\epsilon(t) - f|^2 & = -\frac{1}{2} (\beta(u_\epsilon(t)), \dot{u}_\epsilon(t) + Au_\epsilon(t) - f) \\ & \leq \frac{1}{\epsilon} |\beta(u_\epsilon(t))| \cdot |\dot{u}_\epsilon(t) + Au_\epsilon(t) - f|. \end{aligned}$$

Wir integrieren über die Zeit, verwenden (2.18) und erhalten

$$\begin{aligned} |\dot{u}_\epsilon + Au_\epsilon - f|_T^2 & \leq \frac{1}{\epsilon} |\beta(u_\epsilon)|_T |\dot{u}_\epsilon + Au_\epsilon - f|_T, \\ |\dot{u}_\epsilon + Au_\epsilon - f|_T & \leq (|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T). \end{aligned}$$

Folgendes Korollar ergibt sich im Limes $\epsilon \rightarrow 0$.

Korollar 2.6 *Seien $A\underline{\psi}, A\bar{\psi} \in \mathcal{H}$. Ist u die Lösung von (V), so gilt*

$$|\dot{u} + Au - f|_T \leq (|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T).$$

Nun soll $|\dot{u}|_T$ durch die Daten des Problems abgeschätzt werden. Dazu wird der Differentiationssatz von Lebesgue benötigt, der eine Aussage über Funktionen $g \in L^p_{loc}(\mathbb{R}^n)$ macht. $B(x_0, r)$ bezeichne die Kugel mit Mittelpunkt x_0 und Radius r . $|B(x_0, r)|$ sei das Lebesgue Mass dieser Kugel. Ist $1 \leq p < \infty$, so gilt für fast alle $x_0 \in \mathbb{R}^n$

$$\frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} |g(x) - g(x_0)|^p dx \xrightarrow{r \rightarrow 0} 0.$$

Insbesondere gilt für $p = 1$

$$\frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} g(x) dx \xrightarrow{r \rightarrow 0} g(x_0), \quad \text{f.f.a. } x_0 \in \mathbb{R}^n. \quad (2.19)$$

Ein Beweis des Differentiationssatzes von Lebesgue steht in G. Folland [9].

Lemma 2.7 *Ist u die Lösung der Variationsungleichung (V), so gibt es ein c (vgl. 2.10), das von t und u unabhängig ist, so dass*

$$|\dot{u}|_T \leq |f|_T + \sqrt{\frac{c}{2}} \|u_0\|.$$

Beweis: Sei $\tau > 0$. Mit

$$v(x, t) = \begin{cases} u(x, t - \tau) & \text{für } t \in (\tau, T], \\ u(x, t) & \text{für } t \in [0, \tau] \end{cases}$$

eingesetzt in (V) folgt nach Division durch τ

$$\int_0^T \left(\dot{u}, \frac{u(t - \tau) - u(t)}{\tau} \right) dt + \int_0^T a(u, \frac{u(t - \tau) - u(t)}{\tau}) dt - \int_0^T \left(f, \frac{u(t - \tau) - u(t)}{\tau} \right) dt \geq 0.$$

$u(x, \cdot)$ ist absolut stetig und somit ergibt sich

$$- \int_0^T \left(\dot{u}, \frac{1}{\tau} \int_{t-\tau}^t \dot{u}(s) ds \right) dt - \int_0^T a(u, \frac{1}{\tau} \int_{t-\tau}^t \dot{u}(s) ds) dt + \int_0^T \left(f, \frac{1}{\tau} \int_{t-\tau}^t \dot{u}(s) ds \right) dt \geq 0. \quad (2.20)$$

Nach Satz 2.4 gilt $\dot{u} \in L^\infty(0, T; \mathcal{H})$. Damit sind $|\dot{u}| \cdot \|\dot{u}\|_{L^\infty(0, T; \mathcal{H})}$ und $|f| \cdot \|\dot{u}\|_{L^\infty(0, T; \mathcal{H})}$ obere Schranken für den ersten bzw. dritten Integranden in (2.20). Da $\dot{u}(x, s) = 0$ für alle $(x, s) \in \partial\Omega \times [0, T]$ folgt nach partieller Integration und dem Satz von Fubini

$$\begin{aligned} a(u(t), \frac{1}{\tau} \int_{t-\tau}^t \dot{u}(s) ds) &= \int_{\Omega} \sum_{i,j=1}^n a^{ij} \frac{\partial}{\partial x_i} u(x, t) \frac{\partial}{\partial x_j} \frac{1}{\tau} \int_{t-\tau}^t \dot{u}(x, s) ds dx \\ &= - \int_{\Omega} \sum_{i,j} \frac{\partial}{\partial x_j} \left(a^{ij} \frac{\partial}{\partial x_i} u(x, t) \right) \frac{1}{\tau} \int_{t-\tau}^t \dot{u}(x, s) ds dx \\ &= - \frac{1}{\tau} \int_{t-\tau}^t \int_{\Omega} \sum_{i,j} \frac{\partial}{\partial x_j} \left(a^{ij} \frac{\partial}{\partial x_i} u(x, t) \right) \dot{u}(x, s) dx ds. \end{aligned}$$

Dies führt zu einer endlich integrierbaren oberen Schranke für den Integranden des zweiten Integrals in (2.20). λ_{\max}^{ij} bezeichne den grössten Eigenwert der Matrix a^{ij} . Es gilt

$$\begin{aligned} \left| a \left(u(t), \frac{1}{\tau} \int_{t-\tau}^t \dot{u}(s) ds \right) \right| &\leq \frac{1}{\tau} \int_{t-\tau}^t \int_{\Omega} \left| \sum_{i,j} \frac{\partial}{\partial x_j} \left(a^{ij} \frac{\partial}{\partial x_i} u(x, t) \right) \dot{u}(x, s) \right| dx ds \\ &\leq \frac{1}{\tau} \int_{t-\tau}^t \lambda_{\max}^{ij} \|u(t)\|_{H^2(\Omega)} |\dot{u}(s)| ds \\ &\leq \lambda_{\max}^{ij} \|u(t)\|_{H^2(\Omega)} \|\dot{u}\|_{L^\infty(0, T; \mathcal{H})}. \end{aligned}$$

Wir bilden in (2.20) den Limes $\tau \rightarrow 0$. Wegen des Lebesgueschen Konvergenzsatzes darf der Limes mit den Integralen vertauscht werden. \dot{u} ist endlich integrierbar. Zusammen mit dem Lebesgueschen Differentiationssatz (2.19) folgt aus (2.20)

$$- \int_0^T |\dot{u}|^2 dt - \int_0^T a(u, \dot{u}) dt + \int_0^T (f, \dot{u}) dt \geq 0.$$

Mit $a(u, \dot{u}) = \frac{1}{2} \frac{d}{dt} a(u, u)$ ergibt sich

$$\begin{aligned} |\dot{u}|_T^2 = \int_0^T |\dot{u}|^2 dt &\leq \frac{1}{2} a(u_0, u_0) - \frac{1}{2} a(u(T), u(T)) + \int_0^T (f, \dot{u}) dt \\ &\leq \frac{1}{2} c \|u_0\|^2 + |f|_T |\dot{u}|_T. \end{aligned}$$

Durch Auflösen dieser quadratischen Ungleichung nach $|\dot{u}|_T$ folgt

$$|\dot{u}|_T \leq \frac{1}{2} \left(|f|_T + \sqrt{|f|_T^2 + 2c \|u_0\|^2} \right) \leq |f|_T + \sqrt{\frac{c}{2}} \|u_0\|.$$

□

Lemma 2.8 *Ist u die Lösung der Variationsungleichung (V), so gilt*

$$\begin{aligned} \|\dot{u}\|_{L^\infty(0, T; \mathcal{H})} &\leq |f - Au_0|, \\ \|\dot{u}\|_T &\leq \sqrt{\frac{1}{2\alpha}} |f - Au_0|. \end{aligned}$$

Beweis: Seien $w := \dot{u}_\epsilon$ und $w_I := w$ falls $\beta(u_\epsilon) \neq 0$ und $w_I = 0$ sonst. Leiten wir die Penaltygleichung nach der Zeit ab und multiplizieren mit w , so folgt

$$\begin{aligned} \left(\frac{d}{dt} w, w \right) + a(w, w) + \frac{1}{\epsilon} (w_I, w) &= 0, \\ \frac{1}{2} \frac{d}{dt} |w|^2 + \alpha \|w\|^2 + \frac{1}{\epsilon} |w_I|^2 &\leq 0, \\ \frac{1}{2} \frac{d}{dt} |w|^2 + \alpha \|w\|^2 &\leq 0. \end{aligned} \tag{2.21}$$

$|w(t)|^2$ ist somit monoton fallend und mit $w(0) = \dot{u}_\epsilon(0) = \dot{u}_0 = f - Au_0$ folgt für alle $t \in (0, T]$

$$|\dot{u}_\epsilon(t)| \leq |f - Au_0|. \tag{2.22}$$

Nach Integration von (2.21) über t erhält man zusammen mit (2.22)

$$\|\dot{u}_\epsilon\|_T \leq \sqrt{\frac{1}{2\alpha}} |f - Au_0|.$$

Im Limes $\epsilon \rightarrow 0$ folgen die Behauptungen aus den letzten zwei Ungleichungen. \square

Lemma 2.9 *Seien $A\underline{\psi}, A\overline{\psi} \in \mathcal{H}$. Ist u die Lösung von (V), so gilt für alle $t \in (0, T)$*

$$|Au(t)| \leq |f| + |f - Au_0| + |A\underline{\psi}| + |A\overline{\psi}|.$$

Beweis: Seien $\underline{I}(t), \overline{I}(t)$ die Koinzidenzmengen des oberen bzw. unteren Hindernisses, d.h.

$$\begin{aligned} u(x, t) &= \underline{\psi}(x) && \text{für } x \in \underline{I}(t), \\ u(x, t) &= \overline{\psi}(x) && \text{für } x \in \overline{I}(t). \end{aligned}$$

Wie aus der Penaltygleichung (2.13) leicht zu sehen ist, gilt für ein festes $t \in (0, T)$ und für fast alle $x \in \Omega \setminus (\underline{I} \cup \overline{I})$ die Gleichung

$$\dot{u}(x, t) + Au(x, t) = f(x).$$

Damit folgt

$$\begin{aligned} |Au(t)|^2 &= \|Au(t)\|_{L^2(\Omega \setminus (\underline{I} \cup \overline{I}))}^2 + \|Au(t)\|_{L^2(\underline{I})}^2 + \|Au(t)\|_{L^2(\overline{I})}^2 \\ &= \|f - \dot{u}(t)\|_{L^2(\Omega \setminus (\underline{I} \cup \overline{I}))}^2 + \|A\underline{\psi}\|_{L^2(\underline{I})}^2 + \|A\overline{\psi}\|_{L^2(\overline{I})}^2. \end{aligned}$$

Mit dem Lemma 2.8 und der Dreiecksungleichung folgt die Behauptung. \square

Im folgenden Lemma wird gezeigt, dass die Lösung u von (V) für fast alle $t \in [0, T]$ auch eine vereinfachte Variationsungleichung erfüllt.

Lemma 2.10 *Ist u Lösung von (V), so erfüllt sie für fast alle $t \in [0, T]$ die Variationsungleichung*

$$(\dot{u}(t) - f, \tilde{v} - u(t)) + a(u(t), \tilde{v} - u(t)) \geq 0 \quad \forall \tilde{v} \in \tilde{K} := \{w \in \mathcal{U} : \underline{\psi} \leq w \leq \overline{\psi}, \text{ f.ü. in } \Omega\}.$$

Beweis: Für ein beliebiges $t_0 \in (0, T)$ und $\epsilon > 0$ so, dass $[t_0 - \epsilon, t_0 + \epsilon] \subset (0, T)$, sei

$$w(x, t) = \begin{cases} \tilde{v}(x) & \text{für } t \in [t_0 - \epsilon, t_0 + \epsilon], \\ u(x, t) & \text{sonst,} \end{cases}$$

dabei ist $\tilde{v} \in \tilde{K}$ beliebig. Weil $u \in K$ und $\tilde{v} \in \tilde{K}$, gilt $w \in K$. Einsetzen von w für v in die Variationsungleichung (V) führt zu

$$\int_{t_0 - \epsilon}^{t_0 + \epsilon} (\dot{u} - f, \tilde{v} - u) dt + \int_{t_0 - \epsilon}^{t_0 + \epsilon} a(u, \tilde{v} - u) dt \geq 0 \quad \forall \tilde{v} \in \tilde{K}. \quad (2.23)$$

Es gilt

$$\begin{aligned} \int_0^T |(\dot{u} - f, \tilde{v} - u)| dt &\leq \int_0^T |\dot{u} - f| \cdot |\tilde{v} - u| dt \\ &\leq \frac{1}{2} \int_0^T |\dot{u} - f|^2 + |\tilde{v} - u|^2 dt. \end{aligned}$$

Wegen Satz 2.4 folgt

$$\int_0^T |(\dot{u} - f, \tilde{v} - u)| dt < \infty. \quad (2.24)$$

Für $a(u, \tilde{v} - u)$ gilt wegen der Stetigkeit von $a(.,.)$ und da $u \in L^2(0, T; \mathcal{U})$, dass

$$\begin{aligned} \int_0^T |a(u, \tilde{v} - u)| dt &\leq \int_0^T c \|u\| \cdot \|\tilde{v} - u\| dt \\ &\leq c \left(\frac{1}{2} \int_0^T \|u\|^2 + \|\tilde{v} - u\|^2 dt \right) < \infty \end{aligned} \quad (2.25)$$

Wegen (2.24) und (2.25) gilt $\int_0^T |(\dot{u} - f, \tilde{v} - u) + a(u, \tilde{v} - u)| dt < \infty$. Damit sind die Voraussetzungen des Lebesgueschen Differentiationssatzes (2.19) erfüllt. Wir dividieren (2.23) durch 2ϵ und lassen ϵ gegen 0 streben. Mit dem erwähnten Satz folgt für fast alle $t_0 \in (0, T)$

$$(\dot{u}(t_0) - f, \tilde{v} - u(t_0)) + a(u(t_0), \tilde{v} - u(t_0)) \geq 0 \quad \forall \tilde{v} \in \tilde{K}.$$

□

Schwache Variationsungleichungen und Eigenschaften derer Lösungen

Die Problemstellung (V) macht nur Sinn, falls u in \mathcal{D} liegt. Oft ist es jedoch schwierig, diese Bedingung a priori zu erfüllen. Deshalb diskutiert Lions in [13] eine schwächere Form von (V).

Halbschwache Variationsungleichung: Finde $u \in K$ mit

$$\int_0^T (\dot{v}, v - u) dt + \int_0^T a(u, v - u) dt \geq \int_0^T (f, v - u) dt \quad \forall v \in K \cap \mathcal{D}. \quad (V')$$

Lions beweist in [13] die Existenz und die Eindeutigkeit einer Lösung von (V').

In dieser Arbeit wird eine noch schwächere Form untersucht. Da $a(.,.)$ koerzitiv (2.11) ist, folgt

$$\begin{aligned} \int_0^T a(v, v - u) dt &= \int_0^T a(u, v - u) dt + \int_0^T \underbrace{a(v - u, v - u)}_{\geq 0} dt \\ &\geq \int_0^T a(u, v - u) dt. \end{aligned}$$

Damit motivieren wir die folgende schwache Form von (V).

Schwache Variationsungleichung: Finde $u \in K$ mit $u(x, 0) = u_0(x)$ so, dass

$$\int_0^T (\dot{v}, v - u) dt + \int_0^T a(v, v - u) dt \geq \int_0^T (f, v - u) dt \quad \forall v \in K \cap \mathcal{D}. \quad (\text{V}''')$$

Im folgenden Lemma zeigen wir, dass die Lösung von (V) und (V''') übereinstimmen, wenn sie existieren.

Lemma 2.11 *Seien u und \tilde{u} Lösungen von (V) bzw. (V'''). Es gilt*

$$u = \tilde{u} \text{ in } L^2(0, T; \mathcal{U}) \cap L^\infty(0, T; \mathcal{H}).$$

Beweis: Der Beweis ist in drei Teile gegliedert. Zuerst wird gezeigt, dass die Lösung der starken Form auch Lösung der schwachen Form ist. Im zweiten Teil beweisen wir, dass die schwache Form genau eine Lösung in $K \cap \mathcal{D}$ besitzt. Zum Schluss wird gezeigt, dass alle Lösungen der schwachen Form identisch mit der Lösung in $K \cap \mathcal{D}$ und somit gleich der Lösung der starken Form sind.

1. Sei u Lösung der starken Variationsungleichung (V). Wenn die Testfunktionenmenge auf $K \cap \mathcal{D}$ eingeschränkt wird, gilt für alle $v \in K \cap \mathcal{D}$ immer noch

$$\begin{aligned} \int_0^T (\dot{u} - \dot{v} + \dot{v} - f, v - u) dt + \int_0^T a(u - v + v, v - u) dt &\geq 0, \\ \int_0^T (\dot{v} - f, v - u) dt - \int_0^T \frac{1}{2} \frac{d}{dt} |v - u|^2 dt & \\ + \int_0^T a(v, v - u) dt - \int_0^T a(v - u, v - u) dt &\geq 0. \end{aligned}$$

Da $u, v \in K \cap \mathcal{D}$ und $a(\cdot, \cdot)$ koerzitiv mit Parameter α ist, folgt

$$\begin{aligned} \int_0^T (\dot{v} - f, v - u) dt + \int_0^T a(v, v - u) dt &\geq \\ \frac{1}{2} |v(T) - u(T)|^2 + \alpha \|v - u\|_T^2 &\geq 0. \end{aligned}$$

Somit ist u auch Lösung der schwachen Form (V''').

2. Da die Lösung u in $K \cap \mathcal{D}$ liegt und Lösung der schwachen Form ist, folgt die Existenz einer Lösung von (V''') in $K \cap \mathcal{D}$.

Seien $w, \tilde{w} \in K \cap \mathcal{D}$ Lösungen der schwachen Form (V'''). Sei $\theta \in (0, 1)$. Da $K \cap \mathcal{D}$ konvex ist, kann (V''') für w mit $v = (1 - \theta)w + \theta\tilde{w}$ und (V''') für \tilde{w} mit $v = (1 - \theta)\tilde{w} + \theta w$ getestet werden:

$$\begin{aligned} \int_0^T ((1 - \theta)\dot{w} + \theta\dot{\tilde{w}} - f, \theta(\tilde{w} - w)) dt + \int_0^T a((1 - \theta)w + \theta\tilde{w}, \theta(\tilde{w} - w)) dt &\geq 0, \\ \int_0^T ((1 - \theta)\dot{\tilde{w}} + \theta\dot{w} - f, \theta(w - \tilde{w})) dt + \int_0^T a((1 - \theta)\tilde{w} + \theta w, \theta(w - \tilde{w})) dt &\geq 0. \end{aligned}$$

Durch Addition der beiden Ungleichungen und Division durch θ erhält man

$$\int_0^T ((1-\theta)\dot{w} + \theta\dot{\tilde{w}} - (1-\theta)\dot{w} - \theta\dot{w}, \tilde{w} - w) dt + \int_0^T a((1-\theta)w + \theta\tilde{w} - (1-\theta)\tilde{w} - \theta w, \tilde{w} - w) dt \geq 0.$$

Im Limes $\theta \rightarrow 0$ folgt mit dem Lemma 2.3

$$\begin{aligned} & \int_0^T (\dot{w} - \dot{\tilde{w}}, \tilde{w} - w) dt + \int_0^T a(w - \tilde{w}, \tilde{w} - w) dt \geq 0, \\ & 0 \geq \int_0^T \frac{1}{2} \frac{d}{dt} |\tilde{w} - w|^2 dt + \int_0^T a(\tilde{w} - w, \tilde{w} - w) dt \\ & \geq \frac{1}{2} |\tilde{w}(T) - w(T)|^2 + \alpha \|\tilde{w} - w\|_T^2 \geq 0. \end{aligned}$$

Also ist $\tilde{w} = w$ in $L^2(0, T; \mathcal{U}) \cap L^\infty(0, T; \mathcal{H})$.

3. w sei die eindeutige Lösung der schwachen Form (V'') in $K \cap \mathcal{D}$ gemäss zweitem Teil. Da die Lösung u von (V) in $K \cap \mathcal{D}$ liegt und auch (V'') löst, ist u identisch mit w . Sei nun $\bar{w} \in K$ eine weitere Lösung von (V''). Da $K \cap \mathcal{D}$ in K dicht liegt, existiert eine Folge $(w_k)_{k=1}^\infty \subset K \cap \mathcal{D}$, die gegen \bar{w} strebt. Wir wählen eine Teilfolge aus, die wieder mit (w_k) bezeichnet wird, für die gilt

$$\|w_k - \bar{w}\|_T \leq \frac{1}{k^2} \quad \forall k.$$

Aus (V'') folgt für alle $v \in K \cap \mathcal{D}$

$$\begin{aligned} & \int_0^T (\dot{v} - f, v - w_k) dt + \int_0^T a(v, v - w_k) dt \\ & + \int_0^T (\dot{v} - f, w_k - \bar{w}) dt + \int_0^T a(v, w_k - \bar{w}) dt \geq 0. \end{aligned} \quad (2.26)$$

w erfüllt (V). Durch Einsetzen von $v = w_k$ in diese Ungleichung und $v = (1 - \frac{1}{k})w_k + \frac{1}{k}w$ in (2.26) folgt

$$\begin{aligned} & \int_0^T (\dot{w} - f, w_k - w) dt + \int_0^T a(w, w_k - w) dt \geq 0, \\ & \int_0^T \left(\left(1 - \frac{1}{k}\right) \dot{w}_k + \frac{1}{k} \dot{w} - f, \frac{1}{k}(w - w_k) \right) dt + \int_0^T a \left(\left(1 - \frac{1}{k}\right) w_k + \frac{1}{k} w, \frac{1}{k}(w - w_k) \right) dt \\ & + \int_0^T \left(\left(1 - \frac{1}{k}\right) \dot{w}_k + \frac{1}{k} \dot{w} - f, w_k - \bar{w} \right) dt + \int_0^T a \left(\left(1 - \frac{1}{k}\right) w_k + \frac{1}{k} w, w_k - \bar{w} \right) dt \geq 0. \end{aligned}$$

Multiplikation der zweiten Ungleichung mit k und Addition der beiden führt zu

$$\begin{aligned} & \int \left(1 - \frac{1}{k}\right) (\dot{w}_k - \dot{w}, w - w_k) + \left(\left(1 - \frac{1}{k}\right) \dot{w}_k + \frac{1}{k} \dot{w} - f, k(w_k - \bar{w}) \right) dt \\ & + \int a \left(\left(1 - \frac{1}{k}\right) (w_k - w), w - w_k \right) + a \left(\left(1 - \frac{1}{k}\right) w_k + \frac{1}{k} w, k(w_k - \bar{w}) \right) dt \geq 0. \end{aligned}$$

Da $w_k, w \in \mathcal{D}$ und wegen der Koerzivitat von $a(\cdot, \cdot)$ gilt fur $k \rightarrow \infty$

$$\begin{aligned} 0 &\leq \frac{1}{2} \lim_{k \rightarrow \infty} |w_k(T) - w(T)|^2 + \alpha \lim_{k \rightarrow \infty} \|w_k - w\|_T^2 \\ &\leq \lim_{k \rightarrow \infty} \left[\int_0^T \left(\left(1 - \frac{1}{k}\right) \dot{w}_k + \frac{1}{k} \dot{w} - f, k(w_k - \bar{w}) \right) dt \right. \\ &\quad \left. + \int_0^T a \left(\left(1 - \frac{1}{k}\right) w_k + \frac{1}{k} w, k(w_k - \bar{w}) \right) dt \right] = A \end{aligned}$$

Mit der Schwarzschen Ungleichung und der Stetigkeit von $a(\cdot, \cdot)$ folgt weiter

$$\begin{aligned} |A| &\leq \lim_{k \rightarrow \infty} \left[\left\| \left(1 - \frac{1}{k}\right) \dot{w}_k + \frac{1}{k} \dot{w} - f \right\|_T \underbrace{k \|w_k - \bar{w}\|_T}_{\leq \frac{1}{k^2}} \right. \\ &\quad \left. + c \left\| \left(1 - \frac{1}{k}\right) w_k + \frac{1}{k} w \right\|_T \underbrace{k \|w_k - \bar{w}\|_T}_{\leq \frac{1}{k^2}} \right] = 0. \end{aligned}$$

Schliesslich gilt

$$\begin{aligned} \|w_k - w\|_T &\xrightarrow{k \rightarrow \infty} 0, \\ \|w_k - w\|_{L^\infty(0, T; \mathcal{H})} &\xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Mit der Dreiecksungleichung angewendet auf $\|w - w_k + w_k - \bar{w}\|_T$ folgt die Behauptung. □

2.2 Der Raum der finiten Elemente

Als endlichdimensionaler Unterraum \mathcal{U}_n von \mathcal{U} konnen die Raume der finiten Elemente verwendet werden. Das Gebiet Ω wird in endlich viele Teilgebiete zerlegt. Man betrachtet Funktionen, die auf jedem Teilgebiet Polynome sind. Fur die folgende Beschreibung dieser Funktionenraume sei Ω eine beschrankte, polygonale Teilmenge von \mathbb{R}^2 . Ist $\Omega \subset \mathbb{R}$ oder $\Omega \subset \mathbb{R}^N$, so konnen alle Definitionen angepasst werden.

2.2.1 Basisfunktionen

Definition 2 *Eine zulassige Triangulation \mathcal{T} einer polygonalen Menge $\Omega \subset \mathbb{R}^2$ ist eine Zerlegung in Dreiecke $\mathcal{T} = \{T_1, \dots, T_M\}$, so dass keine Ecke eines Dreiecks im Innern einer Kante eines anderen liegt.*

Im Fall $\Omega \subset \mathbb{R}^3$ musste man die Dreiecke durch Tetraeder ersetzen.

Um zwei Triangulationen bezuglich der Feinheit miteinander vergleichen zu konnen, fuhren wir die Maschenweite h ein. Es sei $h(T_j)$ der Durchmesser und $\rho(T_j)$ der Inkreisradius des Dreiecks $T_j \in \mathcal{T}$. Wir bezeichnen mit $h = \max_{T_j \in \mathcal{T}} h(T_j)$ den grossten Durchmesser und mit $\rho = \min_{T_j \in \mathcal{T}} \rho(T_j)$ den kleinsten Inkreisradius der Dreiecke aus der Triangulation \mathcal{T} .

Definition 3 Eine Triangulation \mathcal{T} heisst quasiuniform, wenn es eine Zahl $\kappa > 0$ mit

$$\frac{h(T_j)}{\rho(T_j)} \leq \kappa \quad \forall T_j \in \mathcal{T}$$

gibt. Eine Triangulation mit $\frac{h}{\rho} \leq \kappa$ heisst uniform.

In einer quasiuniformen Triangulation können die Dreiecke bei konstantem Durchmesser des Umkreises nicht beliebig flach werden.

Eine Funktion $v_n \in \mathcal{U}_n$ soll nun aus Funktionen zusammengesetzt sein, die auf einem Dreieck T_i Polynome sind. $\mathcal{P}_t(G)$ bezeichne die Menge aller Polynome vom Grad kleinergleich t über G . v_n kann durch Funktionswerte an einzelnen Punkten bestimmt werden. Besteht v_n eingeschränkt auf ein T_i zum Beispiel aus einem Polynom vom Grad 1, d.h. $v_n|_{T_i} \in \mathcal{P}_1(T_i)$, so ist $v_n|_{T_i}$ durch die drei Funktionswerte an den Eckpunkten von T_i eindeutig bestimmt. In diesem Fall stimmen die Funktionswerten von v_n am Rand von T_i mit den Funktionswerten auf dem Rand des benachbarten Dreiecks überein, d.h. die Funktion v_n ist in Ω stetig.

Ist $v_n|_{T_i}$ ein Polynom vom Grad t , so kann es eindeutig durch die Vorgabe von Funktionswerten an $(t+1)(t+2)/2$ verschiedenen Stellen bestimmt werden. Werden auch noch die partiellen Ableitungen an diesen Stellen vorgegeben, so kann die Anzahl der Stellen reduziert werden. Die Tabelle 2.1 zeigt eine kleine Auswahl von finiten Elementen. Dabei bedeutet ein Punkt, dass der Funktionswert zur Bestimmung des Polynoms vorgegeben wird. Bei einem bzw. zwei Kreisen um einen Punkt, werden auch die ersten bzw. ersten und zweiten Ableitungen vorgegeben. Wird für das Polynom die Normalenableitung auf einer Kante benötigt, so ist ein Querstrich auf der Kante gezeichnet.

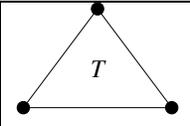
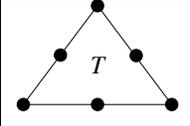
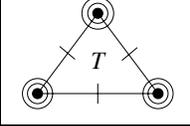
	lineares Element 3 Freiheitsgrade	$v_n _T \in \mathcal{P}_1$	$v_n \in C^0(\Omega)$
	quadratisches Element 6 Freiheitsgrade	$v_n _T \in \mathcal{P}_2$	$v_n \in C^0(\Omega)$
	Argyris Dreieck 21 Freiheitsgrade	$v_n _T \in \mathcal{P}_5$	$v_n \in C^1(\Omega)$

Tabelle 2.1: Einige finite Elemente

Soll die Funktion v durch $v_n \in \mathcal{U}_n$ approximiert werden, so kann v_n durch Vorgabe von Funktionswerten und partiellen Ableitungen von v an bestimmten Stellen eindeutig bestimmt werden. Gilt $v \in H^p(\Omega)$, so müssen die Funktionswerte und die partiellen Ableitungen nicht überall existieren. Wie gross muss p sein, damit diese an beliebigen Stellen existieren? Die Antwort liefert ein Einbettungssatz von Sobolev (vgl. [1]).

Eine Teilmenge $\Omega \subset \mathbb{R}^N$ erfüllt die Kegelbedingung, wenn zu jedem $x \in \partial\Omega$ ein Kegel mit Spitze in x , echt positivem Öffnungswinkel und einer echt positiven Höhe existiert, welcher ganz in Ω liegt. Existieren Öffnungswinkel und Höhe unabhängig von $x \in \partial\Omega$, so erfüllt Ω die

Kegelbedingung gleichmässig. Für ein solches Ω besagt der Einbettungssatz, dass mit $p - \frac{N}{2} > q$ die Einbettung $H^p(\Omega) \hookrightarrow C^q(\bar{\Omega})$ stetig ist.

Sei u die Lösung der Variationsungleichung (V). Der Regularitätssatz 2.4 besagt, dass $u(t)$ f.f.a. $t \in (0, T)$ in $H^2(\Omega)$ liegt. Mit dem Einbettungssatz folgt, dass für $\Omega \subset \mathbb{R}$ der Funktionswert und die erste Ableitung an einer bestimmten Stelle sinnvoll bestimmt werden kann. Für $\Omega \subset \mathbb{R}^2$ bzw. \mathbb{R}^3 können nur die Funktionswerte von $v \in H^2(\Omega)$ sinnvoll bestimmt werden. Daher verwenden wir einen Raum von finiten Elementen, der nur die Funktionswerte an bestimmten Stellen verwendet.

Seien $\{x^i\}_{i=1}^{\bar{n}} \subset \Omega$ die Eckpunkte der Triangulation \mathcal{T} . Die Funktionen $\phi_j(x), j = 1, \dots, \bar{n}$, die bei einem Punkt x^j den Wert 1 bei allen anderen Punkten x^i 0 annehmen und dazwischen linear verlaufen, heissen Pyramidenfunktionen. Es gilt

$$\phi_j(x^i) = \delta_{ij}.$$

Die Pyramidenfunktionen bilden eine Basis für den Raum der stückweise linearen finiten Elemente. Die Dimension dieses endlichdimensionalen Unterraumes ist gleich der Anzahl Eckpunkte x^i . Da $\mathcal{U} = H_0^1(\Omega)$ approximiert werden soll, müssen die Funktionen auf dem Rand von Ω verschwinden. Daher verwenden wir nur die Basisfunktionen, welche an einer Stelle x^i im Innern von Ω den Wert 1 annehmen. Sei n die Anzahl Ecken x^i im Innern von Ω .

Definition 4 *Der endlichdimensionale Unterraum S_h von \mathcal{U} wird aufgespannt durch die Basisfunktionen $\{\phi_j\}_{j=1}^n$. Ist $v_n \in S_h$, so existiert ein $y \in \mathbb{R}^n$ mit $v_n(x) = \sum_{i=1}^n y_i \phi_i(x)$.*

2.2.2 Interpolationsfehler

Der Funktionenraum $H_0^1(\Omega) \cap H^2(\Omega)$ kann durch den oben eingeführten Raum der finiten Elemente S_h approximiert werden. Eine bekannte Eigenschaft von S_h (vgl. P. G. Ciarlet [6]) ist die Folgende. Für alle $v \in H_0^1(\Omega) \cap H^2(\Omega)$ existiert ein Element v_n aus $S_h(\Omega)$ mit

$$\|v_n - v\| \leq ch \|v\|_{H^2}, \quad (2.27)$$

$$|v_n - v| \leq ch^2 \|v\|_{H^2}. \quad (2.28)$$

Bei den Variationsungleichungen muss v_n nicht nur in $S_h(\Omega)$, sondern auch in der konvexen Menge \tilde{K}_n liegen. Dies muss für das v_n in (2.27) und (2.28) nicht zutreffen. Die Abbildung 2.1 zeigt eine Funktion $v \in H^2(0, 1)$ und ihre Projektion bezüglich der L^2 -Norm in den Raum der stückweise linearen finiten Elemente mit $h = \frac{1}{6}$. Dabei erfüllt v die Hindernisbedingung $\underline{\psi}(x) \equiv 0 \leq v(x)$. Die Projektion von v in S_h ist die beste Approximation von v . Sie erfüllt diese Bedingung jedoch nicht und liegt somit nicht in \tilde{K}_n .

v soll durch ein Element aus \tilde{K}_n approximiert werden. Wegen des Einbettungssatzes ist v in den Eckpunkten $\{x^i\}$ der Triangulation \mathcal{T} definiert.

Definition 5 *Die Abbildung*

$$\Pi : \mathcal{U} \rightarrow S_h, \quad v \longmapsto \Pi v = \sum_{i=1}^n v(x^i) \phi_i$$

heisst Interpolationsoperator. Die Funktion Πv ist die Interpolierende von v .

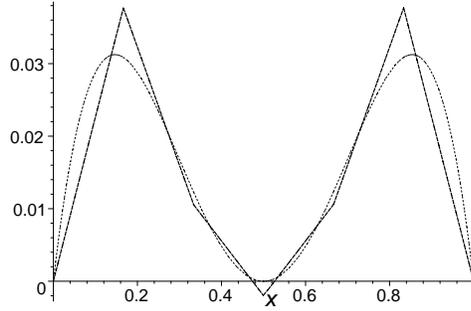


Abbildung 2.1: Projektion in den Raum der stückweise linearen finiten Elemente.

Wie gut approximiert nun Πv die Funktion v ? Dazu folgen wir der Argumentation von Braess [4]. Das nächste Lemma liefert eine Fehlerabschätzung für die Interpolation durch Polynome. Wir bezeichnen mit $|v|_{H^t}$ die Halbnorm in H^t , d.h.

$$|v|_{H^t}^2 = \int_{\Omega} \sum_{|\alpha|=t} (D^\alpha v(x))^2 dx.$$

Lemma 2.12 Sei $G \subset \mathbb{R}^2$ ein Gebiet mit Lipschitz-stetigem Rand. Sei $u \in H^t(G)$ mit $t \geq 2$. Seien $s = t(t+1)/2$ Punkte x^1, \dots, x^s in G gegeben, an denen die Interpolation $\Pi : H^t \rightarrow \mathcal{P}_{t-1}$ wohldefiniert ist. Dann existiert eine Konstante $c(G, x^1, \dots, x^s)$ mit

$$\|u - \Pi u\|_{H^t} \leq c|u|_{H^t} \quad \forall u \in H^t(G).$$

Im Beweis wird gezeigt, dass die Norm

$$\|v\| := |v|_{H^t} + \sum_{i=1}^s |v(x^i)|$$

äquivalent zur Norm $|v|_{H^t}$ ist. Damit folgt für ein $u \in H^t(G)$, da $u(x^i) = \Pi u(x^i)$ und $D^\alpha \Pi u = 0$ mit $|\alpha| = t$,

$$\begin{aligned} \|u - \Pi u\|_{H^t} &\leq c(G) \|u - \Pi u\| \\ &= c(G) \left(|u - \Pi u|_{H^t} + \sum_{i=1}^s |u(x^i) - \Pi u(x^i)| \right) \\ &= c(G) |u - \Pi u|_{H^t} = c(G) |u|_{H^t}. \end{aligned}$$

Um eine Fehlerabschätzung für die Interpolierende zu erhalten, möchten wir das Lemma 2.12 auf die einzelnen Dreiecke der Triangulation \mathcal{T} anwenden und aufsummieren. Das Problem ist nun, dass die Konstante c vom jeweiligen Gebiet abhängt.

Folgende Idee umgeht dieses Problem. Sei T_{ref} das rechtwinklige, gleichschenklige Dreieck mit den Katheten der Länge 1. Sei $\mathcal{T} = \{T_1, \dots, T_M\}$ eine quasiuniforme Triangulation von Ω . Für jedes $T_i \in \mathcal{T}$ werden drei Schritte durchgeführt:

1. Transformieren T_i affin in T_{ref} .
2. Wenden Lemma 2.12 an.
3. Transformieren T_{ref} affin zurück in T_i .

Was passiert bei der Transformation mit den Normen? Seien G, \hat{G} affin äquivalent, d.h. es existiert eine affine Abbildung $F : \hat{G} \rightarrow G, F(\hat{x}) = x_0 + B\hat{x}$ mit B regulär. Für $v \in H^m(G)$ ist dann $\hat{v} := v \circ F \in H^m(\hat{G})$. Es gilt die Transformationsformel

$$|\hat{v}|_{H^m(\hat{G})} \leq c(n, m) \|B\|^m |\det B|^{-1/2} |v|_{H^m(G)}. \quad (2.29)$$

Der Beweis basiert im Wesentlichen auf der Transformationsformel von Mehrfachintegralen.

Beim Hin- und Hertransformieren in quasiuniformen Gittern verlieren wir keine h -Potenzen. Seien $F : \hat{T} \rightarrow T$ die affine Transformation, \hat{r} und r die Umkreisradien von \hat{T} bzw. T und $\hat{\rho}, \rho$ die Inkreisradien der entsprechenden Dreiecke. Für $d \in \mathbb{R}^2$ mit $|d| \leq 2\hat{\rho}$ existieren $\hat{y}, \hat{z} \in \hat{T}$ mit $d = \hat{y} - \hat{z}$. Da $F(\hat{y}), F(\hat{z}) \in T$, gilt

$$\begin{aligned} |F(\hat{y}) - F(\hat{z})| &= |Bd| \leq 2r, \\ \|B\| &\leq \frac{r}{\hat{\rho}}. \end{aligned}$$

Da \mathcal{T} quasiuniform ist, folgt mit der gleichen Überlegung für B^{-1}

$$\|B\| \cdot \|B^{-1}\| \leq \frac{r}{\hat{\rho}} \frac{\hat{r}}{\rho} \leq \kappa \frac{\hat{r}}{\rho}.$$

Satz 2.13 (Braess [4]) *Sei $t \geq 2$ und \mathcal{T} eine quasiuniforme Triangulation von Ω . Dann gilt für die Interpolation durch stückweise Polynome vom Grad $t-1$ mit einer Konstante $c(\Omega, \kappa, t)$:*

$$\|u - \Pi u\|_{H^m(\Omega)} = \sqrt{\sum_{T_i \in \mathcal{T}} \|u - \Pi u\|_{H^m(T_i)}^2} \leq ch^{t-m} |u|_{H^t(\Omega)} \quad \text{für } u \in H^t(\Omega), 0 \leq m \leq t.$$

Beweis: Für T_{ref} gilt $r_{ref} = 1$ und $\rho_{ref} = \frac{1}{1+\sqrt{2}} \geq \frac{2}{5}$. Sei $T_i \in \mathcal{T}$ beliebig und $F : T_i \rightarrow T_{ref}$ sei die affine Transformation. Für den Interpolationsfehler gilt

$$\begin{aligned} |u - \Pi u|_{H^m(T_i)} &\stackrel{(2.29)}{\leq} c_1 \|B\|^m |\det B|^{-1/2} |\hat{u} - \Pi \hat{u}|_{H^m(T_{ref})} \\ &\leq c_1 \|B\|^m |\det B|^{-1/2} \|\hat{u} - \Pi \hat{u}\|_{H^t(T_{ref})} \\ &\stackrel{L2.12}{\leq} c_2 \|B\|^m |\det B|^{-1/2} |\hat{u}|_{H^t(T_{ref})} \\ &\stackrel{(2.29)}{\leq} c_3 \|B\|^m |\det B|^{-1/2} \|B^{-1}\|^t |\det B|^{1/2} |u|_{H^t(T_i)} \\ &\leq c_3 (\|B\| \cdot \|B^{-1}\|)^m \|B^{-1}\|^{t-m} |u|_{H^t(T_i)}. \end{aligned}$$

Mit $\|B\| \cdot \|B^{-1}\| \leq \kappa(1 + \sqrt{2})$ und $\|B^{-1}\| \leq \frac{r}{\rho_{ref}} \leq \frac{5}{2}r \leq 5h$ folgt für alle $l \leq m \leq t$

$$|u - \Pi u|_{H^l(T_i)} \leq ch^{t-l} |u|_{H^t(T_i)}.$$

Durch Quadrieren und Summieren über alle l von 0 bis m folgt für $h < 1$

$$\|u - \Pi u\|_{H^m(T_i)} \leq ch^{t-m}|u|_{H^t(T_i)}.$$

Summation über alle $T_i \in \mathcal{T}$ führt zur Behauptung. \square

Die Verallgemeinerung des Lemmas 2.12 ist das Lemma von Bramble und Hilbert (siehe [4]). Damit kann der Satz 2.13 für Gebiete $\Omega \subset \mathbb{R}^N$ bewiesen werden.

In dieser Arbeit verwenden wir ausschliesslich stückweise lineare finite Elemente. S_h besteht somit aus allen Funktionen, die eingeschränkt auf ein Dreieck $T_i \in \mathcal{T}$ linear und in ganz Ω stetig sind.

2.2.3 Eigenwertabschätzung des elliptischen Operators in S_h

$a(.,.)$ beschreibt eine elliptische Bilinearform auf S_h . Nach dem Darstellungssatz von Riesz existiert ein $A_n : S_h \rightarrow S'_h$ mit

$$a(u_n, v_n) = (A_n u_n, v_n) \quad \forall u_n, v_n \in S_h. \quad (2.30)$$

Nach der Definition 4 lassen sich u_n, v_n schreiben als $u_n(x) = \sum_{i=1}^n y_i \phi_i(x)$ und $v_n(x) = \sum_{i=1}^n z_i \phi_i(x)$. Damit kann (2.30) mit $A_n u_n := \sum_{i=1}^n y_{A_n, i} \phi_i$ geschrieben werden als

$$\begin{aligned} a(u_n, v_n) &= a\left(\sum_{i=1}^n y_i \phi_i, \sum_{j=1}^n z_j \phi_j\right) = \sum_{i,j=1}^n y_i z_j a(\phi_i, \phi_j) \\ a(u_n, v_n) &= \left(\sum_{i=1}^n y_{A_n, i} \phi_i, \sum_{j=1}^n z_j \phi_j\right) = \sum_{i,j=1}^n y_{A_n, i} z_j (\phi_i, \phi_j). \end{aligned} \quad (2.31)$$

Zur Vereinfachung der Schreibweise führen wir die Massematrix $M_{ij} := (\phi_i, \phi_j)$ und die Steifigkeitsmatrix $S_{ij} := a(\phi_i, \phi_j)$ ein. M und S sind symmetrisch und positiv definit. Beschreibt $(.,.)$ das Standardskalarprodukt in \mathbb{R}^n , so entspricht (2.31)

$$(y, Sz) = (y_{A_n}, Mz) \quad (2.32)$$

und es folgt $y_{A_n} = M^{-1}Sy$. Für die spezielle Bilinearform mit $a^{kl} = \mathbf{1}$ führen wir die Matrix $G_{ij} := (\nabla \phi_i, \nabla \phi_j)$ ein.

Lemma 2.14 *Sind A, B positiv definite, symmetrische und reelle Matrizen, so besitzen $A^{-1}B$ und $\sqrt{A^{-1}}^\top B \sqrt{A^{-1}}$ die gleichen Eigenwerte.*

Beweis: Da A symmetrisch und reell ist, existiert eine orthogonale Matrix U und eine Diagonalmatrix D mit

$$A = UDU^\top.$$

D ist auch positiv definit. Damit kann die Quadratwurzel von A definiert werden durch $\sqrt{A} := U\sqrt{D}U^\top = \sqrt{A}^\top$.

Sei ξ der Eigenvektor zum Eigenwert λ von $A^{-1}B$. A ist regulär also auch A^{-1} und $\sqrt{A^{-1}}$. Somit existiert ein $\tilde{\xi}$ mit $\sqrt{A^{-1}}\tilde{\xi} = \xi$. Es folgt

$$\begin{aligned} A^{-1}B\xi &= \lambda\xi, \\ A^{-1}B\sqrt{A^{-1}}\tilde{\xi} &= \lambda\sqrt{A^{-1}}\tilde{\xi}. \end{aligned}$$

Nach der Multiplikation von links mit $(\sqrt{A^{-1}})^{-1}$ folgt die Behauptung. \square

Lemma 2.15 Sind $\lambda_{\min}^{a^{ij}}, \lambda_{\max}^{a^{ij}}$ der kleinste bzw. grösste Eigenwert von a^{ij} , so gilt

$$\lambda_{\min}^{a^{ij}}\lambda_{\min}^G \leq \lambda_{\min}^S \leq \lambda_{\max}^S \leq \lambda_{\max}^{a^{ij}}\lambda_{\max}^G.$$

Beweis: Es ist $(y, Sy) = \sum_{i,j=1}^n y_i y_j \underbrace{a(\phi_i, \phi_j)}_{S_{ij}}$ und

$$\begin{aligned} S_{ij} &= \int_{\Omega} \sum_{k,l} a^{kl} \frac{\partial}{\partial x_k} \phi_i(x) \frac{\partial}{\partial x_l} \phi_j(x) dx \\ &= (\nabla \phi_i, (a^{kl}) \nabla \phi_j). \end{aligned}$$

Für (y, Sy) folgt

$$\begin{aligned} (y, Sy) &= \sum_{i,j=1}^n y_i y_j (\nabla \phi_i, (a^{kl}) \nabla \phi_j) \\ &= \left(\sum_{i=1}^n y_i \nabla \phi_i, (a^{kl}) \sum_{j=1}^n y_j \nabla \phi_j \right) \\ &\leq \lambda_{\max}^{a^{kl}} (y, Gy). \end{aligned}$$

Damit folgt mit der Quotientenregel von Rayleigh $\lambda_{\max}^S \leq \lambda_{\max}^{a^{kl}} \lambda_{\max}^G$. Die Ungleichung für λ_{\min}^S folgt analog. \square

Nun soll der kleinste und der grösste Eigenwert von $M^{-1}S$ abgeschätzt werden.

Satz 2.16 Sei S_h der Raum der linearen finiten Elemente mit einer quasiuniformen Triangulation von Ω mit Maschenweite h . Für den kleinsten Eigenwert $\lambda_{\min}^{M^{-1}S}$ und den grössten $\lambda_{\max}^{M^{-1}S}$ von $M^{-1}S$ gilt

$$\lambda_{\min}^{M^{-1}S} \geq c(\Omega) \quad \text{und} \quad \lambda_{\max}^{M^{-1}S} \leq \frac{c(\kappa)}{h^2}.$$

Beweis: Betrachten wir für Funktionen aus \mathcal{U} das Eigenwertproblem

$$\begin{aligned} a(u, v) &= \lambda(u, v) && \text{in } \Omega \text{ und} \\ u &= 0 && \text{auf } \partial\Omega, \end{aligned}$$

so gilt für den kleinsten Eigenwert nach der Quotientenregel von Rayleigh

$$\lambda_1 = \min_{\substack{u \in \mathcal{U} \\ u \neq 0}} \frac{a(u, u)}{|u|^2} \leq \min_{\substack{u_n \in \mathcal{U}_n \\ u_n \neq 0}} \frac{a(u_n, u_n)}{|u_n|^2} = \min_{\substack{y \in \mathbb{R}^n \\ y \neq 0}} \frac{(y, Sy)}{(y, My)}.$$

Ersetzt man y durch $\sqrt{M^{-1}}z$, so folgt

$$\lambda_1 \leq \min_{\substack{z \in \mathbb{R}^n \\ z \neq 0}} \frac{(\sqrt{M^{-1}}z, S\sqrt{M^{-1}}z)}{(z, z)} = \lambda_{\min}^{\sqrt{M^{-1}}S\sqrt{M^{-1}}}.$$

λ_1 ist eine nur von der Menge Ω abhängende Konstante. Damit folgt mit dem Lemma 2.14 die erste Behauptung.

Die Abschätzung des grössten Eigenwertes kann wie folgt aufgeteilt werden. Es gilt

$$\begin{aligned} \lambda_{\max}^{M^{-1}S} &= \max_{\substack{y \in \mathbb{R}^n \\ y \neq 0}} \frac{(\sqrt{M^{-1}}y, S\sqrt{M^{-1}}y)}{(y, y)} = \max_{\substack{y \in \mathbb{R}^n \\ y \neq 0}} \frac{(\sqrt{M^{-1}}y, S\sqrt{M^{-1}}y)}{(y, M^{-1}y)} \frac{(y, M^{-1}y)}{(y, y)} \\ &\leq \max_{\substack{z \in \mathbb{R}^n \\ z \neq 0}} \frac{(z, Sz)}{(z, z)} \max_{\substack{y \in \mathbb{R}^n \\ y \neq 0}} \frac{(y, M^{-1}y)}{(y, y)}. \end{aligned}$$

Fried [10] hat gezeigt, dass für jede Wahl der Triangulation $\frac{\lambda_{\max}^S}{\lambda_{\min}^M} \leq \frac{c}{h^2}$ gilt. Die Konstante hängt von der Geometrie der Dreiecke ab. Sie wächst, je kleiner ein Winkel eines Dreiecks der Triangulation wird. In quasiuniformen Gittern können die kleinsten Winkel mit Hilfe von κ nach unten und damit die Konstante c nach oben beschränkt werden. \square

Wird zum Beispiel im eindimensionalen Fall das Intervall $[0, 1]$ in $n+1$ Teilintervalle der Länge $h = \frac{1}{n+1}$ aufgeteilt, so können die Eigenwerte von M und G alle exakt berechnet werden. Es gilt

$$M = \frac{h}{6} \begin{pmatrix} 4 & 1 & & 0 \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & 4 \end{pmatrix}, \quad G = \frac{1}{h} \begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix}.$$

M und G besitzen die Eigenwerte

$$\lambda_i^M = \frac{h}{3} \left(2 + \cos \frac{i\pi}{n+1} \right), \quad \lambda_i^G = \frac{1}{h} \left(2 + 2 \cos \frac{i\pi}{n+1} \right) \quad \text{für } i = 1, \dots, n.$$

Damit kann $\lambda_{\max}^{M^{-1}G}$ abgeschätzt werden durch

$$\lambda_{\max}^{M^{-1}G} \leq \frac{2 + 2 \cos \frac{\pi}{n+1}}{2 - \cos \frac{\pi}{n+1}} \frac{3}{h^2} \leq \frac{12}{h^2}.$$

2.3 Approximation der Variationsungleichung

2.3.1 Die semidiskrete Variationsungleichung

Eine mögliche Approximation der Variationsungleichung (V) erhalten wir durch das Galerkinverfahren. Dabei wird das ganze Problem in einem endlichdimensionalen Unterraum von \mathcal{U} betrachtet. Sei $\mathcal{U}_n(\Omega)$ ein n -dimensionaler Unterraum von \mathcal{U} . Eine Möglichkeit für \mathcal{U}_n ist der Raum aufgespannt durch finite Elemente. Die Variationsungleichung (V) habe ihren Ursprung in einem Hindernisproblem mit den Hindernisfunktionen $\underline{\psi}$ und $\bar{\psi}$. Zunächst werden die Funktionen $\underline{\psi}$ und $\bar{\psi}$ durch geeignete Funktionen $\underline{\psi}_n$ und $\bar{\psi}_n$ aus \mathcal{U}_n approximiert. Dann wird in \mathcal{U}_n die Menge K ersetzt durch

$$K_n := \{v_n \in L^2(0, T; \mathcal{U}_n) : \underline{\psi}_n(x) \leq v_n(x, t) \leq \bar{\psi}_n(x) \text{ f.f.a. } (x, t) \in Q\}.$$

Mit der Anfangsbedingung $u_{n,0} \in K_n$ und der Definition von $D_n = \{v_n \in L^2(0, T; \mathcal{U}_n) : \dot{v}_n \in L^2(0, T; \mathcal{U}_n), v_n(0) = u_{n,0}\}$ erhalten wir die (V) entsprechende semidiskrete Variationsungleichung.

Finde $u_n \in K_n \cap D_n$ so, dass

$$\int_0^T (\dot{u}_n, v_n - u_n) dt + \int_0^T a(u_n, v_n - u_n) dt \geq \int_0^T (f, v_n - u_n) dt \quad \forall v_n \in K_n. \quad (2.33)$$

Seien $\phi_i(x)$ für $i = 1 \dots n$ die Basisfunktionen des Unterraumes \mathcal{U}_n . v_n und u_n können als Linearkombination der Basisfunktionen dargestellt werden. Mit $v_n(x, t) = \sum_{i=1}^n z_i(t)\phi_i(x)$ und $u_n(x, t) = \sum_{i=1}^n y_i(t)\phi_i(x)$ folgt aus (2.33)

$$\begin{aligned} \int_0^T \sum_{i=1}^n \sum_{j=1}^n \dot{y}_i(t)(z_j(t) - y_j(t))(\phi_i, \phi_j) dt + \int_0^T \sum_{i=1}^n \sum_{j=1}^n y_i(t)(z_j(t) - y_j(t))a(\phi_i, \phi_j) dt \\ \geq \int_0^T \sum_{j=1}^n (z_j(t) - y_j(t))(f, \phi_j) dt. \end{aligned} \quad (2.34)$$

Mit $l_j := (f, \phi_j)$, der Massematrix M und der Steifigkeitsmatrix S lässt sich die Variationsungleichung (2.34) schreiben als

$$\int_0^T (\dot{y}, M(z - y)) dt + \int_0^T (y, S(z - y)) dt \geq \int_0^T (l, z - y) dt,$$

wobei $y(t) \in \mathbb{R}^n$. Die Elemente in $K_n \cap D_n$ sind nach Lemma 2.3 stetig bezüglich der Zeit. Sei

$$z(t) \in \tilde{K}_n := \{w \in \mathbb{R}^n : \underline{\psi}_n(x) \leq \sum_{i=1}^n w_i \phi_i(x) \leq \bar{\psi}_n(x) \text{ f.f.a. in } \Omega\}.$$

Zudem sind alle Komponenten von z in $L^2(0, T)$. Mit der Anfangsbedingung $y_0 \in \tilde{K}_n$ definieren wir

$$\begin{aligned} \bar{K}_n &:= \{w \in L^2(0, T; \mathbb{R}^n) : w(t) \in \tilde{K}_n \text{ f.f.a. } t \in [0, T]\}, \\ \bar{D}_n &:= \{w \in L^2(0, T; \mathbb{R}^n) : \dot{w} \in L^2(0, T; \mathbb{R}^n), w(0) = y_0\}. \end{aligned}$$

Damit kann die semidiskrete Variationsungleichung für die Koeffizientenvektoren formuliert werden.

Semidiskrete Variationsungleichung: Finde $y \in \overline{K}_n \cap \overline{D}_n$ mit

$$\int_0^T (M\dot{y} + Sy - l, z - y) dt \geq 0 \quad \forall z \in \overline{K}_n. \quad (\text{S})$$

Mit den gleichen Argumenten wie im Beweis von Lemma 2.10, kann man zeigen, dass die Lösung $y(t)$ für fast alle $t \in [0, T]$ auch die Variationsungleichung

$$(M\dot{y}(t) + Sy(t) - l, z - y(t)) \geq 0 \quad \forall z \in \tilde{K}_n \quad (\text{S}')$$

erfüllt.

2.3.2 Eigenschaften der semidiskreten Lösung

In diesem Abschnitt werden folgende Bezeichnungen verwendet. Es steht $y(t_1-)$ für $\lim_{\epsilon \rightarrow 0} y(t_1 - \epsilon)$ und $y(t_1+)$ für $\lim_{\epsilon \rightarrow 0} y(t_1 + \epsilon)$. $|\cdot|$ bezeichnet die Standardnorm in \mathbb{R}^n und $\|\cdot\|_M$ die Norm bezüglich dem Skalarprodukt $(\cdot, \cdot)_M := (M\cdot, \cdot)$.

In einem beliebigen Hilbertraum H ist die Projektion auf eine konvexe, abgeschlossene Teilmenge K so definiert, dass zu einem $z \in H$ das $y \in K$ gesucht wird mit $\|z - y\|_H = \inf_{v \in K} \|z - v\|_H$. In jedem Hilbertraum gilt die Parallelogrammidentität, d.h. für $a, b \in H$ gilt

$$\|a + b\|_H^2 + \|a - b\|_H^2 = 2(\|a\|_H^2 + \|b\|_H^2). \quad (2.35)$$

Sei y^k eine Minimalfolge in K , d.h. $\lim_{k \rightarrow \infty} \|y^k - z\|_H = d = \inf_{v \in K} \|z - v\|_H$. Mit (2.35) folgt

$$\|y^k - y^l\|_H^2 = 2\|z - y^k\|_H^2 + 2\|z - y^l\|_H^2 - 4\left\|z - \frac{1}{2}(y^k + y^l)\right\|_H^2.$$

Wegen der Konvexität von K liegt $\frac{1}{2}(y^k + y^l)$ in K , und es gilt

$$\|y^k - y^l\|_H^2 \leq 2\|z - y^k\|_H^2 + 2\|z - y^l\|_H^2 - 4d^2.$$

Somit ist $\lim_{k, l \rightarrow \infty} \|y^k - y^l\|_H^2 = 0$. Da H vollständig ist, existiert ein $y \in K$ mit $\lim_{k \rightarrow \infty} y^k = y$. Seien y_1, y_2 zwei Projektionen von z auf K , d.h. es gilt $\|z - y_1\|_H = \|z - y_2\|_H = \inf_{v \in K} \|z - v\|_H = d$. Durch Einsetzen von $a = z - y_1$ und $b = z - y_2$ in (2.35) folgt

$$\begin{aligned} \|y_2 - y_1\|_H^2 &= 2 \left(\|z - y_1\|_H^2 + \|z - y_2\|_H^2 - 2 \left\|z - \frac{y_1 + y_2}{2}\right\|_H^2 \right) \\ &\leq 2(\|z - y_1\|_H^2 + \|z - y_2\|_H^2 - 2d^2) = 0. \end{aligned}$$

Somit ist die Projektion auf eine konvexe, abgeschlossene Teilmenge von H eindeutig. Diese Projektion kann durch eine Variationsungleichung charakterisiert werden (vgl. [12]).

Proposition 2.17 *Sei K eine konvexe, abgeschlossene Teilmenge des Hilbertraumes H und $z \in H$. $y \in K$ ist genau dann die Projektion von z auf K , $y = \text{Pr}_K z$, wenn*

$$(y, v - y) \geq (z, v - y) \quad \forall v \in K$$

gilt.

Beweis: Seien $y = \text{Pr}_K z$ und $v \in K$. Für alle $\theta \in [0, 1]$ liegt $y + \theta(v - y)$ in K . Die Funktion

$$G(\theta) = \|z - y - \theta(v - y)\|_H^2 = \|z - y\|_H^2 - 2\theta(z - y, v - y) + \theta^2\|v - y\|_H^2$$

nimmt ihr Minimum in $\theta = 0$ an. Somit ist $G'(0) \geq 0$, d.h.

$$\begin{aligned} -(z - y, v - y) &\geq 0 & \forall v \in K, \\ (y, v - y) &\geq (z, v - y) & \forall v \in K. \end{aligned}$$

Umgekehrt gilt für alle $v \in K$

$$0 \leq (y - z, v - z + z - y) = -\|y - z\|_H^2 + (y - z, v - z).$$

Daraus folgt

$$\|y - z\|_H^2 \leq (y - z, v - z) \leq \|y - z\|_H \|v - z\|_H.$$

Nach der Division durch $\|y - z\|_H$ wird klar, dass für y der Abstand zu z am kleinsten wird. Somit ist $y = \text{Pr}_K z$. \square

Nun führen wir den Normal- und den Tangentialkegel einer konvexen, abgeschlossenen Teilmenge eines Hilbertraumes H mit Hilfe dieser Projektion ein.

Definition 6 Sei $K \subset H$ eine konvexe, abgeschlossene Menge und $y \in \partial K$. Wir definieren

$$\begin{aligned} N_K(y) &:= \text{Pr}_K^{-1}(y) - y, \\ T_K(y) &:= \{v \in H : (w, v) \leq 0, \quad \forall w \in N_K(y)\}. \end{aligned}$$

$N_K(y)$ heisst der Normalkegel der Menge K im Punkt y und $T_K(y)$ der Tangentialkegel im Punkt y .

Geometrisch kann der Tangentialkegel als Durchschnitt aller Halbräume begrenzt durch Stützhyperebenen der Menge K im Punkt $y \in \partial K$ definiert werden. Damit wird klar, dass $T_K(y)$ nicht vom gewählten Skalarprodukt abhängt. Hingegen hängt $N_K(y)$ sehr wohl vom Skalarprodukt ab. $N_K(y)$ ist die Vereinigung der Normalenvektoren aller Stützhyperebenen im Punkt y . In dieser Arbeit wird $N_K(y)$ ausschliesslich bezüglich dem Standardskalarprodukt definiert. Die Abbildung 2.2 zeigt in einem 2-dimensionalen Beispiel die beiden Kegel.

Grundsätzlich werden für die Diskussion der Variationsungleichung (S) zwei Fälle unterschieden, je nachdem ob $y(t)$ im Innern oder auf dem Rand von \tilde{K}_n liegt.

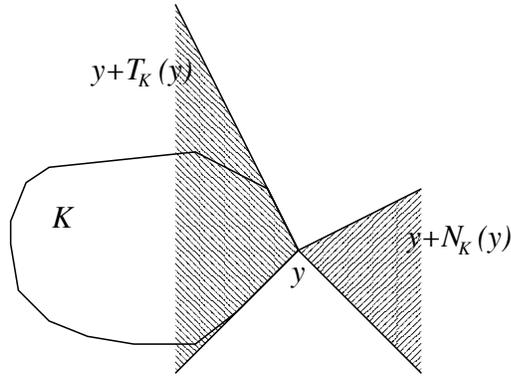
Lemma 2.18 Sei y die Lösung der semidiskreten Variationsungleichung (S). Liegt $y(t_0)$ für ein $t_0 \in (0, T)$ im Innern von \tilde{K}_n , so gilt für ein $\delta > 0$

$$M\dot{y}(t) + Sy(t) - l = 0$$

für fast alle $t \in [t_0 - \delta, t_0 + \delta]$.

Beweis: Da $y(t)$ stetig ist, existiert ein Intervall $I := [t_0 - \delta, t_0 + \delta]$ so, dass für $t \in I$ $y(t)$ immer noch im Innern von \tilde{K}_n liegt. Der minimale Abstand von $y(I)$ zum Rand von \tilde{K}_n

$$d := \min_{t \in I} \min_{w \in \partial \tilde{K}_n} |w - y(t)|$$

Abbildung 2.2: Die Mengen $N_K(y)$ und $T_K(y)$.

ist echt positiv. Sei $\mu : I \rightarrow \mathbb{R}^n$ eine beliebige, stetige Funktion nicht identisch 0. Mit

$$z(t) = \begin{cases} y(t) & \text{für } t \notin I, \\ y(t) + \frac{d}{\max_{s \in I} |\mu(s)|} \mu(t) & \text{für } t \in I, \end{cases}$$

liegt $z(t)$ für alle $t \in [0, T]$ in \tilde{K}_n . Wird dieses $z(t)$ in die semidiskrete Variationsungleichung (S) eingesetzt, so folgt

$$\int_{t_0-\delta}^{t_0+\delta} (M\dot{y} + Sy - l, \mu) dt \geq 0. \quad (2.36)$$

Wir können $\mu(t)$ ersetzen durch $-\mu(t)$ und erhalten zusammen mit (2.36)

$$\int_{t_0-\delta}^{t_0+\delta} (M\dot{y} + Sy - l, \mu) dt = 0.$$

Da μ beliebig gewählt werden konnte, folgt unmittelbar

$$M\dot{y}(t) + Sy(t) - l = 0 \quad \text{für fast alle } t \in I.$$

□

Liegt $y(t_0)$ im Innern von \tilde{K}_n , so kann mit diesem Lemma die Lösungsfunktion y als Lösung eines Systemes gewöhnlicher Differentialgleichungen eindeutig bestimmt werden, bis $y(t)$ in einem Zeitpunkt t_1 den Rand von \tilde{K}_n berührt. Liegt $\dot{y}(t_1-) = M^{-1}(l - Sy(t_1))$ nicht in $T_{\tilde{K}_n}(y(t_1))$, so muss sich die Richtung $\dot{y}(t_1+)$ unstetig ändern, da $y(t)$ in \tilde{K}_n bleiben muss. Was kann im Zeitpunkt t_1 passieren?

Lemma 2.19 *Trifft die semidiskrete Lösung $y(t)$ im Zeitpunkt t_1 auf den Rand von \tilde{K}_n und ist $\dot{y}(t_1-) \notin T_{\tilde{K}_n}(y(t_1))$, dann existiert ein $\delta > 0$ mit $y(t) \in \partial\tilde{K}_n$ für $t \in [t_1, t_1 + \delta]$.*

Beweis: Annahme: $y(t)$ liege für jedes t aus einer Umgebung von t_1 mit $t \neq t_1$ im Innern von \tilde{K}_n und $\dot{y}(t_1-) \neq \dot{y}(t_1+)$.

Nach dem Lemma 2.18 gilt

$$\begin{aligned} M\dot{y}(t_1-) + Sy(t_1) - l &= 0, \\ M\dot{y}(t_1+) + Sy(t_1) - l &= 0. \end{aligned}$$

Durch Subtraktion folgt

$$M(\dot{y}(t_1-) - \dot{y}(t_1+)) = 0.$$

Da M positiv definit ist, folgt ein Widerspruch zur Annahme. \square

Satz 2.20 (Existenz und Eindeutigkeit) Für fast alle $t \in [0, T]$ mit $y(t) \in \partial\tilde{K}_n$ und $\dot{y}(t) = M^{-1}(l - Sy(t)) \notin T_{\tilde{K}_n}(y(t))$ existiert eine eindeutige Richtung $q(t) \in T_{\tilde{K}_n}(y(t))$ mit $y(t + \Delta t) = y(t) + \int_t^{t+\Delta t} q(s)ds$ und

$$(Mq(t) + Sy(t) - l, z - y(t)) \geq 0 \quad \forall z \in \tilde{K}_n.$$

Die Richtung $q(t)$ ist die Projektion bezüglich dem Skalarprodukt (M, \cdot) von $\dot{y}(t)$ auf die Menge $T_{\tilde{K}_n}(y(t))$.

Beweis: Sei $q(t)$ die Projektion bezüglich dem Skalarprodukt (M, \cdot) von $\dot{y}(t)$ auf $T_{\tilde{K}_n}(y(t))$. Nach der Proposition 2.17 gilt für $q(t)$

$$(M(q(t) - \dot{y}(t)), w - q(t)) \geq 0 \quad \forall w \in T_{\tilde{K}_n}(y(t)). \quad (2.37)$$

$q(t)$ liegt auf dem Rand von $T_{\tilde{K}_n}(y(t))$. Die Ungleichung in (2.37) beschreibt eine Stützhyperebene zu $T_{\tilde{K}_n}(y(t))$ im Punkt $q(t)$. Da $T_{\tilde{K}_n}(y(t))$ ein Kegel mit Spitze in $y(t)$ ist, folgt $-M(q(t) - \dot{y}(t)) \in N_{\tilde{K}_n}(y(t))$ und es existiert ein $n \in N_{\tilde{K}_n}$ mit

$$M(q(t) - \dot{y}(t)) = Mq(t) + Sy(t) - l = -n. \quad (2.38)$$

Wird $w = 0 \in T_{\tilde{K}_n}(y(t))$ in (2.37) eingesetzt, so folgt, da $q(t) \in T_{\tilde{K}_n}(y(t))$ und $n \in N_{\tilde{K}_n}(y(t))$,

$$\begin{aligned} 0 &\leq (M(q(t) - \dot{y}(t)), -q(t)) = (n, q(t)) \leq 0, \\ &(n, q(t)) = 0. \end{aligned}$$

Da für $y(t) \in \partial\tilde{K}_n$ und $z \in \tilde{K}_n$ die Differenz $z - y(t)$ in $T_{\tilde{K}_n}(y(t))$ liegt, gilt nach (2.38)

$$(Mq(t) + Sy(t) - l, z - y(t)) = \underbrace{-(n, z - y(t))}_{\leq 0} \geq 0 \quad \forall z \in \tilde{K}_n.$$

Damit ist die Existenz einer Richtung $q(t) \in T_{\tilde{K}_n}(y(t))$ gezeigt.

(S') im Zeitpunkt $t + \Delta t$ mit $y(t + \Delta t) = y(t) + \int_t^{t+\Delta t} q(s)ds$ und $z = y(t)$ liefert

$$\left(Mq(t + \Delta t) + S \left(y(t) + \int_t^{t+\Delta t} q(s)ds \right) - l, - \int_t^{t+\Delta t} q(s)ds \right) \geq 0.$$

q ist integrierbar, somit folgt mit dem Differentiationssatz von Lebesgue (2.19) im Limes $\Delta t \rightarrow 0$ für fast alle $t \in [0, T]$

$$(Mq(t) + Sy(t) - l, -q(t)) \geq 0.$$

Da $-(Mq(t) + Sy(t) - l) \in N_{\tilde{K}_n}(y(t))$ und $q(t) \in T_{\tilde{K}_n}(y(t))$, folgt

$$(Mq(t) + Sy(t) - l, q(t)) = 0.$$

Seien $q_1(t)$ und $q_2(t)$ Richtungen aus $T_{\tilde{K}_n}(y(t))$ mit

$$\begin{aligned} Mq_1(t) + Sy(t) - l &= -n_1 & \text{und } (n_1, q_1(t)) &= 0, \\ Mq_2(t) + Sy(t) - l &= -n_2 & \text{und } (n_2, q_2(t)) &= 0. \end{aligned}$$

Nach Subtraktion gilt

$$M(q_1(t) - q_2(t)) = n_2 - n_1.$$

Für die M -Norm von $q_1(t) - q_2(t)$ folgt

$$\begin{aligned} 0 \leq \|q_1(t) - q_2(t)\|_M^2 &= (n_2, q_1(t) - q_2(t)) - (n_1, q_1(t) - q_2(t)) \\ &= \underbrace{(n_2, q_1(t))}_{\leq 0} + \underbrace{(n_1, q_2(t))}_{\leq 0} \leq 0. \end{aligned}$$

Damit schliessen wir $q_1(t) = q_2(t)$ für fast alle $t \in [0, T]$. □

In fast jedem Zeitpunkt $t \in [0, T]$ in dem $y(t)$ auf dem Rand $\partial\tilde{K}_n$ liegt, bestimmt der Satz 2.20 die eindeutige Richtung $\dot{y}(t)$. Daraus folgt zusammen mit dem Lemma 2.18 die Existenz und die Eindeutigkeit der Lösung der semidiskreten Variationsungleichung (S).

Korollar 2.21 $y(t)$ sei die Lösung des semidiskreten Problems (S). Ist $u_{n,0} = \Pi u_0$, so gibt es ein c unabhängig von h und u mit

$$\sup_{t \in (0, T)} \|\dot{y}(t)\|_M \leq |f| + c|u_0|_{H^2} + |Au_0|.$$

Beweis: Wie im Lemma 2.8 kann gezeigt werden, dass $\|\dot{y}(t)\|_M$ monoton fallend ist. Damit folgt für alle $t \in (0, T)$

$$\|\dot{y}(t)\|_M \leq \|M^{-1}l - M^{-1}Sy_0\|_M. \quad (2.39)$$

Für ein beliebiges $v_n(x) = \sum z_i \phi_i(x) \in S_h$ gilt

$$|(f, v_n)| \leq |f| \cdot |v_n| = |f| \cdot \|z\|_M. \quad (2.40)$$

Es ist $|(f, v_n)| = |(l, z)| = |(M^{-1}l, z)_M|$. Durch Einsetzen von $z = M^{-1}l$ erhält man zusammen mit (2.40)

$$\begin{aligned} \|M^{-1}l\|_M^2 &\leq |f| \cdot \|M^{-1}l\|_M, \\ \|M^{-1}l\|_M &\leq |f|. \end{aligned} \quad (2.41)$$

Da in \mathcal{U} $\|\cdot\|$ und die Halbnorm $|\cdot|_{\mathcal{U}}$ äquivalente Normen sind, kann (2.10) für $u, v \in \mathcal{U}$ geschrieben werden als

$$|a(u, v)| \leq c|u|_{\mathcal{U}}|v|_{\mathcal{U}}. \quad (2.42)$$

Für $\|M^{-1}Sy_0\|_M \neq 0$ folgt mit $u_{n,0}(x) = \sum_{i=1}^n y_{0i}\phi_i(x)$ und (2.42)

$$\begin{aligned} \|M^{-1}Sy_0\|_M^2 &= (Sy_0, M^{-1}Sy_0) = a \left(u_{n,0}, \sum_{i=1}^n (M^{-1}Sy_0)_i \phi_i \right) \\ &= a \left(u_{n,0} - u_0, \sum_{i=1}^n (M^{-1}Sy_0)_i \phi_i \right) + a \left(u_0, \sum_{i=1}^n (M^{-1}Sy_0)_i \phi_i \right). \end{aligned}$$

Die Normen können geschrieben werden als $|\sum_{i=1}^n (M^{-1}Sy_0)_i \phi_i|_{\mathcal{U}} = \sqrt{(GM^{-1}Sy_0, M^{-1}Sy_0)}$ und $|\sum_{i=1}^n (M^{-1}Sy_0)_i \phi_i| = \sqrt{(Sy_0, M^{-1}Sy_0)}$. Damit ergibt sich die Abschätzung

$$\|M^{-1}Sy_0\|_M^2 \leq c \|u_{n,0} - u_0\| \sqrt{(GM^{-1}Sy_0, M^{-1}Sy_0)} + |Au_0| \sqrt{(Sy_0, M^{-1}Sy_0)}.$$

Nach der Division durch $\|M^{-1}Sy_0\|_M$ gilt mit $\frac{(GM^{-1}Sy_0, M^{-1}Sy_0)}{(Sy_0, M^{-1}Sy_0)} \leq \lambda_{\max}^{\sqrt{M^{-1}G}\sqrt{M^{-1}}}$

$$\|M^{-1}Sy_0\|_M \leq c \|u_{n,0} - u_0\| \sqrt{\lambda_{\max}^{\sqrt{M^{-1}G}\sqrt{M^{-1}}} + |Au_0|}.$$

Da $u_{n,0} = \Pi u_0$, folgt mit dem Satz 2.13, den Lemmata 2.14, 2.15 und dem Satz 2.16

$$\|M^{-1}Sy_0\|_M \leq c |u_0|_{H^2} + |Au_0|. \quad (2.43)$$

Aus (2.39) folgt zusammen mit (2.41) und (2.43) die Behauptung. \square

Wir können $|\dot{u}_n|_T$ wie im Lemma 2.7 abschätzen.

Lemma 2.22 *Sei u_n die Lösung der semidiskreten Variationsungleichung (2.33) und $u_{n,0} = \Pi u_0$, so gibt es eine von h und u unabhängige Konstante c mit*

$$\begin{aligned} |\dot{u}_n|_T &\leq |f|_T + c \|u_{n,0}\| \\ &\leq |f| + c(h|u_0|_{H^2} + \|u_0\|). \end{aligned}$$

Beweis: Sei $\tau > 0$. Mit

$$v_n(x, t) = \begin{cases} u_n(x, t - \tau) & \text{für } t \in (\tau, T], \\ u_n(x, t) & \text{für } t \in [0, \tau] \end{cases}$$

eingesetzt in (2.33) folgt nach Division durch τ

$$\begin{aligned} \int_{\tau}^T \left(\dot{u}_n(t) - f, \frac{u_n(t - \tau) - u_n(t)}{\tau} \right) + a \left(u_n(t), \frac{u_n(t - \tau) - u_n(t)}{\tau} \right) dt &\geq 0, \\ \int_{\tau}^T \left(M\dot{y}(t), \frac{y(t - \tau) - y(t)}{\tau} \right) + \left(Sy(t) - l, \frac{y(t - \tau) - y(t)}{\tau} \right) dt &\geq 0. \end{aligned}$$

Die gleichen Argumente wie im Beweis vom Lemma 2.7 führen zu

$$|\dot{u}_n|_T = \sqrt{\int_0^T \|\dot{y}\|_M^2 dt} \leq \sqrt{\int_0^T \|M^{-1}l\|_M^2 dt} + \frac{1}{\sqrt{2}} \|y(0)\|_S. \quad (2.44)$$

Mit (2.41) kann $\|M^{-1}l\|_M$ durch $|f|$ abgeschätzt werden. Zudem gilt wegen der Stetigkeit der Bilinearform $a(\cdot, \cdot)$

$$\|y(0)\|_S^2 = a(u_{n,0}, u_{n,0}) \leq c\|u_{n,0}\|^2. \quad (2.45)$$

Es gilt mit dem Satz 2.13 $\|u_{n,0}\| \leq \|\Pi u_0 - u_0\| + \|u_0\| \leq ch|u_0|_{H^2} + \|u_0\|$. Zusammen mit (2.44) und (2.45) folgt die Behauptung. \square

Lemma 2.23 *Sei y die Lösung der semidiskreten Variationsungleichung (S) im Unterraum S_h . Sind $\underline{\psi}_n = \Pi \underline{\psi}$ und $\overline{\psi}_n = \Pi \overline{\psi}$ die Approximationen zu $\underline{\psi}$ und $\overline{\psi}$ in S_h und ist $u_{n,0} = \Pi u_0$, so gilt*

$$\sup_{t \in (0, T)} \|M^{-1}Sy(t)\|_M \leq 2|f| + c(|u_0|_{H^2} + |\underline{\psi}|_{H^2} + |\overline{\psi}_n|_{H^2}) + |Au_0| + |A\underline{\psi}| + |A\overline{\psi}|.$$

Beweis: Nach (2.32) ist $y_{A_n} = M^{-1}Sy$. $\underline{I}(t)$ und $\overline{I}(t)$ seien wieder die Koizidenzmengen, nun aber zu den Hindernisfunktionen $\underline{\psi}_n$ und $\overline{\psi}_n$. Analog zum Beweis von Lemma 2.9 folgt

$$|A_n u_n(t)| \leq \|A_n u_n(t)\|_{L^2(\Omega \setminus (\underline{I} \cup \overline{I}))}^2 + \|A_n u_n(t)\|_{L^2(\underline{I})}^2 + \|A_n u_n(t)\|_{L^2(\overline{I})}^2. \quad (2.46)$$

Aus (2.33) folgt für $t \in (0, T)$ und für fast alle $x \in \Omega \setminus (\underline{I}(t) \cup \overline{I}(t))$

$$\dot{u}_n(x, t) + A_n u_n(x, t) = f(x).$$

Damit erhält man mit dem Korollar 2.21

$$\begin{aligned} |A_n u_n(t)|_{L^2(\Omega \setminus (\underline{I} \cup \overline{I}))} &\leq |f| + |\dot{u}_n(t)| = |f| + \|\dot{y}(t)\|_M \\ &\leq 2|f| + c|u_0|_{H^2} + |Au_0|. \end{aligned} \quad (2.47)$$

Wie im Beweis von Korollar 2.21 folgt für die Hindernisfunktion $\underline{\psi}_n$ mit $\underline{\psi}_n(x) = \sum_{i=1}^n \underline{\Psi}_i \psi_i(x)$

$$\begin{aligned} |A_n \underline{\psi}_n|^2 &= (S\underline{\Psi}, M^{-1}S\underline{\Psi}) = a(\underline{\psi}_n, \sum_{i=1}^n (M^{-1}S\underline{\Psi})_i \phi_i) \\ &= a(\underline{\psi}_n - \underline{\psi}, \sum_{i=1}^n (M^{-1}S\underline{\Psi})_i \phi_i) + a(\underline{\psi}, \sum_{i=1}^n (M^{-1}S\underline{\Psi})_i \phi_i) \\ &\leq c\|\underline{\psi}_n - \underline{\psi}\| \sqrt{(GM^{-1}S\underline{\Psi}, M^{-1}S\underline{\Psi})} + |A\underline{\psi}|(S\underline{\Psi}, M^{-1}S\underline{\Psi}). \end{aligned}$$

Damit erhält man nach der Division durch $|A_n \underline{\psi}_n|$ mit dem Satz 2.13, den Lemmata 2.14, 2.15 und dem Satz 2.16

$$|A_n \underline{\psi}_n| \leq c|\underline{\psi}|_{H^2} + |A\underline{\psi}|. \quad (2.48)$$

$\|M^{-1}Sy(t)\|_M = |A_n u_n(t)|$ und (2.46) bestätigen zusammen mit (2.47) und (2.48) die Behauptung. \square

Wenn y den Rand von \tilde{K}_n trifft, ändert sich \dot{y} unstetig. Im nächsten Korollar wird der Fall untersucht, wo y den Rand wieder verlässt.

Korollar 2.24 *y sei die Lösung der semidiskreten Variationsungleichung (S), welche im Zeitpunkt $t_2 \in (0, T)$ den Rand $\partial\tilde{K}_n$ verlasse. Dann ist \dot{y} in t_2 stetig, d.h.*

$$\dot{y}(t_2-) = \dot{y}(t_2+).$$

Beweis: $y(t)$ liegt für $t \in (t_2, t_2 + \epsilon)$ im Innern von \tilde{K}_n . Daher gilt nach Lemma 2.18

$$\dot{y}(t_2+) = M^{-1}(l - Sy(t_2)) = \dot{\tilde{y}}(t_2). \quad (2.49)$$

$\dot{y}(t_2-)$ ist nach Satz 2.20 die Projektion von $\dot{\tilde{y}}(t_2)$ auf die Menge $T_{\tilde{K}_n}(y(t_2))$. Da $\dot{\tilde{y}}(t_2)$ wegen (2.49) bereits in $T_{\tilde{K}_n}(y(t_2))$ liegt, folgt

$$\dot{y}(t_2-) = \dot{y}(t_2+).$$

□

Ein Unterschied zwischen parabolischen Variationsgleichungen und Variationsungleichungen besteht darin, dass bei Variationsungleichungen die stationäre Lösung in endlicher Zeit erreicht werden kann. Die stationäre Lösung y^* erfüllt die Variationsungleichung

$$(Sy^* - l, z - y^*) \geq 0 \quad \forall z \in \tilde{K}_n.$$

Damit y^* in endlicher Zeit erreicht werden kann, darf y^* nicht im Innern von \tilde{K}_n liegen, d.h. $S^{-1}l \notin \tilde{K}_n$.

Korollar 2.25 *Trifft die Lösung y der semidiskreten Variationsungleichung (S) im Zeitpunkt $t_1 \in (0, T]$ auf den Rand von \tilde{K}_n und gilt $l - Sy(t_1) \in N_{\tilde{K}_n}(y(t_1))$, so ist $y(t_1)$ die stationäre Lösung von (S), d.h. $\dot{y}(t_1) = 0$.*

Beweis: Nach (2.37) gilt mit $\dot{\tilde{y}} = M^{-1}(l - Sy(t_1))$

$$(M\dot{y}, w - \dot{y}) \geq (M\dot{\tilde{y}}, w - \dot{y}) \quad \forall w \in T_{\tilde{K}_n}(y(t_1)).$$

Durch Einsetzen von $w = 0$ folgt, da $M\dot{\tilde{y}} = l - Sy(t_1) \in N_{\tilde{K}_n}(y(t_1))$ und $\dot{y} \in T_{\tilde{K}_n}(y(t_1))$, dass

$$\begin{aligned} -(M\dot{y}, \dot{y}) &\geq -(M\dot{\tilde{y}}, \dot{y}), \\ 0 \leq \|\dot{y}\|_M^2 &\leq (M\dot{\tilde{y}}, \dot{y}) \leq 0. \end{aligned}$$

□

Eigenschaften von \dot{y} , wenn \tilde{K}_n ein Polytop ist

Definition 7 *Ein Polytop ist eine konvexe abgeschlossene Teilmenge von \mathbb{R}^n , die als Durchschnitt endlich vieler Halbräume geschrieben werden kann. Ist $B \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$, so ist ein Polytop gegeben durch*

$$P = \{y \in \mathbb{R}^n : By \leq b\}.$$

Die Zeilenvektoren von B sind gerade die äusseren Normalenvektoren zu den Hyperebenen, welche die Halbräume begrenzen. Ist das Polytop P beschränkt und nicht leer, so ist der Rang der Matrix B gleich n .

Für die weiteren Untersuchungen spielt der Rand von \tilde{K}_n eine wichtige Rolle. Daher müssen die Randpunkte genau beschrieben werden. Liegt $y(t_1)$ auf einer $(n-k)$ -dimensionalen Randfläche von \tilde{K}_n , so existiert eine $(k \times n)$ -Untermatrix $B_{\mathcal{F}}$ von B mit dem Rang k , $B_{\mathcal{F}}y(t_1) = b_{\mathcal{F}}$ und $B_{\mathcal{F}^c}y(t_1) < b_{\mathcal{F}^c}$. $\mathcal{F} \subset \{1, \dots, 2n\}$ ist die Indexmenge aller Zeilen von B , die in $B_{\mathcal{F}}$ übernommen werden. $\mathcal{F}^c := \{1, \dots, 2n\} \setminus \mathcal{F}$ bezeichne das Komplement von \mathcal{F} . $b_{\mathcal{F}}$ und $b_{\mathcal{F}^c}$ erhalten wir analog aus b . Ein Vektor aus dem Normalkegel kann als positive Linearkombination der äusseren Normalenvektoren beschrieben werden. Daher gilt

$$\begin{aligned} N_{\tilde{K}_n} &= \{v \in \mathbb{R}^n : B_{\mathcal{F}}^T \xi_{\mathcal{F}} = v, \xi_{\mathcal{F}} \in \mathbb{R}_+^k\}, \\ T_{\tilde{K}_n} &= \{v \in \mathbb{R}^n : B_{\mathcal{F}}v \leq 0\}. \end{aligned}$$

Satz 2.26 *Sei y die Lösung von (S) und \tilde{K}_n sei ein Polytop $P = \{y \in \mathbb{R}^n : By \leq b\}$. Für ein $t \in [0, T]$ liege $y(t)$ auf einer $(n-k)$ -dimensionalen Randfläche von P . Sei \mathcal{F} die Indexmenge mit $\text{Rang}(B_{\mathcal{F}}) = k$, $B_{\mathcal{F}}y(t) = b_{\mathcal{F}}$ und $B_{\mathcal{F}^c}y(t) < b_{\mathcal{F}^c}$. Ist $\dot{y}(t) \notin T_{\tilde{K}_n}(y(t))$, so existiert eine eindeutige Teilmenge $\mathcal{F}' \subset \mathcal{F}$ mit*

$$\dot{y}(t) = \check{y}(t) + M^{-1}B_{\mathcal{F}'}^T \xi_{\mathcal{F}'}$$

wobei $-(B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^T)^{-1}B_{\mathcal{F}'}\check{y}(t) = \xi_{\mathcal{F}'}$.

Beweis: Nach Satz 2.20 ist $\dot{y}(t)$ die Projektion von $\check{y}(t)$ auf $T_{\tilde{K}_n}(y(t))$ bezüglich dem Skalarprodukt (M, \cdot) . Da $\check{y}(t) \notin T_{\tilde{K}_n}(y(t))$, muss $\dot{y}(t)$ auf dem Rand von $T_{\tilde{K}_n}(y(t))$ liegen, d.h. es existiert ein $\mathcal{F}' \subset \mathcal{F}$ mit

$$B_{\mathcal{F}'}\dot{y}(t) = 0, \quad B_{\mathcal{F} \setminus \mathcal{F}'}\dot{y}(t) < 0.$$

Diese Indexmenge \mathcal{F}' ist wegen der Eindeutigkeit der Projektion eindeutig. Nach (2.38) gilt

$$M(\dot{y}(t) - \check{y}(t)) = B_{\mathcal{F}'}^T \xi_{\mathcal{F}'} \quad \text{mit } \xi_{\mathcal{F}'} < 0. \quad (2.50)$$

Mit Hilfe von $B_{\mathcal{F}'}\dot{y}(t) = 0$ wird (2.50) nach $\xi_{\mathcal{F}'}$ aufgelöst,

$$\begin{aligned} \dot{y}(t) - \check{y}(t) &= M^{-1}B_{\mathcal{F}'}\xi_{\mathcal{F}'}, \\ -B_{\mathcal{F}'}\check{y}(t) &= B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^T \xi_{\mathcal{F}'}. \end{aligned} \quad (2.51)$$

$B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^T$ ist invertierbar, da der Kern nur aus dem Nullvektor besteht. Es sei $\zeta \in \mathbb{R}^k$ mit $\zeta^T B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^T \zeta = 0$. $B_{\mathcal{F}'}^T \zeta$ liegt somit im Kern von M^{-1} . Da M regulär ist, gilt $B_{\mathcal{F}'}^T \zeta = 0$. Aus $\text{Rang}(B_{\mathcal{F}'}) = k$ folgt $\zeta = 0$.

Auflösen von (2.51) führt zur gewünschten Gleichung für $\xi_{\mathcal{F}'}$,

$$-(B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^T)^{-1}B_{\mathcal{F}'}\check{y}(t) = \xi_{\mathcal{F}'}$$

Durch Einsetzen in (2.50) ergibt sich für $\dot{y}(t)$

$$\begin{aligned} M(\dot{y}(t) - \check{y}(t)) &= -B_{\mathcal{F}'}^T (B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^T)^{-1} B_{\mathcal{F}'}\check{y}(t), \\ \dot{y}(t) &= \check{y}(t) - M^{-1}B_{\mathcal{F}'}^T (B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^T)^{-1} B_{\mathcal{F}'}\check{y}(t). \end{aligned}$$

□

Korollar 2.27 *Ist $y(t)$ die Lösung des semidiskreten Problems (S), so gilt für $t \in [0, T]$*

$$\|\dot{y}(t)\|_M^2 = \|\dot{\tilde{y}}(t)\|_M^2 - (B_{\mathcal{F}'}\dot{\tilde{y}}(t), (B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^\top)^{-1}B_{\mathcal{F}'}\dot{\tilde{y}}(t)).$$

Beweis: Nach Satz 2.26 gilt

$$\dot{y}(t) = \dot{\tilde{y}}(t) - M^{-1}B_{\mathcal{F}'}^\top(B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^\top)^{-1}B_{\mathcal{F}'}\dot{\tilde{y}}(t).$$

Für die M -Norm von $\dot{y}(t)$ folgt

$$\begin{aligned} \|\dot{y}(t)\|_M^2 &= (\dot{\tilde{y}}(t), M\dot{\tilde{y}}(t)) - 2(\dot{\tilde{y}}(t), B_{\mathcal{F}'}^\top(B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^\top)^{-1}B_{\mathcal{F}'}\dot{\tilde{y}}(t)) \\ &\quad + (M^{-1}B_{\mathcal{F}'}^\top(B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^\top)^{-1}B_{\mathcal{F}'}\dot{\tilde{y}}(t), B_{\mathcal{F}'}^\top(B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^\top)^{-1}B_{\mathcal{F}'}\dot{\tilde{y}}(t)) \\ &= \|\dot{\tilde{y}}(t)\|_M^2 - 2(B_{\mathcal{F}'}\dot{\tilde{y}}(t), (B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^\top)^{-1}B_{\mathcal{F}'}\dot{\tilde{y}}(t)) \\ &\quad + (B_{\mathcal{F}'}\dot{\tilde{y}}(t), (B_{\mathcal{F}'}M^{-1}B_{\mathcal{F}'}^\top)^{-1}B_{\mathcal{F}'}\dot{\tilde{y}}(t)). \end{aligned}$$

□

2.3.3 Diskrete Variationsungleichung

Damit die semidiskrete Variationsungleichung (S) durch eine Quadraturformel approximiert werden kann, muss auch die Zeitachse diskretisiert werden. Durch die Unterteilung des Zeitintervalls $[0, T]$ in Abschnitte der Länge τ kann die Ableitung durch einen Differenzenquotienten approximiert werden. y^k sei die Lösung der folgenden Variationsungleichung und approximiere $y(t)$ im Zeitpunkt $t = k\tau$.

Diskrete Variationsungleichung: Finde $y^k \in \tilde{K}_n$ für $k = 1 \dots \frac{T}{\tau}$ mit $y^0 = y_0$ und

$$\left(\frac{1}{\tau}M(y^k - y^{k-1}) + Sy^k - l, z - y^k \right) \geq 0 \quad \forall z \in \tilde{K}_n. \quad (\text{D})$$

Die Funktion $y(t)$ kann durch die lineare Interpolation der Punkte y^k approximiert werden.

2.3.4 Eigenschaften der diskreten Lösung

Für jedes feste k kann die diskrete Variationsungleichung (D) als Variationsungleichung in \mathbb{R}^n aufgefasst werden. Kinderlehrer und Stampacchia haben in [12] die Existenz und Eindeutigkeit der Lösung solcher Variationsungleichungen untersucht.

Es sei K eine konvexe, abgeschlossene und beschränkte Teilmenge von \mathbb{R}^n und $F : K \rightarrow \mathbb{R}^n$ eine stetige Abbildung. Auch in diesem Abschnitt bezeichne (\cdot, \cdot) das Standardskalarprodukt und $|\cdot|$ die entsprechende Norm in \mathbb{R}^n .

Definition 8 *Die stetige Abbildung $F : K \rightarrow \mathbb{R}^n$ heisst stark monoton, falls ein $\alpha > 0$ existiert mit*

$$(F(y) - F(z), y - z) \geq \alpha|y - z|^2 \quad \forall y, z \in K.$$

Lemma 2.28 $K \subset \mathbb{R}^n$ sei konvex und beschränkt. Ist $F : K \rightarrow \mathbb{R}^n$ stark monoton, so besitzt die Variationsungleichung

$$(F(y), z - y) \geq 0 \quad \forall z \in K \quad (2.52)$$

genau eine Lösung y in K .

Beweis: Die Abbildung $\text{Pr}_K \circ (\mathbb{I} - F) : K \rightarrow K$ ist stetig. Nach dem Browserschen Fixpunktsatz besitzt sie einen Fixpunkt y , d.h.

$$\text{Pr}_K(y - F(y)) = y.$$

Die Projektion wird nach der Proposition 2.17 charakterisiert durch $(y, z - y) \geq (y - F(y), z - y) \quad \forall z \in K$. Somit ist der Fixpunkt y eine Lösung der Variationsungleichung (2.52).

Angenommen y_1 und y_2 seien Lösungen von (2.52), so gilt

$$(F(y_1), z - y_1) \geq 0, \quad \forall z \in K, \quad (2.53)$$

$$(F(y_2), z - y_2) \geq 0, \quad \forall z \in K. \quad (2.54)$$

Einsetzen von $z = y_2$ in (2.53) und $z = y_1$ in (2.54) führt nach der Addition der beiden Ungleichungen zu

$$\begin{aligned} (F(y_1) - F(y_2), y_2 - y_1) &\geq 0, \\ 0 &\geq (F(y_1) - F(y_2), y_1 - y_2) \geq \alpha |y_1 - y_2|^2 \geq 0. \end{aligned}$$

Somit ist die Lösung auch eindeutig. \square

Die Abbildung $F_k : \tilde{K}_n \rightarrow \mathbb{R}^n$ definiert durch

$$F_k(y) := \left(\frac{\mathbb{I}}{\tau} + M^{-1}S \right) y - \frac{1}{\tau} y^{k-1} - M^{-1}l$$

ist stetig. Damit lässt sich die diskrete Variationsungleichung (D) mit dem Skalarprodukt $(\cdot, \cdot)_M$ in der Form (2.52) schreiben. Die Existenz und Eindeutigkeit einer Lösung von (D) folgen nun direkt aus dem Lemma 2.28.

Korollar 2.29 Für $\tau > 0$ und festes k besitzt die diskrete Variationsungleichung (D) eine eindeutige Lösung.

Beweis: Mit der Extremaleigenschaft des Rayleigh-Quotienten folgt

$$\lambda_{\min}^{\frac{1}{\tau} + M^{-1}S} = \min_{|z|=1} \left(z, \left(\frac{\mathbb{I}}{\tau} + M^{-1}S \right) z \right) \geq \frac{1}{\tau} + \min_{|z|=1} (z, M^{-1}S z) = \frac{1}{\tau} + \lambda_{\min}^{M^{-1}S}. \quad (2.55)$$

Für die Abbildung F_k gilt mit (2.55)

$$(F_k(y) - F_k(z), y - z)_M = \left(\left(\frac{\mathbb{I}}{\tau} + M^{-1}S \right) (y - z), y - z \right)_M \geq \underbrace{\left(\frac{1}{\tau} + \lambda_{\min}^{M^{-1}S} \right)}_{=: \alpha} \|y - z\|_M^2.$$

Somit ist die Abbildung F_k stark monoton und das Lemma 2.28 kann angewendet werden. \square

Die Lösung y^k lässt sich auch als Projektion auffassen.

Lemma 2.30 Die Lösung y^k der diskreten Variationsungleichung (D) ist die Projektion von

$$\left(\frac{1}{\tau}M + S\right)^{-1} \left(l + \frac{1}{\tau}My^{k-1}\right)$$

auf \tilde{K}_n bezüglich dem Skalarprodukt $((\frac{1}{\tau}M + S) \cdot, \cdot)$.

Beweis: Nach der Proposition 2.17 lässt sich die Projektion als Variationsungleichung schreiben,

$$\left(\left(\frac{1}{\tau}M + S\right)y^k, z - y^k\right) \geq \left(l + \frac{1}{\tau}My^{k-1}, z - y^k\right) \quad \forall z \in \tilde{K}_n.$$

Damit ist y^k die Lösung von (D). □

Lemma 2.31 Seien $\{y^k\}_{k=1}^{T/\tau}$ die Lösungen der diskreten Variationsungleichung (D). Sind $u_{n,0} = \Pi u_0$, $\underline{\psi}_n = \Pi \underline{\psi}$ und $\bar{\psi}_n = \Pi \bar{\psi}$, so gilt für alle k und für eine von h und τ unabhängige Konstante c

$$\|M^{-1}Sy^k\|_M \leq 2|f| + c(|u_0|_{H^2} + |\underline{\psi}|_{H^2} + |\bar{\psi}|_{H^2}) + |Au_0| + |A\underline{\psi}| + |A\bar{\psi}|.$$

Beweis: Zunächst wird gezeigt, dass $\left\|\frac{y^k - y^{k-1}}{\tau}\right\|_M$ monoton fallend in k ist. Testen wir (D) für y^k mit $z = y^{k-1}$, (D) für y^{k-1} mit $z = y^k$ und addieren die beiden Ungleichungen, so folgt

$$\begin{aligned} & \left(\frac{y^k - y^{k-1}}{\tau} - \frac{y^{k-1} - y^{k-2}}{\tau} + M^{-1}S(y^k - y^{k-1}), y^{k-1} - y^k\right)_M \geq 0, \\ & \left(\frac{y^k - y^{k-1}}{\tau} - \frac{y^{k-1} - y^{k-2}}{\tau}, y^k - y^{k-1}\right)_M + \underbrace{\left\|y^k - y^{k-1}\right\|_S^2}_{\geq 0} \leq 0. \end{aligned} \quad (2.56)$$

Mit der Gleichheit

$$\begin{aligned} & \left(\frac{y^k - y^{k-1}}{\tau} - \frac{y^{k-1} - y^{k-2}}{\tau}, y^k - y^{k-1} - \frac{1}{2}(y^{k-1} - y^{k-2}) + \frac{1}{2}(y^{k-1} - y^{k-2})\right)_M \\ & = \frac{1}{2}\tau \left(\left\|\frac{y^k - y^{k-1}}{\tau}\right\|_M^2 - \left\|\frac{y^{k-1} - y^{k-2}}{\tau}\right\|_M^2\right) + \frac{1}{2}\tau \left\|\frac{y^k - y^{k-1}}{\tau} - \frac{y^{k-1} - y^{k-2}}{\tau}\right\|_M^2 \end{aligned}$$

folgt aus (2.56), dass für alle k

$$\left\|\frac{y^k - y^{k-1}}{\tau}\right\|_M \leq \left\|\frac{y^{k-1} - y^{k-2}}{\tau}\right\|_M. \quad (2.57)$$

Wird die semidiskrete Variationsungleichung (S') für y_0 mit $z = y^1$ getestet, ergibt sich nach der Division durch τ

$$\left(\dot{y}(0) + M^{-1}Sy_0 - M^{-1}l, \frac{y^1 - y_0}{\tau}\right)_M \geq 0.$$

Wird dazu die diskrete Variationsungleichung (D) mit $k = 1$ und $z = y_0$ addiert, so folgt

$$\begin{aligned} \left(\dot{y}(0) - \frac{y^1 - y_0}{\tau} + M^{-1}S(y_0 - y^1), \frac{y^1 - y_0}{\tau} \right)_M &\geq 0, \\ \left(\dot{y}(0), \frac{y^1 - y_0}{\tau} \right)_M - \left\| \frac{y^1 - y_0}{\tau} \right\|_M^2 - \tau \lambda_{\min}^{M^{-1}S} \left\| \frac{y^1 - y_0}{\tau} \right\|_M^2 &\geq 0. \end{aligned}$$

Mit der Schwarzischen Ungleichung erhält man

$$\|\dot{y}(0)\|_M \geq \frac{1}{1 + \tau \lambda_{\min}^{M^{-1}S}} \|\dot{y}(0)\|_M \geq \left\| \frac{y^1 - y_0}{\tau} \right\|_M. \quad (2.58)$$

Aus (2.57) folgt zusammen mit (2.58) und dem Korollar 2.21

$$\left\| \frac{y^k - y^{k-1}}{\tau} \right\|_M \leq 2|f| + c|u_0|_{H^2} + |Au_0|. \quad (2.59)$$

Mit den gleichen Argumenten wie im Beweis vom Lemma 2.23 folgt zusammen mit (2.59) die Behauptung. \square

2.4 Fehlerabschätzung

2.4.1 Fehler zwischen den Lösungen der starken Variationsungleichung und der semidiskreten Variationsungleichung

Zunächst zeigen wir eine Abschätzung, die den Fehler zwischen den Lösungen der starken Variationsungleichung und der semidiskreten Variationsungleichung auf die Approximationseigenschaft des Raumes \mathcal{U} durch den Unterraum \mathcal{U}_n zurückführt.

Lemma 2.32 *Seien u und u_n die Lösungen der starken Variationsungleichung (V) bzw. der semidiskreten Variationsungleichung (2.33). Gilt $(\dot{u} + Au - f) \in \mathcal{H}$, so existiert eine Konstante C unabhängig vom Unterraum \mathcal{U}_n und der Menge K_n mit*

$$\begin{aligned} \int_0^T \|u - u_n\|^2 dt + |u_n(T) - u(T)|^2 &\leq C \{ |u_{n,0} - u_0|^2 \\ &+ \int_0^T \inf_{v_n \in K_n} ((|\dot{u} + Au - f| + |\dot{u}_n - \dot{u}|) |v_n - u| + \|v_n - u\|^2) \\ &+ \inf_{v \in K} (|\dot{u} + Au - f| \cdot |v - u_n|) dt \}. \end{aligned}$$

Beweis: Da $a(\cdot, \cdot) : \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$ eine koerzitive Bilinearform (2.11) ist, gilt

$$\alpha \|u - u_n\|^2 \leq a(u - u_n, u - u_n) = a(u, u) - a(u, u_n) - a(u_n, u) + a(u_n, u_n). \quad (2.60)$$

Für u und u_n folgt aus (V) und (2.33)

$$\begin{aligned} \int_0^T a(u, u) dt &\leq \int_0^T (\dot{u} - f, v - u) + a(u, v) dt \quad \forall v \in K, \\ \int_0^T a(u_n, u_n) dt &\leq \int_0^T (\dot{u}_n - f, v_n - u_n) + a(u_n, v_n) dt \quad \forall v_n \in K_n. \end{aligned}$$

Durch Einsetzen in (2.60) erhält man für alle $v \in K$ und $v_n \in K_n$

$$\begin{aligned} \int_0^T \alpha \|u - u_n\|^2 dt &\leq \int_0^T (\dot{u} - f, v - u) + a(u, v) - a(u, u_n) - a(u_n, u) \\ &\quad + (\dot{u}_n - f, v_n - u_n) + a(u_n, v_n) dt \\ &= \int_0^T (\dot{u} - f, v - u) + a(u, v - u_n) + (\dot{u} - f, v_n - u) + (\dot{u}_n - \dot{u}, v_n - u_n) \\ &\quad + a(u, v_n - u) + a(u_n - u, v_n - u) dt \\ &= \int_0^T (\dot{u} - f, v - u_n) + a(u, v - u_n) + (\dot{u} - f, v_n - u) + (\dot{u}_n - \dot{u}, v_n - u_n) \\ &\quad + a(u, v_n - u) + a(u_n - u, v_n - u) dt = A \end{aligned}$$

Nach dem Satz 2.4 gilt $u \in L^2(0, T; H^2(\Omega))$. $a(u, \cdot)$ kann durch (Au, \cdot) ersetzt werden. Mit $\int_0^T (\dot{u}_n - \dot{u}, u_n - u) dt = \frac{1}{2}(|u_n(T) - u(T)|^2 - |u_{n,0} - u_0|^2)$ folgt weiter

$$\begin{aligned} A &= \int_0^T (\dot{u} + Au - f, v - u_n) + (\dot{u} + Au - f, v_n - u) \\ &\quad + (\dot{u}_n - \dot{u}, v_n - u) + (\dot{u}_n - \dot{u}, u - u_n) + a(u_n - u, v_n - u) dt \\ &\leq \int_0^T \left(|\dot{u} + Au - f| (|v - u_n| + |v_n - u|) + |\dot{u}_n - \dot{u}| \cdot |v_n - u| \right. \\ &\quad \left. + c \|u_n - u\| \cdot \|v_n - u\| \right) dt - \frac{1}{2} (|u_n(T) - u(T)|^2 - |u_{n,0} - u_0|^2) \\ &\leq \int_0^T \left(|\dot{u} + Au - f| (|v - u_n| + |v_n - u|) + |\dot{u}_n - \dot{u}| \cdot |v_n - u| + \frac{1}{2} \alpha \|u_n - u\|^2 \right. \\ &\quad \left. + \frac{c^2}{2\alpha} \|v_n - u\|^2 \right) dt - \frac{1}{2} |u_n(T) - u(T)|^2 + \frac{1}{2} |u_{n,0} - u_0|^2. \end{aligned}$$

Weitere Umformungen führen zu

$$\begin{aligned} \frac{1}{2} \int_0^T \alpha \|u - u_n\|^2 dt + \frac{1}{2} |u_n(T) - u(T)|^2 \\ \leq \frac{1}{2} |u_{n,0} - u_0|^2 + \int_0^T |\dot{u} + Au - f| (|v - u_n| + |v_n - u|) \\ + |\dot{u}_n - \dot{u}| \cdot |v_n - u| + \frac{c^2}{2\alpha} \|v_n - u\|^2 dt. \end{aligned}$$

Falls $\alpha \leq 1$ folgt

$$\begin{aligned}
& \int_0^T \|u - u_n\|^2 dt + |u_n(T) - u(T)|^2 \\
& \leq \frac{1}{\alpha} \left(|u_{n,0} - u_0|^2 + 2 \int_0^T |\dot{u} + Au - f| (|v - u_n| + |v - u_n|) \right. \\
& \quad \left. + |\dot{u}_n - \dot{u}| \cdot |v_n - u| + \frac{c^2}{\alpha} \|v_n - u\|^2 dt \right) \\
& \leq \frac{1}{\alpha} \left(2 + \frac{c^2}{\alpha} \right) \left(|u_{n,0} - u_0|^2 + \int_0^T |\dot{u} + Au - f| (|v - u_n| + |v_n - u|) \right. \\
& \quad \left. + |\dot{u}_n - \dot{u}| \cdot |v_n - u| + \|v_n - u\|^2 dt \right). \tag{2.61}
\end{aligned}$$

Für $\alpha \geq 1$ gilt

$$\begin{aligned}
& \int_0^T \|u - u_n\|^2 dt + |u_n(T) - u(T)|^2 \\
& \leq (2 + c^2) \left(|u_{n,0} - u_0|^2 + \int_0^T |\dot{u} + Au - f| (|v - u_n| + |v_n - u|) \right. \\
& \quad \left. + |\dot{u}_n - \dot{u}| \cdot |v_n - u| + \|v_n - u\|^2 dt \right). \tag{2.62}
\end{aligned}$$

Da die Ungleichungen (2.61) und (2.62) für alle $v \in K$ und $v_n \in K_n$ gelten, folgt die Behauptung. \square

Die Voraussetzung $(\dot{u} + Au - f) \in \mathcal{H}$ ist für die Lösung u von (V) erfüllt. Dies folgt aus der Regularität der Lösung u (Satz 2.4, (2.12)) und der Voraussetzung $f \in \mathcal{H}$.

Mit dem Korollar 2.6 kann $|\dot{u} - Au - f|_T$ durch die Problemdata abgeschätzt werden. $|\dot{u}|_T$ und $|\dot{u}_n|_T$ schätzen wir mit dem Lemma 2.7 bzw. dem Korollar 2.22 ab. Das Lemma 2.32 kann auf (V) angewendet werden. Dadurch reduziert sich die Fehlerabschätzung auf die Approximationseigenschaft des Raumes \mathcal{U} durch den Raum \mathcal{U}_n .

Wird als n -dimensionaler Unterraum \mathcal{U}_n der Raum der finiten Elemente S_h mit stückweise linearen Funktionen verwendet, so ist die Approximationseigenschaft des Raumes \mathcal{U} durch \mathcal{U}_n bekannt. Approximiert u_n die Funktion u aus \mathcal{U} , so muss diese jedoch nicht in K liegen.

$\Pi \underline{\psi}$ und $\Pi \bar{\psi}$ liegen in S_h und approximieren $\underline{\psi}$ und $\bar{\psi}$. In den Eckpunkten x^i der Triangulation gilt $\Pi \underline{\psi}(x^i) = \underline{\psi}(x^i)$ und $\Pi \bar{\psi}(x^i) = \bar{\psi}(x^i)$ für $i = 1, \dots, n$. Da $\Pi \underline{\psi}$, $\Pi \bar{\psi}$ und v_n zwischen den Eckpunkten linear sind, liegt $v_n(x, t)$ genau dann für alle $x \in \Omega$ zwischen $\Pi \underline{\psi}(x)$ und $\Pi \bar{\psi}(x)$, wenn

$$\underline{\psi}(x^i) \leq v_n(x^i, t) \leq \bar{\psi}(x^i) \quad \text{für } i = 1, \dots, n.$$

Damit kann K_n geschrieben werden als

$$K_n = \{v_n \in L^2(0, T; S_h) : \underline{\psi}(x^i) \leq v_n(x^i, t) \leq \bar{\psi}(x^i), \quad i = 1, \dots, n, \text{ f.f.a. } t \in [0, T]\}.$$

Mit $u \in K$ folgt $\underline{\psi}(x^i) \leq u(x^i, t) \leq \overline{\psi}(x^i)$ für $i = 1, \dots, n$ und somit liegt Πu in K_n . Es gilt nun mit dem Satz 2.13

$$\begin{aligned} \inf_{v_n \in K_n} & \left((|\dot{u} + Au - f| + |\dot{u}_n - \dot{u}|) |v_n - u| + \|v_n - u\|^2 \right) \\ & \leq (|\dot{u} + Au - f| + |\dot{u}_n - \dot{u}|) |\Pi u - u| + \|\Pi u - u\|^2 \\ & c \left(\leq (|\dot{u} + Au - f| + |\dot{u}_n - \dot{u}|) h^2 |u|_{H^2} + h^2 |u|_{H^2}^2 \right). \end{aligned} \quad (2.63)$$

Um den Term $\inf_{v \in K} |v - u_n|$ abzuschätzen, folgen wir der Argumentation von Ciarlet [6]. Wir benötigen folgende Hilfsfunktion. Sei

$$\tilde{u}_n := u_n + (\underline{\psi} - u_n)^+ - (\overline{\psi} - u_n)^-.$$

Damit gilt $\underline{\psi}(x) \leq \tilde{u}_n(x, t) \leq \overline{\psi}(x)$ f.f.a. $(x, t) \in Q$. Mit $u_n \in L^2(0, T; \mathcal{U})$ liegt auch \tilde{u}_n in $L^2(0, T; \mathcal{U})$. Somit ist $\tilde{u}_n \in K$. Seien

$$\begin{aligned} \underline{\Lambda}(t) & := \{x \in \Omega : u_n(x, t) \leq \underline{\psi}(x)\}, \\ \overline{\Lambda}(t) & := \{x \in \Omega : u_n(x, t) \geq \overline{\psi}(x)\}. \end{aligned}$$

Die Lösung u_n von (S) liegt in K_n und daher gilt

$$\Pi \underline{\psi}(x) \leq u_n(x, t) \leq \Pi \overline{\psi}(x) \quad \text{f.f.a. } (x, t) \in Q.$$

Damit ergibt sich

$$0 \leq \underline{\psi}(x) - u_n(x) \leq \underline{\psi}(x) - \Pi \underline{\psi}(x) \quad \forall x \in \underline{\Lambda}, \quad (2.64)$$

$$0 \leq u_n(x) - \overline{\psi}(x) \leq \Pi \overline{\psi}(x) - \overline{\psi}(x) \quad \forall x \in \overline{\Lambda}. \quad (2.65)$$

Da $u_n - \tilde{u}_n = 0$ auf $\Omega \setminus (\overline{\Lambda} \cup \underline{\Lambda})$, folgt mit (2.64) und (2.65)

$$\begin{aligned} |u_n - \tilde{u}_n|^2 & = \int_{\underline{\Lambda}} (u_n(x) - \underline{\psi}(x))^2 dx + \int_{\overline{\Lambda}} (u_n(x) - \overline{\psi}(x))^2 dx \\ & \leq \int_{\underline{\Lambda}} (\underline{\psi}(x) - \Pi \underline{\psi}(x))^2 dx + \int_{\overline{\Lambda}} (\Pi \overline{\psi}(x) - \overline{\psi}(x))^2 dx \\ & \leq |\underline{\psi} - \Pi \underline{\psi}|^2 + |\Pi \overline{\psi} - \overline{\psi}|^2. \end{aligned}$$

Mit dem Satz 2.13 kann $\inf_{v \in K} |v - u_n|$ abgeschätzt werden durch

$$\begin{aligned} \inf_{v \in K} |v - u_n| & \leq \sqrt{|\underline{\psi} - \Pi \underline{\psi}|^2 + |\Pi \overline{\psi} - \overline{\psi}|^2} \\ & \leq ch^2 (|\underline{\psi}|_{H^2} + |\overline{\psi}|_{H^2}). \end{aligned} \quad (2.66)$$

Satz 2.33 Sei Ω ein polygonales, konvexes, beschränktes Gebiet in \mathbb{R}^N . Seien u bzw. u_n die Lösungen von (V) bzw. (2.33). Sind $\underline{\psi}, \overline{\psi} \in H^2(\Omega)$, wird \mathcal{U} durch den Raum der finiten Elemente S_h approximiert und ist $N \leq 3$, so gilt mit einer von h und u unabhängigen Konstante C

$$\|u - u_n\|_T + \|u - u_n\|_{L^\infty(0, T; \mathcal{H})} \leq C (|u_{n,0} - u_0| + h(|u|_{L^2(0, T; H^2)} + 1)).$$

Dabei bezeichnet $|u|_{L^2(0, T; H^2)}$ die Halbnorm bezüglich dem Ort, d.h. $|u|_{L^2(0, T; H^2)}^2 = \int_0^T |u|_{H^2}^2 dt$.

Beweis: Wird im Ergebnis vom Lemma 2.32 v_n durch Πu ersetzt und die Abschätzungen (2.63) und (2.66) verwendet, so folgt

$$\|u - u_n\|_T^2 + |u(T) - u_n(T)|^2 \leq C \left(|u_{n,0} - u_0|^2 + h^2 \int_0^T (|\dot{u} + Au - f| + |\dot{u}_n - \dot{u}|) |u|_{H^2} + |u|_{H^2}^2 + |\dot{u} + Au - f| (|\underline{\psi}|_{H^2} + |\bar{\psi}|_{H^2}) dt \right).$$

$|\dot{u} + Au - f|_T$, $|\dot{u}|_T$ und $|\dot{u}_n|_T$ können mit dem Korollar 2.6, dem Lemma 2.7 und dem Lemma 2.22 durch die Problemdaten abgeschätzt werden. Somit gilt

$$\begin{aligned} \|u - u_n\|_T^2 + |u(T) - u_n(T)|^2 &\leq C \left(|u_{n,0} - u_0|^2 + h^2 (|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T \right. \\ &\quad \left. + |f|_T + \|u_0\|_T + h\|u_0\|_T) |u|_{L^2(0,T;H^2)} + h^2 |u|_{L^2(0,T;H^2)}^2 \right. \\ &\quad \left. + h^2 (|f|_T + |A\underline{\psi}|_T + |A\bar{\psi}|_T) (|\underline{\psi}|_{L^2(0,T;H^2)} + |\bar{\psi}|_{L^2(0,T;H^2)}) \right) \\ &\leq C \left(|u_{n,0} - u_0|^2 + h^2 (|u|_{L^2(0,T;H^2)} + |u|_{L^2(0,T;H^2)}^2 + 1) \right). \end{aligned}$$

□

Wird für die Anfangsbedingung $u_{n,0}$ der semidiskreten Variationsungleichung die Interpolierende von u_0 verwendet, so folgt aus Satz 2.13 und dem Satz 2.33

$$\|u - u_n\|_T + \|u - u_n\|_{L^\infty(0,T;\mathcal{H})} \leq Ch (|u_0|_{H^1} + |u|_{L^2(0,T;H^2)} + 1). \quad (2.67)$$

2.4.2 Fehler zwischen den Lösungen der semidiskreten und der diskreten Variationsungleichung

Im nächsten Lemma untersuchen wir den Fehler zwischen den Lösungen von (S) und (D). Im Rest dieses Kapitels bezeichne $\lambda_{\max} := \lambda_{\max}^{\sqrt{M^{-1}T} S \sqrt{M^{-1}}}$ und $\lambda_{\min} := \lambda_{\min}^{\sqrt{M^{-1}T} S \sqrt{M^{-1}}}$.

Lemma 2.34 *Sei $y(t) \in \mathbb{R}^n \times (0, T)$ die Lösung der semidiskreten Variationsungleichung (S). Für die Lösung y^k von (D) gibt es eine von h und τ unabhängige Konstante c mit*

$$\|y(k\tau) - y^k\|_M \leq c\sqrt{\tau}.$$

Beweis: Seien $t_j := j\tau$ die Zeitpunkte der diskretisierten Zeitachse. Sei nun $z(t) \equiv y^k \in \tilde{K}_n$ konstant für $t \in (t_{k-1}, t_k]$ und $z(t) = y(t)$ ausserhalb $(t_{k-1}, t_k]$, so folgt aus (S)

$$\underbrace{\int_{t_{k-1}}^{t_k} (\dot{y}(t), M(y^k - y(t))) dt}_A + \underbrace{\int_{t_{k-1}}^{t_k} (y(t), S(y^k - y(t))) dt}_B - \underbrace{\int_{t_{k-1}}^{t_k} (l, y^k - y(t)) dt}_C \geq 0. \quad (2.68)$$

Die Integrale A, B und C werden nun durch eine Quadraturformel genähert. Es gilt

$$M(y^k - y(t)) = M(y^k - y(t_k)) + \int_t^{t_k} M \dot{y}(s) ds.$$

Damit folgt für A

$$A = \int_{t_{k-1}}^{t_k} (\dot{y}(t), M(y^k - y(t_k))) dt + \int_{t_{k-1}}^{t_k} \int_t^{t_k} (\dot{y}(t), M\dot{y}(s)) ds dt.$$

Wegen des Korollars 2.21 gilt

$$\int_{t_{k-1}}^{t_k} \int_t^{t_k} \|\dot{y}(t)\|_M \|\dot{y}(s)\|_M ds dt \leq \tau^2 c_A.$$

Für A erhält man somit die Abschätzung

$$A \leq (y(t_k) - y(t_{k-1}), M(y^k - y(t_k))) + c_A \tau^2.$$

Mit dem Korollar 2.21, dem Lemma 2.23 und dem Lemma 2.31 gelten folgende Abschätzungen

$$\begin{aligned} (\dot{y}(s), S(y^k - y(t_k))) &\leq \|\dot{y}(s)\|_M \|M^{-1}S(y^k - y(t_k))\|_M \\ &\leq \|\dot{y}(s)\|_M \left(\|M^{-1}S y^k\|_M + \|M^{-1}S y(t_k)\|_M \right) \leq c \end{aligned} \quad (2.69)$$

$$(y(t), S\dot{y}(s)) \leq \|\dot{y}(s)\|_M \|M^{-1}S y(t)\|_M \leq c. \quad (2.70)$$

Mit (2.69) und (2.70) schätzen wir B ab durch

$$\begin{aligned} B &= \int_{t_{k-1}}^{t_k} \left((y(t_k), S(y^k - y(t_k))) + \int_t^{t_k} (y(t), S\dot{y}(s)) ds - \int_t^{t_k} (\dot{y}(s), S(y^k - y(t_k))) ds \right) dt \\ &\leq (y(t_k), S(y^k - y(t_k)))\tau + c_B \tau^2. \end{aligned}$$

Für C folgt analog mit (2.41)

$$\begin{aligned} C &= (l, z - y(t_k))\tau + \int_{t_{k-1}}^{t_k} \int_t^{t_k} \underbrace{(l, \dot{y}(s))}_{\leq \|M^{-1}l\|_M \|\dot{y}(s)\|_M} ds dt \\ &\leq (l, z - y(t_k))\tau + c_C \tau^2. \end{aligned}$$

Nun kann die Variationsungleichung (2.68) für $z = y^k$ formuliert werden als

$$(M(y(t_k) - y(t_{k-1})) + \tau S y(t_k) - \tau l, y^k - y(t_k)) + \underbrace{(c_A + c_B + c_C)}_{=:c} \tau^2 \geq 0. \quad (2.71)$$

Einsetzen von $z = y(t_k)$ in die diskrete Variationsungleichung (D) führt nach Multiplikation von (D) mit τ und Addition mit (2.71) zu

$$\begin{aligned} (M((y(t_k) - y(t_{k-1})) - (y^k - y^{k-1})) + \tau S(y(t_k) - y^k), y^k - y(t_k)) + c\tau^2 &\geq 0, \\ ((M + \tau S)(y(t_k) - y^k) + M(y^{k-1} - y(t_{k-1})), y^k - y(t_k)) + c\tau^2 &\geq 0. \end{aligned}$$

Mit $e(k) := \sqrt{(M(y^k - y(t_k)), y^k - y(t_k))}$ gilt

$$e(k-1)e(k) - \lambda_{\min}^{\mathbb{1} + \tau\sqrt{M^{-1}S}\sqrt{M^{-1}}} e(k)^2 + c\tau^2 \geq 0.$$

Nach (2.55) ist $\lambda_{\min}^{\mathbb{1} + \tau\sqrt{M^{-1}S}\sqrt{M^{-1}}} \geq 1 + \tau\lambda_{\min}$ und somit

$$e(k-1)e(k) - (1 + \tau\lambda_{\min})e(k)^2 + c\tau^2 \geq 0.$$

Durch Auflösen nach $e(k)$ erhält man die iterative obere Schranke

$$e(k) \leq \frac{e(k-1) + \sqrt{e(k-1)^2 + 4(1 + \tau\lambda_{\min})c\tau^2}}{2(1 + \tau\lambda_{\min})},$$

$$\frac{e(k)}{\sqrt{\tau}} \leq \frac{\frac{e(k-1)}{\sqrt{\tau}} + \sqrt{\frac{e(k-1)^2}{\tau} + 4(1 + \tau\lambda_{\min})c\tau}}{2(1 + \tau\lambda_{\min})} =: G(k).$$

Ist $e(0) \leq \sqrt{\frac{c\tau}{\lambda_{\min}}}$, so ist die Folge $\{G(k)\}_{k=0}^{\infty}$ monoton steigend und beschränkt. Die Beschränktheit wird induktiv bewiesen. Die Verankerung gilt nach Voraussetzung, denn aus

$$e(0) \leq \sqrt{\frac{c}{\lambda_{\min}}}\sqrt{\tau} \quad \text{folgt} \quad G(0) = \frac{e(0)}{\sqrt{\tau}} \leq \sqrt{\frac{c}{\lambda_{\min}}}.$$

Für den Induktionsschritt sei $G(k-1) \leq \sqrt{\frac{c}{\lambda_{\min}}}$, dann gilt

$$\begin{aligned} \frac{G(k-1) + \sqrt{G(k-1)^2 + 4(1 + \tau\lambda_{\min})c\tau}}{2(1 + \tau\lambda_{\min})} &\leq \frac{\sqrt{\frac{c}{\lambda_{\min}}} + \sqrt{\frac{c}{\lambda_{\min}} + 4(1 + \tau\lambda_{\min})c\tau}}{2(1 + \tau\lambda_{\min})} \\ &= \sqrt{\frac{c}{\lambda_{\min}} \frac{1 + \sqrt{1 + 4(1 + \tau\lambda_{\min})\tau\lambda_{\min}}}{2(1 + \tau\lambda_{\min})}} \\ &= \sqrt{\frac{c}{\lambda_{\min}} \frac{1 + 1 + 2\tau\lambda_{\min}}{2(1 + \tau\lambda_{\min})}} = \sqrt{\frac{c}{\lambda_{\min}}}. \end{aligned}$$

Mit $e(0) \leq \sqrt{\frac{c\tau}{\lambda_{\min}}}$ ist somit für alle k $e(k) \leq \sqrt{\frac{c}{\lambda_{\min}}}\sqrt{\tau}$. Nach dem Satz 2.16 kann λ_{\min} von h und τ unabhängig nach unten abgeschätzt werden. Der Fehler zu einem festen Zeitpunkt $T_0 \in \{t_1, t_2, \dots, t_{\frac{T}{\tau}}\}$ kann damit durch Verkleinerung der Zeitschritte τ beliebig klein gemacht werden. Er geht mit der Ordnung $\sqrt{\tau}$ gegen 0. \square

2.4.3 Gesamtfehler

$\{y^k\}_{k=0}^{T/\tau}$ sei die Lösung von (D). Daraus wird wie folgt eine Funktion in $L^2(0, T; S_h)$ durch Interpolation konstruiert.

Für $t \in (t_{k-1}, t_k]$ sei

$$Y(t) := y^{k-1} + \frac{t - t_{k-1}}{\tau}(y^k - y^{k-1}). \quad (2.72)$$

Damit lässt sich eine Funktion aus $L^2(0, T; S_h)$ definieren durch

$$U_n(x, t) := \sum_{i=0}^n Y_i(t)\phi_i(x). \quad (2.73)$$

Im nächsten Satz soll der Fehler zwischen der Lösung u von (V) und U_n untersucht werden.

Satz 2.35 (Gesamtfehler) *Sei u die Lösung von (V) und U_n gegeben durch (2.72) und (2.73). Seien $u_{n,0} = \Pi u_0$ und $u_0, \underline{\psi}, \overline{\psi} \in H^2$. Es existiert eine Konstante c , die unabhängig von h und τ ist, mit*

$$|u - U_n|_T \leq c(h + \tau + \sqrt{\tau}).$$

Beweis: Sei u_n die Lösung von (2.33), d.h. $y(t)$ ist Lösung von (S). Die zeitliche Interpolierende $Iu_n(x, t)$ von $u_n(x, t)$ sei definiert durch

$$Iy(t) := y(t_{k-1}) + \frac{t - t_{k-1}}{\tau} (y(t_k) - y(t_{k-1})) \quad \text{für } t \in (t_{k-1}, t_k],$$

$$Iu_n(x, t) := \sum_{i=1}^n Iy_i(t) \phi_i(x).$$

Mit der Dreiecksungleichung gilt

$$|u - U_n|_T \leq \underbrace{|u - u_n|_T}_A + \underbrace{|u_n - Iu_n|_T}_B + \underbrace{|Iu_n - U_n|_T}_C.$$

Für A folgt mit (2.67)

$$A \leq c (|u_0|_{H^1} + |u|_{L^2(0,T;H^2)} + 1) h. \quad (2.74)$$

Die Norm B kann geschrieben werden als

$$B^2 = \int_0^T (y(t) - Iy(t), M(y(t) - Iy(t))) dt$$

$$\leq \int_0^T \underbrace{\|y(t) - Iy(t)\|_M^2}_{D^2} dt.$$

Nach dem Korollar 2.21 ist $\dot{y}(t)$ für alle t beschränkt. Damit liegen die Koordinatenfunktionen $y_i(t)$ in $W^{1,\infty}(0, T)$ für $i = 1, \dots, n$. Für den Interpolationsfehler D verwenden wir die folgende Fehlerabschätzung für die Interpolation von Funktionen aus $W^{1,\infty}$ durch lineare Splinefunktionen (vgl. z.B. Schumaker [20])

$$\sup_{t \in (0,T)} \|y(t) - Iy(t)\|_M \leq \tau \sup_{t \in (0,T)} \|\dot{y}(t)\|_M.$$

Es folgt mit dem Korollar 2.21

$$B \leq \sqrt{T} (|f| + c|u_0|_{H^2} + |Au_0|) \tau. \quad (2.75)$$

Die Abschätzung der Norm C kann wie bei der Norm B auf die Abschätzung der Koeffizientenfunktionen reduziert werden.

$$C^2 = \int_0^T (Iy(t) - Y(t), M(Iy(t) - Y(t))) dt$$

$$\leq \int_0^T \underbrace{\|Iy(t) - Y(t)\|_M^2}_{E^2} dt.$$

In den Zeitpunkten $k\tau$ lässt sich E mit dem Lemma 2.34 abschätzen. $Iy(t)$ und $Y(t)$ sind stückweise linear. Somit kann E für einen beliebigen Zeitpunkt $t \in (0, T)$ abgeschätzt werden durch

$$E \leq c\sqrt{\tau}.$$

Damit folgt für C

$$C \leq \sqrt{T} c \sqrt{\tau}. \quad (2.76)$$

Aus (2.74), (2.75) und (2.76) erhält man die Behauptung. \square

Kapitel 3

Lösungsmethode für stark monotone Variationsungleichungen

Seien \mathcal{H} ein beliebiger Hilbertraum, K eine konvexe, beschränkte Teilmenge von \mathcal{H} und $F : K \rightarrow \mathcal{H}$ ein Operator auf K . F heisst stark monoton in $K \subset \mathcal{H}$, falls für ein $\alpha > 0$ und für alle $u, v \in K$

$$(F(u) - F(v), u - v) \geq \alpha|u - v|^2$$

gilt. Gesucht wird ein $u \in K$, das die Variationsungleichung

$$(F(u), v - u) \geq 0 \quad \forall v \in K \quad (3.1)$$

erfüllt. Die Existenz und Eindeutigkeit der Lösung u kann wie im Lemma 2.28 bewiesen werden.

Definition 9 Für ein beliebiges $v \in K$ heisst die Menge

$$C_v := \{w \in \mathcal{H} : t_v(w) := (F(v), v - w) - \alpha|v - w|^2 \geq 0\}$$

quadratischer Schnitt an der Stelle v .

Ein quadratischer Schnitt C_v ist eine Kugel mit Mittelpunkt $v - \frac{1}{2\alpha}F(v)$ und Radius $\frac{1}{2\alpha}|F(v)|$. Mit der Charakterisierung dieser Kugel folgt nämlich

$$\begin{aligned} \left| w - \left(v - \frac{1}{2\alpha}F(v) \right) \right|^2 &\leq \frac{1}{4\alpha^2}|F(v)|^2, \\ 2 \left(w - v, \frac{1}{2\alpha}F(v) \right) + |w - v|^2 + \frac{1}{4\alpha^2}|F(v)|^2 &\leq \frac{1}{4\alpha^2}|F(v)|^2, \\ (F(v), w - v) + \alpha|w - v|^2 &\leq 0. \end{aligned}$$

Lemma 3.1 Die Lösung u von (3.1) ist für beliebiges $v \in K$ in C_v enthalten.

Beweis: Für $t_v(u)$ gilt mit (3.1) und da F stark monoton ist

$$\begin{aligned} t_v(u) &= (F(v), v - u) - \alpha|v - u|^2 \\ &= \underbrace{(F(v) - F(u), v - u)}_{\geq \alpha|v - u|^2} + (F(u), v - u) - \alpha|v - u|^2 \\ &\geq (F(u), v - u) \geq 0. \end{aligned}$$

□

Definition 10 Eine implizite Mengencharakterisierung der Menge K heisst Orakel, falls sie für jedes $v \in K$ den Wert „true“ und für $v \notin K$ den Wert „false“ liefert. Zudem liefert sie im Fall $v \notin K$ einen Halbraum H_v mit $v \notin H_v$ und $K \subset H_v$.

Besteht K aus allen Funktionen aus $\mathcal{H} = L^2(\Omega)$, die fast überall in Ω zwischen den Hindernisfunktionen $\underline{\psi}$ und $\overline{\psi}$ liegen, so kann wie folgt ein Orakel für K definiert werden. Gilt $v \notin K$, so ist

$$H_v := \{w \in \mathcal{H} : ((v - \underline{\psi})^- + (v - \overline{\psi})^+, v - (v - \underline{\psi})^- - (v - \overline{\psi})^+ - w) \geq 0\}$$

ein Halbraum von \mathcal{H} , der die gewünschte Trennungseigenschaft besitzt. Es ist $v \notin H_v$, da

$$((v - \underline{\psi})^- + (v - \overline{\psi})^+, -(v - \underline{\psi})^- - (v - \overline{\psi})^+) = -|(v - \underline{\psi})^-|^2 - |(v - \overline{\psi})^+|^2 < 0.$$

Für ein beliebiges $w \in K$ gilt

$$\begin{aligned} & ((v - \underline{\psi})^- + (v - \overline{\psi})^+, v - (v - \underline{\psi})^- - (v - \overline{\psi})^+ - w) \\ &= \underbrace{((v - \underline{\psi})^-, \underline{\psi} - w)}_{\leq 0} + \underbrace{((v - \overline{\psi})^+, \overline{\psi} - w)}_{\geq 0} \geq 0. \end{aligned}$$

Die Lösungsmethode ist ein iteratives Verfahren, das eine Folge u^j konstruiert, welche gegen u konvergiert. Zu Beginn wird die Menge K durch eine geeignete einfache Obermenge $K^0 \supset K$, zum Beispiel eine Kugel oder ein Polytop, approximiert. Im Iterationsschritt j wird zunächst ein Element u^j im Innern von K^j bestimmt. Nun können zwei Fälle eintreten:

1. **u^j ist in K enthalten (Optimalitätsschnitt):** Mit dem quadratischen Schnitt C_{u^j} wird ein Teil von K^j weggeschnitten, $K^{j+1} := K^j \cap C_{u^j}$. Nach dem Lemma 3.1 liegt die Lösung u auch in K^{j+1} . Es folgt die nächste Iteration.
2. **u^j ist nicht in K enthalten (Zulässigkeitsschnitt):** Mit Hilfe des Orakels wird u^j von K abgetrennt. Die Approximation K^0 von K wird verbessert durch $K^0 := K^0 \cap H_{u^j}$. Damit ändert sich auch K^j . Erneut wird ein Element u^j in K^j bestimmt. Dies wird so lange wiederholt, bis wir im 1. Fall sind.

3.1 Numerische Lösungsmethode mit quadratischen Schnitten

In diesem Kapitel werden die M -Norm und das M -Skalarprodukt als Standard verwendet. Zur Vereinfachung der Schreibweise bezeichne $|\cdot| := \|\cdot\|_M$ und $(\cdot, \cdot) := (\cdot, \cdot)_M$. Es bezeichne $T_j := \sum_{i=0}^{j-1} (t_i(v^j))^{-1}$, $L := \max_{z \in K^0} |F(z)|$ und $D_K := \max_i (\overline{\psi}_i - \underline{\psi}_i)$. Die übrigen Bezeichnungen des Kapitels 2 werden übernommen. Für zwei Matrizen A, B gilt $A \geq B$, falls $A - B$ eine positiv semidefinite Matrix ist. $\text{Diag}(a_1, \dots, a_n)$ bezeichne die Diagonalmatrix mit den Diagonalelementen a_1, \dots, a_n .

Es soll nun die diskrete Variationsungleichung (D) numerisch gelöst werden. Wir nehmen an, dass y^{k-1} bekannt ist und werden daraus eine Näherung zu y^k bestimmen.

Ist $F(y) := \left(\frac{1}{\tau} + M^{-1}S\right)y - \frac{1}{\tau}y^{k-1} - M^{-1}l$, so erfüllt $y^k \in \tilde{K}_n$ die Variationsungleichung

$$\left(F(y^k), z - y^k\right) \geq 0 \quad \forall z \in \tilde{K}_n. \quad (3.2)$$

Geometrisch kann man diese Variationsungleichung wie folgt auffassen. F beschreibt ein Vektorfeld auf \tilde{K}_n . Gesucht wird entweder ein Punkt in \tilde{K}_n , wo das Vektorfeld verschwindet, oder ein Punkt auf dem Rand von \tilde{K}_n , so dass die Menge \tilde{K}_n auf der positiven Seite der Hyperebene liegt, welche in diesem Randpunkt durch das Vektorfeld bestimmt wird. Die Abbildung 3.1 zeigt die geometrische Situation in einem zweidimensionalen Beispiel.

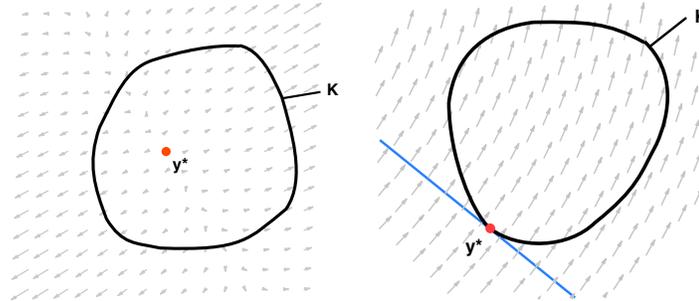


Abbildung 3.1: $y^* \in K$ ist Lösung von $(F(y), z - y) \geq 0 \quad \forall z \in K$.

Im Beweis des Korollars 2.29 wurde gezeigt, dass $F : \tilde{K}_n \rightarrow \mathbb{R}^n$ stark monoton ist mit $\alpha = \frac{1}{\tau} + \lambda_{\min}^{M^{-1}S}$. Dies ist auch der Grund für die Eindeutigkeit der Lösung von (D). Die quadratischen Schnitte sind in diesem Fall Teilmengen von \mathbb{R}^n ,

$$C_y := \{z \in \mathbb{R}^n : t_y(z) := (F(y), y - z) - \alpha|y - z|^2 \geq 0\}$$

Die Idee des Algorithmus besteht nun darin, an einem vorgegebenen Startpunkt v^0 in K^0 mit dem quadratischen Schnitt C_{v^0} einen Teil der Menge K^0 wegzuschneiden, anschliessend in der Restmenge $K^1 = K^0 \cap C_{v^0}$ einen inneren Punkt v^1 zu bestimmen, erneut mit einem quadratischen Schnitt einen Teil der Menge K^1 wegzuschneiden und so weiter. Die Punktfolge v^j konvergiert gegen die gesuchte Lösung y^k . Die Abbildung 3.2 zeigt die geometrische Idee an Hand eines zweidimensionalen Beispiels. Ist \tilde{K}_n ein Polytop, so kann K^0 gerade gleich \tilde{K}_n gewählt werden. In diesem Fall liegen alle Punkte v^j in \tilde{K}_n und es sind keine Zulässigkeitschnitte nötig, da \tilde{K}_n exakt durch K^0 „approximiert“ wird.

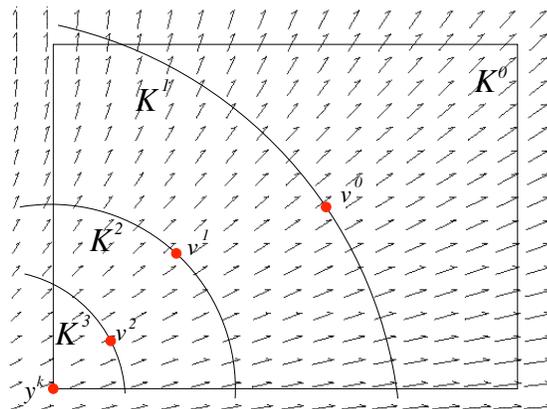


Abbildung 3.2: Geometrische Idee des Algorithmus

Im Fall der Hindernisprobleme ist \tilde{K}_n ein n -dimensionaler Quader, der durch $2n$ lineare Ungleichungen bestimmt wird. Ist z_i die i -te Komponente des Vektors z , $\underline{\psi}_i := \underline{\psi}(x^i)$ und $\overline{\psi}_i := \overline{\psi}(x^i)$,

so gilt für $z \in \tilde{K}_n$

$$\underline{\psi}_i \leq z_i \leq \overline{\psi}_i \quad \text{für } i = 1, \dots, n.$$

Die Menge $K^0 = \tilde{K}_n$ kann auch durch n quadratische Restriktionen beschrieben werden. Seien

$$s_i(z) := (z_i - \underline{\psi}_i)(\overline{\psi}_i - z_i) \quad \text{für } i = 1, \dots, n.$$

Damit gilt

$$K^0 = \{z \in \mathbb{R}^n : s_i(z) \geq 0, i = 1, \dots, n\}. \quad (3.3)$$

Definition 11 Für $j \geq 1$ heisst die Funktion

$$B_j(z) := -\frac{1}{n} \sum_{i=1}^n \log s_i(z) - \sum_{i=0}^{j-1} \log t_{v^i}(z)$$

logarithmische Barrierefunktion. Zur Vereinfachung wird $t_{v^i}(z)$ durch $t_i(z)$ ersetzt.

Diese Barrierefunktion entspricht der, bei Schnittebenenverfahren üblichen, Barrierefunktion mit n -fachen Schnitten. Die Restriktionen der Ausgangsmenge K^0 werden weniger stark gewichtet als die neuen quadratischen Schnitte.

Für den Gradienten und die Hessische Matrix von $B_j(z)$ benötigt man

$$\begin{aligned} (\nabla s_i(z))_j &= \left(-2z_i + (\underline{\psi}_i + \overline{\psi}_i)\right) \delta_{ij} \quad \text{für } j = 1, \dots, n, \\ (D^2 s_i(z))_{jl} &= -2\delta_{ji}\delta_{li}, \\ \nabla t_i(z) &= -F(v^i) - 2\alpha(z - v^i), \\ D^2 t_i(z) &= -2\alpha \mathbb{I}. \end{aligned} \quad (3.4)$$

Damit folgt für $B_j(z)$

$$\begin{aligned} \nabla B_j(z) &= -\frac{1}{n} \sum_{i=1}^n \frac{\nabla s_i(z)}{s_i(z)} - \sum_{i=0}^{j-1} \frac{\nabla t_i(z)}{t_i(z)}, \\ D^2 B_j(z) &= -\frac{1}{n} \sum_{i=1}^n \frac{D^2 s_i(z) s_i(z) - \nabla s_i(z) \nabla s_i(z)^\top}{s_i(z)^2} - \sum_{i=0}^{j-1} \frac{D^2 t_i(z) t_i(z) - \nabla t_i(z) \nabla t_i(z)^\top}{t_i(z)^2} \\ &= \frac{1}{n} \text{Diag} \left(\frac{2}{s_1(z)}, \dots, \frac{2}{s_n(z)} \right) + \sum_{i=1}^n \frac{\nabla s_i(z) \nabla s_i(z)^\top}{n s_i(z)^2} + \sum_{i=0}^{j-1} \frac{2\alpha \mathbb{I}}{t_i(z)} + \sum_{i=0}^{j-1} \frac{\nabla t_i(z) \nabla t_i(z)^\top}{t_i(z)^2}. \end{aligned} \quad (3.6)$$

Nach der Quotientenregel von Rayleigh ist die Summe positiv semidefiniter Matrizen wieder positiv semidefinit. Der kleinste Eigenwert der Summe ist grösser oder gleich der Summe der kleinsten Eigenwerte. Somit ist $D^2 B_j(z)$ für alle z aus dem Innern von K^j positiv definit. Die Definitionsmenge von B_j ist gerade das Innere von K^j . Daher existiert ein eindeutiger Punkt $v^j \in \text{Int}(K^j)$, wo B_j minimiert wird.

Definition 12 Der Punkt v^j in K^j mit

$$B_j(v^j) = \min_{z \in \text{Int}(K^j)} B_j(z)$$

heisst analytisches Zentrum von K^j . Der Minimierer wird auch mit der Schreibweise $v^j = \text{argmin}_{z \in K^j} B_j(z)$ bezeichnet.

Formal kann der Algorithmus zur numerischen Lösung der diskreten Variationsungleichung wie folgt beschrieben werden. Sei v^* die gesuchte Lösung von (3.2).

ACQCM: ANALYTIC CENTER QUADRATIC CUT METHOD
(i) Setze $\epsilon > 0$. Seien $K^0 := \tilde{K}_n$, $B_0(z) := -\frac{1}{n} \sum_{i=1}^n \log(s_i(z))$, $j = 0$.
(ii) Berechne das analytische Zentrum $v^j = \operatorname{argmin}_{z \in K^j} B_j(z)$, sei $B_{j+1}(z) := B_j(z) - \log(t_j(z))$ und $K^{j+1} := K^j \cap C_{v^j}$.
(iii) Ist $ v^j - v^* > \epsilon$, so setze $j := j + 1$ und gehe zu (ii), sonst halt.

Mit $y^k = v^j$ kann $F(y) := \left(\frac{\mathbb{1}}{\tau} + M^{-1}S\right)y - \frac{1}{\tau}y^k - M^{-1}l$ neu definiert werden. Der gleiche Algorithmus liefert nun eine Näherung für y^{k+1} .

Lemma 3.2 *Sei v^* die Lösung der Variationsungleichung (3.2). Ist*

$$\min_{z \in K^0} (F(v^j), z - v^j) \geq -\epsilon^2 \alpha, \quad (3.7)$$

so folgt

$$|v^j - v^*| \leq \epsilon.$$

Beweis: Da v^* die Lösung von (3.2) ist, erhält man für $z = v^*$ aus (3.7) mit der starken Monotonie von F

$$\begin{aligned} -\epsilon^2 \alpha &\leq (F(v^j), v^* - v^j) \\ &\leq (F(v^j), v^* - v^j) + (F(v^*), v^j - v^*) \\ &= -(F(v^j) - F(v^*), v^j - v^*) \\ &\leq -\alpha |v^j - v^*|^2. \end{aligned}$$

□

Das Abbruchkriterium $|v^j - v^*| \leq \epsilon$ im Algorithmus ACQCM kann nicht direkt berechnet werden, da v^* nicht bekannt ist. Das Lemma 3.2 liefert eine Möglichkeit, wie das Kriterium trotzdem berechnet werden kann. Für festes v^j ist (3.7) ein lineares Minimierungsproblem über dem Polytop K^0 und somit berechenbar. Wir werden im Folgenden eine Konvergenzrate der Punktfolge v^j gegen v^* herleiten, woraus sich eine obere Schranke für die Anzahl Iterationen j ableiten lässt. Das heisst, dass der Algorithmus ACQCM zu einer beliebigen Fehlertoleranz $\epsilon > 0$ in endlich vielen Iterationen eine Näherung zu v^* berechnet. Lüthi und Büeler [15] haben bewiesen, dass mit $\theta := \frac{1}{2}(1 + \sqrt{5})$ der Fehler in der j -ten Iteration abgeschätzt werden kann durch

$$|v^j - v^*| \leq \frac{L}{\alpha} \sqrt{\frac{3 + \sqrt{5}}{j}} 2^{\frac{n}{j} - \frac{1}{2}} e^{\sqrt{n}\theta + 1}. \quad (3.8)$$

Für $j > n$ ist die Konvergenzrate von v^j gegen v^* $\mathcal{O}\left(\frac{1}{\sqrt{j}}\right)$. Bei Lüthi und Büeler bezeichnet n die Anzahl Restriktionen, die K^0 beschreiben, und ist unabhängig von der Dimension des zu

Grunde liegenden Raumes. Da im Fall der Hindernisprobleme die Anzahl Restriktionen von K^0 gerade gleich der Dimension des Raumes \mathbb{R}^n ist (vgl. (3.3)), scheint das Resultat (3.8) nicht geeignet zu sein. Wird die Maschenweite h der räumlichen Diskretisierung verkleinert, so steigt n und damit $e^{\sqrt{n}\theta}$ in (3.8) stark an. Wir werden den Beweis von (3.8) aus [15] abändern, um ein besseres Ergebnis zu erhalten. Eine wesentliche Änderung besteht darin, dass die quadratischen Schnitte mit der Dimension n gewichtet werden. Das Hauptresultat dieses Teils ist der folgende Satz:

Satz 3.3 *Sei v^* die Lösung der Variationsungleichung (3.2). Ist v^j die mit dem Algorithmus ACQCM berechnete Näherung in der j -ten Iteration, so gilt mit einer von n und j unabhängigen Konstante c*

$$|v^j - v^*| < c \frac{L}{\alpha} n^{\frac{1}{3j}} \frac{1}{\sqrt{j}}.$$

Der Beweis verläuft in mehreren Schritten und benötigt einige technische Lemmata.

Lemma 3.4 *Sei v^j das analytische Zentrum von K^j . Für jede Komponente von v^j gilt*

$$\underline{\psi}_i + \frac{1}{nj+2}(\bar{\psi}_i - \underline{\psi}_i) \leq v_i^j \leq \bar{\psi}_i - \frac{1}{nj+2}(\bar{\psi}_i - \underline{\psi}_i) \quad i = 1, \dots, n.$$

Beweis: Das Lemma wird für eine beliebige Komponente v_i^j bewiesen. Der Beweis der anderen Komponenten von v^j folgt analog. Sei e_i der Einheitsvektor in der i -ten Koordinatenrichtung. Wir schränken die Suche nach dem analytischen Zentrum auf die Gerade $v^j + re_i$ ein. Da v^j im Innern der Kugeln C_{v^l} $l = 0, \dots, j-1$ liegt, schneidet die Gerade $v^j + re_i$ die Kugeloberflächen jeweils in zwei Punkten. Sei $z^l := v^l - \frac{1}{2\alpha}F(v^l)$ der Mittelpunkt und $R^l := \frac{1}{2\alpha}|F(v^l)|$ der Radius von C_{v^l} . Damit lässt sich $t_l(w)$ schreiben als $t_l(w) = \alpha((R^l)^2 - |z^l - w|^2)$. Für den Parameter r gelten mit $l = 0, \dots, j-1$ die Restriktionen

$$\begin{aligned} r &\leq \bar{q} := \bar{\psi}_i - v_i^j, \\ -r &\leq \underline{p} := v_i^j - \underline{\psi}_i, \\ r &\leq \bar{q}^l := z_i^l - v_i^l + \sqrt{(R^l)^2 - |v^j - z^l|^2 + |z_i^l - v_i^l|^2} > 0, \\ -r &\leq \underline{p}^l := -(z_i^l - v_i^l) + \sqrt{(R^l)^2 - |v^j - z^l|^2 + |z_i^l - v_i^l|^2} > 0. \end{aligned} \tag{3.9}$$

Die logarithmische Barrierefunktion für die Restriktionen (3.9) ist

$$\tilde{B}(r) = -\frac{1}{n} \log(\bar{q} - r) - \frac{1}{n} \log(\underline{p} + r) - \sum_{l=0}^{j-1} \left(\log(\bar{q}^l - r) + \log(r + \underline{p}^l) \right).$$

Sei $\tilde{r} = \operatorname{argmin}_r \tilde{B}(r)$ das analytische Zentrum von (3.9). Es ist a priori nicht klar, dass

$\tilde{v} = v^j + \tilde{r}e_i$ auch das analytische Zentrum von K^0 ist, d.h. dass $\tilde{r} = 0$ gilt. Es ist

$$\begin{aligned}
\tilde{B}(r) &= -\frac{1}{n} \log(-r^2 + r(\bar{q} - \underline{p}) + \bar{q}\underline{p}) - \sum_{l=0}^{j-1} \log(-r^2 + r(\bar{q}^l - \underline{p}^l) + \bar{q}^l \underline{p}^l) \\
&= -\frac{1}{n} \log(-r^2 + r(\bar{\psi}_i + \underline{\psi}_i - 2v_i^j) - \bar{\psi}_i \underline{\psi}_i - (v_i^j)^2 + v_i^j(\bar{\psi}_i + \underline{\psi}_i)) \\
&\quad - \sum_{l=0}^{j-1} \log(-r^2 + 2r(z_i^l - v_i^j) + (R^l)^2 - |v^j - z^l|^2) \\
&= -\frac{1}{n} \log((v_i^j + r - \underline{\psi}_i)(\bar{\psi}_i - (v_i^j + r))) - \sum_{l=0}^{j-1} \log((R^l)^2 - |v^j + re_i - z^l|^2) \\
&= \underbrace{-\frac{1}{n} \log(s_i(v^j + re_i)) - \sum_{l=0}^{j-1} \log(t_l(v^j + re_i))}_{=B(v^j + re_i)} + j \log \alpha.
\end{aligned}$$

$\tilde{B}(r)$ und $B(v^j + re_i)$ unterscheiden sich nur durch einen von r unabhängigen Term. Somit besitzen beide den gleichen Minimierer. Seien $q = \min(\{\bar{q}^l : l = 0, \dots, j-1\}, \bar{q})$ und $p = \min(\{\underline{p}^l : l = 0, \dots, j-1\}, \underline{p})$. Die Behauptung des Lemmas, auf das mit r parametrisierte Problem übertragen, lautet

$$\frac{1}{nj+2} \leq \frac{\underline{p}}{\underline{p} + \bar{q}} \quad \text{und} \quad \frac{1}{nj+2} \leq \frac{\bar{q}}{\underline{p} + \bar{q}}.$$

Wir beweisen die stärkeren Aussagen

$$\frac{1}{nj+2} \leq \frac{p}{p+q} \quad \text{und} \quad \frac{1}{nj+2} \leq \frac{q}{p+q}.$$

Für den Widerspruchsbeweis wird $\frac{1}{nj+2} > \frac{q}{p+q}$ angenommen. Damit gilt

$$\frac{1}{q} > \frac{nj+2}{p+q}, \tag{3.10}$$

$$p - \left(\frac{p+q}{nj+2} - q \right) < p,$$

$$(p+q) \left(1 - \frac{1}{nj+2} \right) < p. \tag{3.11}$$

Die Ableitung von \tilde{B} an der Stelle $r = 0$ kann abgeschätzt werden durch

$$\begin{aligned}
n \frac{d}{dr} \tilde{B}(0) &= \frac{1}{\bar{q}} - \frac{1}{\underline{p}} + n \sum_{l=0}^{j-1} \left(\frac{1}{\bar{q}^l} - \frac{1}{\underline{p}^l} \right) \\
&> \frac{1}{q} - \frac{1+nj}{p} \\
&\stackrel{(3.10)}{>} \frac{1}{p+q} - \frac{1+nj}{(p+q) \left(1 - \frac{1}{nj+2} \right)} \\
&\stackrel{(3.11)}{=} \frac{(nj+1)(nj+2) - (1+nj)(nj+2)}{(p+q)(nj+1)} = 0.
\end{aligned}$$

Dies steht im Widerspruch dazu, dass $\tilde{B}(r)$ an der Stelle $r = 0$ ihr Minimum annimmt. Die Behauptung $\frac{1}{nj+2} \leq \frac{p}{p+q}$ wird analog bewiesen. \square

Nesterov und Nemirovskii zeigen in [17], dass für logarithmische Barrierefunktionen gemäss Definition 11 mit $v, z \in K^0$

$$(\nabla B(z) - \nabla B(v), z - v) \geq \frac{\|z - v\|_{D^2 B(v)}^2}{1 + \|z - v\|_{D^2 B(v)}} \quad (3.12)$$

gilt. Der Vollständigkeits halber geben wir das folgende Lemma aus [15] mit Beweis wieder.

Lemma 3.5 Sei $\theta := \frac{1}{2}(1 + \sqrt{5})$. Für $j \geq 1$ gilt

$$\|v^j - v^{j-1}\|_{D^2 B_{j-1}(v^j)} \leq \theta, \quad (3.13)$$

$$|v^j - v^{j-1}| \leq \sqrt{\frac{1+\theta}{2\alpha}} \frac{1}{\sqrt{T_j}}, \quad (3.14)$$

$$|v^j - v^{j-1}| \leq \frac{L}{\alpha} \sqrt{\frac{1+\theta}{8}} \frac{1}{\sqrt{j}}. \quad (3.15)$$

Beweis: Nach der Definition von v^j gilt $\nabla B_j(v^j) = 0$. Für $j \geq 1$ folgt aus (3.5)

$$\nabla B_{j-1}(v^j) = \frac{\nabla t_{j-1}(v^j)}{t_{j-1}(v^j)}. \quad (3.16)$$

Mit $t_{j-1}(v^{j-1}) = 0$ erhält man aus der Taylorentwicklung von $t_{j-1}(z)$ zusammen mit (3.16)

$$\begin{aligned} 0 &= t_{j-1}(v^{j-1}) = t_{j-1}(v^j) + (\nabla t_{j-1}(v^j), v^{j-1} - v^j) + \frac{1}{2} (v^{j-1} - v^j, D^2 t_{j-1}(v^j)(v^{j-1} - v^j)), \\ &(\nabla t_{j-1}(v^j), v^j - v^{j-1}) = t_{j-1}(v^j) - \alpha |v^{j-1} - v^j|^2, \\ &(\nabla B_{j-1}(v^j), v^j - v^{j-1}) = 1 - \alpha \frac{|v^{j-1} - v^j|^2}{t_{j-1}(v^j)}. \end{aligned} \quad (3.17)$$

Wegen der Minimaleigenschaft von v^{j-1} ist $\nabla B_{j-1}(v^{j-1}) = 0$. Wird $(\nabla B_{j-1}(v^{j-1}), v^j - v^{j-1}) = 0$ auf der linken Seite von (3.17) subtrahiert, so folgt mit (3.12)

$$1 - \alpha \frac{|v^{j-1} - v^j|^2}{t_{j-1}(v^j)} \geq \frac{\|v^{j-1} - v^j\|_{D^2 B_{j-1}(v^j)}^2}{1 + \|v^{j-1} - v^j\|_{D^2 B_{j-1}(v^j)}}. \quad (3.18)$$

v^j liegt im Innern von K^j und damit ist $t_{j-1}(v^j) > 0$. Es folgt

$$0 \geq \|v^{j-1} - v^j\|_{D^2 B_{j-1}(v^j)}^2 - \|v^{j-1} - v^j\|_{D^2 B_{j-1}(v^j)} - 1.$$

Auflösen dieser quadratischen Ungleichung führt zur Behauptung (3.13).

Aus (3.18) folgt für $j = 1$

$$\begin{aligned} \alpha \frac{|v^1 - v^0|^2}{t_0(v^1)} &\leq 1, \\ |v^1 - v^0|^2 &\leq \frac{1}{\alpha T_1} \leq \frac{1+\theta}{2\alpha} \frac{1}{T_1}. \end{aligned}$$

Damit ergibt sich (3.14) für $j = 1$. Ist $j \geq 2$, so gilt nach (3.6)

$$\begin{aligned} D^2 B_{j-1}(v^j) &= \underbrace{D^2 B_0(v^j)}_{\geq 0} - \sum_{i=0}^{j-2} \frac{D^2 t_i(v^j)}{t_i(v^j)} + \underbrace{\sum_{i=0}^{j-2} \frac{\nabla t_i(v^j) \nabla t_i(v^j)^\top}{t_i(v^j)^2}}_{\geq 0} \\ &\geq 2\alpha \sum_{i=0}^{j-2} \frac{\mathbb{I}}{t_i(v^j)}. \end{aligned}$$

Wird nun $v^j - v^{j-1}$ bezüglich diesen Matrizen genormt, so ergibt sich mit (3.18)

$$\begin{aligned} \|v^j - v^{j-1}\|_{D^2 B_{j-1}(v^j)}^2 &\geq |v^j - v^{j-1}| \sum_{i=0}^{j-2} \frac{2\alpha}{t_i(v^j)}, \\ 1 - \alpha \frac{|v^j - v^{j-1}|}{t_{j-1}(v^j)} &\stackrel{(3.18)}{\geq} \frac{1}{1+\theta} |v^j - v^{j-1}| \sum_{i=0}^{j-2} \frac{2\alpha}{t_i(v^j)}, \\ 1 &\geq \frac{1}{1+\theta} \left(\sum_{i=0}^{j-2} \frac{2\alpha}{t_i(v^j)} + \frac{\alpha(1+\theta)}{t_{j-1}(v^j)} \right) |v^j - v^{j-1}|^2. \end{aligned}$$

Mit $1 + \theta \geq 2$ folgt (3.14) für $j \geq 2$.

Gemäss der Definition 9 und der dort anschliessenden Bemerkung sind die Niveaumengen von $t_i(z)$ konzentrische Kugeloberflächen. Daher nimmt $t_i(z)$ ihr Maximum im Mittelpunkt $v^i - \frac{1}{2\alpha} F(v^i)$ an. Es gilt

$$\begin{aligned} \max_{z \in \mathbb{R}^n} t_i(z) &= \left(F(v^i), v^i - v^i + \frac{1}{2\alpha} F(v^i) \right) - \alpha \left| v^i - v^i + \frac{1}{2\alpha} F(v^i) \right|^2 \\ &= \frac{1}{2\alpha} |F(v^i)|^2 - \frac{1}{4\alpha} |F(v^i)|^2 \\ &\leq \frac{1}{4\alpha} L^2. \end{aligned} \tag{3.19}$$

Die Behauptung (3.15) folgt aus (3.14) zusammen mit der Schranke (3.19). \square

Lemma 3.6 Für $j \geq 1$ gilt

$$-\sum_{i=0}^{j-1} \log t_i(v^j) \geq -\sum_{i=0}^{j-1} \log t_i(v^{i+1}) - \log nj - \frac{3}{nj} + 2 \log 2.$$

Beweis: Es gilt

$$-\sum_{i=0}^{j-1} \log t_i(v^j) = \underbrace{B_j(v^j) - B_0(v^0)}_{(*)} + \underbrace{B_0(v^0) - B_0(v^j)}_{(**)}. \tag{3.20}$$

Mit der Definition der Barrierefunktion und der Minimaleigenschaft von v^{i-1} folgt für (*)

$$\begin{aligned} B_i(v^i) &= B_{i-1}(v^i) - \log t_{i-1}(v^i) \geq B_{i-1}(v^{i-1}) - \log t_{i-1}(v^i), \\ B_i(v^i) - B_{i-1}(v^{i-1}) &\geq -\log t_{i-1}(v^i). \end{aligned}$$

Die Summation über $i = 1, \dots, j$ führt zu

$$B_j(v^j) - B_0(v^0) \geq - \sum_{i=0}^{j-1} \log t_i(v^{i+1}). \quad (3.21)$$

Für (**) wird $B_0(v^j)$ nach oben abgeschätzt. Nach dem Lemma 3.4 kann das analytische Zentrum v^j nicht beliebig nahe am Rand von K^0 liegen. $s_i(v^j)$ ist somit immer echt grösser als 0. Es folgt mit dem Lemma 3.4

$$\begin{aligned} s_i(v^j) &\geq \frac{1}{nj+2} (\bar{\psi}_i - \underline{\psi}_i)^2 \frac{nj+1}{nj+2}, \\ -\log s_i(v^j) &\leq \log \frac{(nj+2)^2}{nj+1} - 2 \log(\bar{\psi}_i - \underline{\psi}_i). \end{aligned} \quad (3.22)$$

Da die Logarithmusfunktion strikt konkav ist, gilt $\log(nj+2) < \log(nj+1) + \frac{1}{nj+1}$. Damit kann der Logarithmus in (3.22) abgeschätzt werden durch

$$\log \frac{(nj+2)^2}{nj+1} = 2 \log(nj+2) - \log(nj+1) < \log(nj+1) + \frac{2}{nj+1} < \log nj + \frac{3}{nj}.$$

Aus (3.22) folgt nach Summation über $i = 1, \dots, n$

$$B_0(v^j) \leq \log nj + \frac{3}{nj} - 2 \frac{1}{n} \sum_{i=1}^n \log(\bar{\psi}_i - \underline{\psi}_i). \quad (3.23)$$

Da $s_i(v)$ ihr Maximum für $v_i = \frac{\bar{\psi}_i + \underline{\psi}_i}{2}$ annimmt und $B_0(v)$ symmetrisch bezüglich diesem v ist, kann $B_0(v^0)$ exakt angegeben werden. Es ist

$$B_0(v^0) = -\frac{1}{n} \sum_{i=1}^n \log s_i(v^0) = -\frac{1}{n} \sum_{i=1}^n \log \left(\frac{\bar{\psi}_i - \underline{\psi}_i}{2} \right)^2 = \frac{-2}{n} \sum_{i=1}^n \log(\bar{\psi}_i - \underline{\psi}_i) + 2 \log 2. \quad (3.24)$$

Die Behauptung folgt nun aus (3.20) zusammen mit (3.21), (3.23) und (3.24). \square

Lemma 3.7 Für $j \geq 1$ und $s \geq 1$ gelten die Summenabschätzungen

$$\sum_{i=0}^{j-1} \frac{1}{i+1} (\log(i+1))^s \leq \frac{1}{s+1} (\log j)^{s+1} + \frac{1}{2} (\log j)^s, \quad (3.25)$$

$$\sum_{i=1}^j \frac{1}{i} \leq \log j + 1, \quad (3.26)$$

$$\sum_{i=1}^j \frac{1}{i^2} \leq 2 - \frac{1}{j}, \quad (3.27)$$

$$\sum_{i=1}^j \log i > j \log j - j. \quad (3.28)$$

Beweis: Für $j = 1$ ist (3.25) trivial. Im Fall $j \geq 2$ beweisen wir die stärkere Aussage

$$\sum_{i=0}^{j-1} \frac{1}{i+1} (\log(i+1))^s < \frac{1}{s+1} (\log j)^{s+1} + \frac{1}{j} (\log j)^s. \quad (3.29)$$

Es wird gezeigt, dass die Zunahme der linken Seite beim Übergang $j \rightarrow j+1$ kleiner als die Zunahme der rechten Seite ist. Da die Logarithmusfunktion konkav ist, folgt für die Zunahme der rechten Seite von (3.29)

$$\begin{aligned} & \frac{(\log(j+1))^{s+1}}{s+1} + \frac{(\log(j+1))^s}{j+1} - \frac{(\log j)^{s+1}}{s+1} - \frac{(\log j)^s}{j} \\ & \geq \frac{(\log(j+1))^s}{j+1} + \frac{1}{s+1} \left(\log j + \frac{1}{j+1} \right)^{s+1} - \frac{(\log j)^{s+1}}{s+1} - \frac{(\log j)^s}{j+1} \\ & = \frac{(\log(j+1))^s}{j+1} + \frac{1}{s+1} \left((\log j)^{s+1} + (s+1) \frac{(\log j)^s}{j+1} \right. \\ & \quad \left. + \underbrace{\binom{s+1}{2} \frac{(\log j)^{s-1}}{(j+1)^2} + \dots + \frac{1}{(j+1)^{s+1}}}_{>0} \right) - \frac{(\log j)^{s+1}}{s+1} - \frac{(\log j)^s}{j+1}. \end{aligned}$$

Damit ist (3.25) gezeigt. (3.26) gilt, da $\frac{1}{x}$ monoton fallend ist. Somit ist

$$\sum_{i=1}^j \frac{1}{i} \leq \int_1^j \frac{1}{x} dx + 1 = \log j + 1.$$

Das gleiche Argument führt zu (3.27). (3.28) wird mit Hilfe der Formel von Stirling gezeigt. Es gilt

$$\begin{aligned} x! &= \sqrt{2\pi x} x^{x+\frac{1}{2}} e^{-x+\frac{\vartheta}{12x}} \quad \text{mit } 0 < \vartheta < 1, \\ \log x! &> x \log x - x. \end{aligned}$$

Damit folgt aus $\sum_{i=1}^j \log i = \log j!$ die Behauptung (3.28). □

Lemma 3.8 Mit $\eta := \frac{L^2}{\alpha} \sqrt{\frac{1+\theta}{8}}$ gilt für $j \geq 1$

$$-\sum_{i=0}^{j-1} \log t_i(v^j) \geq \frac{1}{2} j \log j - j \left(\log \eta + \frac{1}{2} \right) - \log nj - \frac{3}{nj}.$$

Beweis: Aus der Definition von $t_i(z)$ folgt mit (3.15)

$$t_i(v^{i+1}) \leq L|v^i - v^{i+1}| \quad (3.30)$$

$$\leq \underbrace{\frac{L^2}{\alpha} \sqrt{\frac{1+\theta}{8}}}_{=\eta} \frac{1}{\sqrt{i+1}}. \quad (3.31)$$

Da v^{i+1} im Innern von K^i liegt, ist $t_i(v^{i+1}) > 0$ und wir können (3.31) logarithmieren. Dies führt zu

$$\begin{aligned} -\log t_i(v^{i+1}) &\geq -\log \eta + \frac{1}{2} \log(i+1), \\ -\sum_{i=0}^{j-1} \log t_i(v^{i+1}) &\geq -j \log \eta + \frac{1}{2} \sum_{i=1}^j \log i. \end{aligned} \quad (3.32)$$

Mit (3.28) erhält man

$$-\sum_{i=0}^{j-1} \log t_i(v^{i+1}) \geq -j \log \eta + \frac{1}{2} j \log j - \frac{j}{2}.$$

Zusammen mit dem Lemma 3.6 folgt die Behauptung. \square

Die bewiesene untere Schranke für $\mathcal{B} := -\sum_{i=0}^{j-1} \log t_i(v^j)$ soll nun mit einem iterativen Schema verbessert werden. Wir nehmen $\gamma := L\sqrt{\frac{1+\theta}{2\alpha}} \geq 1$ an. Die Definition von L rechtfertigt diese Annahme.

Lemma 3.9 *Für alle $j \geq 1$ gilt*

$$\mathcal{B} = -\sum_{i=0}^{j-1} \log t_i(v^j) > j \log j - j(2 \log \gamma + 2) - j^{\frac{2}{3}} \left(\frac{4}{n} + 2 \log n + 2 \right) - \frac{2}{jn} + 2.$$

Beweis: Da das geometrische Mittel kleiner oder gleich dem arithmetischen Mittel ist, folgt

$$\begin{aligned} T_j &= j \frac{1}{j} \sum_{i=0}^{j-1} (t_i(v^j))^{-1} \\ &\geq j \sqrt[j]{\prod_{i=0}^{j-1} (t_i(v^j))^{-1}} \\ &= j \exp \left(\frac{-1}{j} \sum_{i=0}^{j-1} \log t_i(v^j) \right). \end{aligned} \quad (3.33)$$

Wir wenden nun das von Lüthi und Büeler vorgeschlagene iterative Schema an.

- (o) Das Lemma 3.8 liefert eine erste untere Schranke für $\mathcal{B} = -\sum_{i=0}^{j-1} \log t_i(v^j)$.
- (i) Mit (3.33) wird T_j von unten beschränkt.
- (ii) Diese untere Schranke eingesetzt in (3.14) ergibt eine obere Schranke für $|v^j - v^{j-1}|$ und mit (3.30) eine obere Schranke für $t_i(v^{i+1})$.
- (iii) Das Lemma 3.6 liefert mit dieser Schranke für $t_i(v^{i+1})$ eine neue, bessere untere Schranke für \mathcal{B} . Mit dieser gehen wir zu (i).

Im ersten Durchgang gilt:

(i) Für alle $j \geq 1$ ist

$$\begin{aligned} T_j &\geq j \exp \left(\frac{1}{2} \log j - \underbrace{\left(\log \eta + \frac{1}{2} \right)}_{=: C_1} - \frac{1}{j} \log nj - \frac{3}{nj^2} \right) \\ &= \exp \left(\frac{3}{2} \log j - C_1 - \frac{1}{j} \log nj - \frac{3}{nj^2} \right). \end{aligned}$$

(ii) Daraus folgt

$$\begin{aligned} |v^j - v^{j-1}| &\stackrel{(3.14)}{\leq} \sqrt{\frac{1+\theta}{2\alpha}} \exp \left(-\frac{3}{4} \log j + \frac{1}{2} C_1 + \frac{1}{2} \frac{1}{j} \log nj + \frac{3}{2} \frac{1}{nj^2} \right), \\ \log t_i(v^{i+1}) &\stackrel{(3.30)}{\leq} \log \gamma - \frac{3}{4} \log(i+1) + \frac{1}{2} C_1 + \frac{1}{2} \frac{1}{i+1} \log(n(i+1)) + \frac{3}{2} \frac{1}{n(i+1)^2}. \end{aligned}$$

(iii) Summation über $i = 0, \dots, j-1$ führt zu

$$\begin{aligned} - \sum_{i=0}^{j-1} \log t_i(v^{i+1}) &\geq -j \log \gamma + \underbrace{\frac{3}{4} \sum_{i=1}^j \log i}_{(A)} - j \frac{1}{2} C_1 - \frac{1}{2} \log n \underbrace{\sum_{i=1}^j \frac{1}{i}}_{(B)} \\ &\quad - \underbrace{\frac{1}{2} \sum_{i=1}^{j-1} \frac{1}{i+1} \log(i+1)}_{(C)} - \underbrace{\frac{3}{2n} \sum_{i=1}^j \frac{1}{i^2}}_{(D)}. \end{aligned} \tag{3.34}$$

(A) kann mit (3.28), (B) mit (3.26), (C) mit (3.25) und (D) mit (3.27) abgeschätzt werden. Damit folgt aus (3.34)

$$\begin{aligned} - \sum_{i=0}^{j-1} \log t_i(v^{i+1}) &> -j \log \gamma + \frac{3}{4} (j \log j - j) - j \frac{1}{2} C_1 - \frac{1}{2} \log n (\log j + 1) \\ &\quad - \frac{1}{2} \left(\frac{1}{2} (\log j)^2 + \frac{1}{2} \log j \right) - \frac{3}{2n} \left(2 - \frac{1}{j} \right) \\ &= \frac{3}{4} j \log j - j \underbrace{\left(\log \gamma + \frac{3}{4} + \frac{1}{2} C_1 \right)}_{=: C_2} - \log j \left(\frac{1}{2} \log n + \frac{1}{4} \right) \\ &\quad - \frac{(\log j)^2}{2} \frac{1}{2} - \frac{1}{j} \left(-\frac{3}{2n} \right) - \left(\frac{1}{2} \log n + \frac{3}{n} \right). \end{aligned}$$

Mit dem Lemma 3.6 erhält man

$$\mathcal{B} > \frac{3}{4} j \log j - j C_2 - \log j \left(\frac{1}{2} \log n + \frac{5}{4} \right) - \frac{(\log j)^2}{2} \frac{1}{2} - \frac{1}{j} \frac{3}{2n} - \left(\frac{3}{2} \log n + \frac{3}{n} \right).$$

Damit ist eine erste verbesserte untere Schranke für \mathcal{B} gefunden worden. Für den Iterationsschritt $l \rightarrow l+1$ machen wir die Annahme

$$\mathcal{B} > \nu_l j \log j - j C_l - \sum_{s=1}^l \frac{(\log j)^s}{s} Q_l^s - \frac{1}{j} R_l^1 - R_l^0. \tag{3.35}$$

(i) Mit (3.33) folgt

$$\begin{aligned} T_j &\geq j \exp \left(\nu_l \log j - C_l - \sum_{s=1}^l \frac{(\log j)^s}{j^s} Q_l^s - \frac{1}{j^2} R_l^1 - \frac{1}{j} R_l^0 \right) \\ &= \exp \left((1 + \nu_l) \log j - C_l - \sum_{s=1}^l \frac{(\log j)^s}{j^s} Q_l^s - \frac{1}{j^2} R_l^1 - \frac{1}{j} R_l^0 \right). \end{aligned}$$

(ii) Einsetzen dieser Schranke in (3.14) führt mit (3.30) zu

$$\begin{aligned} \log t_i(v^{i+1}) &\leq \log \gamma - \frac{1 + \nu_l}{2} \log(i+1) + \frac{1}{2} C_l + \frac{1}{2} \sum_{s=1}^l \frac{(\log(i+1))^s}{(i+1)^s} Q_l^s \\ &\quad + \frac{1}{(i+1)^2} \frac{R_l^1}{2} + \frac{1}{i+1} \frac{R_l^0}{2}. \end{aligned}$$

(iii) Summation über $i = 0, \dots, j-1$ führt zu

$$\begin{aligned} - \sum_{i=0}^{j-1} \log t_i(v^{i+1}) &\geq -j \log \gamma + \frac{1 + \nu_l}{2} \sum_{i=1}^j \log i - j \frac{C_l}{2} - \sum_{s=1}^l \sum_{i=1}^{j-1} \frac{(\log(i+1))^s}{s(i+1)} \frac{Q_l^s}{2} \\ &\quad - \frac{R_l^1}{2} \sum_{i=1}^j \frac{1}{i^2} - \frac{R_l^0}{2} \sum_{i=1}^j \frac{1}{i}. \end{aligned}$$

Mit dem Lemma 3.7 lassen sich die Summen abschätzen. Es gilt

$$\begin{aligned} - \sum_{i=0}^{j-1} \log t_i(v^{i+1}) &> -j \log \gamma + \frac{1 + \nu_l}{2} (j \log j - j) - \sum_{s=1}^l \left(\frac{(\log j)^{s+1}}{s(s+1)} + \frac{(\log j)^s}{2s} \right) \frac{Q_l^s}{2} \\ &\quad - j \frac{C_l}{2} - \frac{R_l^1}{2} \left(2 - \frac{1}{j} \right) - \frac{R_l^0}{2} (\log j + 1). \end{aligned}$$

Das Lemma 3.6 liefert nun für \mathcal{B} die neue untere Schranke

$$\begin{aligned} \mathcal{B} &> \overbrace{\frac{1 + \nu_l}{2}}{=: \nu_{l+1}} j \log j - j \overbrace{\left(\log \gamma + \frac{1}{2} C_l + \frac{1 + \nu_l}{2} \right)}{=: C_{l+1}} - \sum_{s=1}^{l-1} \frac{(\log j)^{s+1}}{s+1} \overbrace{\left(\frac{1}{2s} Q_l^s + \frac{1}{4} Q_l^{s+1} \right)}{=: Q_{l+1}^{s+1}} \\ &\quad - \log j \underbrace{\left(\frac{R_l^0}{2} + \frac{Q_l^1}{4} + 1 \right)}{=: Q_{l+1}^1} - \frac{(\log j)^{l+1}}{l+1} \underbrace{\frac{Q_l^l}{2l}}{=: Q_{l+1}^{l+1}} - \frac{1}{j} \underbrace{\left(\frac{3}{n} - \frac{R_l^1}{2} \right)}{=: R_{l+1}^1} - \underbrace{\left(\frac{R_l^0}{2} + R_l^1 + \log n \right)}{=: R_{l+1}^0}. \end{aligned}$$

Für die Folgen der Koeffizienten gelten folgende Rekursionsformeln und Anfangsbedingungen:

$$\begin{aligned} \nu_{l+1} &= \frac{1 + \nu_l}{2} & \nu_1 &= \frac{1}{2}, \\ C_{l+1} &= \log \gamma + \frac{C_l}{2} + \nu_{l+1} & C_1 &= \log \eta + \frac{1}{2}, \\ Q_{l+1}^1 &= \frac{R_l^0}{2} + \frac{1}{4} Q_l^1 + 1 & Q_2^1 &= \frac{1}{2} \log \eta + \frac{5}{4}, \end{aligned}$$

$$\begin{aligned}
Q_{l+1}^{l+1} &= \frac{1}{2l} Q_l^l & Q_2^2 &= \frac{1}{2}, \\
Q_{l+1}^s &= \frac{1}{2(s-1)} Q_l^{s-1} + \frac{1}{4} Q_l^s & s &= 2, \dots, l, \\
R_{l+1}^1 &= \frac{3}{n} - \frac{1}{2} R_l^1 & R_2^1 &= \frac{3}{2n}, \\
R_{l+1}^0 &= \frac{1}{2} R_l^0 + R_l^1 + \log n & R_2^0 &= \frac{3}{2} \log n + \frac{3}{n}.
\end{aligned}$$

Im Limes $l \rightarrow \infty$ erhält man für die Koeffizienten

$$\begin{aligned}
\nu_\infty &= 1, & C_\infty &= 2 \log \gamma + 2, \\
R_\infty^1 &= \frac{2}{n}, & R_\infty^0 &= \frac{4}{n} + 2 \log n, \\
Q_\infty^1 &= \frac{8}{3n} + \frac{4}{3} \log n + \frac{4}{3}, & Q_\infty^\infty &= 0.
\end{aligned} \tag{3.36}$$

Für Q_∞^s gilt die Gleichung

$$Q_\infty^s = \frac{2}{3(s-1)} Q_\infty^{s-1} = \left(\frac{2}{3}\right)^{s-1} \frac{1}{(s-1)!} Q_\infty^1. \tag{3.37}$$

(3.36) und (3.37) eingesetzt in (3.35) führt zu

$$\mathcal{B} > j \log j - j(2 \log \gamma + 2) - \frac{3}{2} Q_\infty^1 \underbrace{\sum_{s=1}^{\infty} \frac{(\frac{2}{3} \log j)^s}{s!}}_{=\exp(\frac{2}{3} \log j) - 1} - \frac{1}{j} \frac{2}{n} - \frac{4}{n} - 2 \log n.$$

□

Nun kann das Hauptresultat bewiesen werden.

Beweis von Satz 3.3: Der Beweis basiert darauf, dass die untere Schranke für \mathcal{B} aus dem Lemma 3.9 schneller wächst als die folgende obere Schranke.

Sei $\bar{v}^j := \frac{1}{2}(v^j + v^*)$. Die Taylorentwicklung von $t_i(v)$ um v^j führt zu

$$t_i(v^*) - t_i(v^j) + \alpha |v^* - v^j|^2 = (\nabla t_i(v^j), v^* - v^j).$$

Mit $\bar{v}^j - v^j = \frac{1}{2}(v^* - v^j)$ folgt

$$\frac{1}{2} (t_i(v^*) - t_i(v^j) + \alpha |v^* - v^j|^2) = (\nabla t_i(v^j), \bar{v}^j - v^j). \tag{3.38}$$

Wird \bar{v}^j in die Taylorentwicklung eingesetzt, so ergibt sich mit (3.38)

$$\begin{aligned}
t_i(\bar{v}^j) &= t_i(v^j) + (\nabla t_i(v^j), \bar{v}^j - v^j) - \alpha |\bar{v}^j - v^j|^2 \\
&\stackrel{(3.38)}{=} t_i(v^j) + \frac{1}{2} (t_i(v^*) - t_i(v^j) + \alpha |v^* - v^j|^2) - \frac{1}{4} \alpha |v^* - v^j|^2 \\
&= \frac{1}{2} (t_i(v^j) + t_i(v^*)) + \frac{1}{4} \alpha |v^* - v^j|^2.
\end{aligned}$$

Da der Logarithmus konkav ist, folgt nach Streichung des positiven Terms $t_i(v^*)$

$$\log t_i(\bar{v}^j) \geq \frac{1}{2} \log t_i(v^j) + \frac{1}{2} \log \left(\frac{1}{2} \alpha |v^* - v^j|^2 \right). \quad (3.39)$$

Da $s_i(v)$ konkav ist, gilt

$$s_i(\bar{v}^j) = s_i \left(\frac{1}{2} v^* + \frac{1}{2} v^j \right) \geq \frac{1}{2} s_i(v^*) + \frac{1}{2} s_i(v^j) \geq \frac{1}{2} s_i(v^j). \quad (3.40)$$

v^j minimiert B_j . Somit gilt mit (3.39) und (3.40)

$$\begin{aligned} B_j(v^j) &\leq -\frac{1}{n} \sum_{i=1}^n \log s_i(\bar{v}^j) - \sum_{i=0}^{j-1} \log t_i(\bar{v}^j) \\ &\leq -\frac{1}{n} \sum_{i=1}^n \log \left(\frac{1}{2} s_i(v^j) \right) - \frac{1}{2} \sum_{i=0}^{j-1} \log t_i(v^j) - \frac{1}{2} j \log \left(\frac{\alpha}{2} |v^* - v^j|^2 \right) \\ &= \log 2 - \frac{1}{n} \sum_{i=1}^n \log s_i(v^j) - \frac{1}{2} \sum_{i=0}^{j-1} \log t_i(v^j) - \frac{1}{2} j \log \left(\frac{\alpha}{2} |v^* - v^j|^2 \right), \\ \mathcal{B} = B_j(v^j) - B_0(v^j) &\leq 2 \log 2 - j \log \left(\frac{\alpha}{2} |v^* - v^j|^2 \right). \end{aligned}$$

Diese obere Schranke wird nun mit der unteren Schranke aus dem Lemma 3.9 verglichen. Es ist

$$j \log j - j(2 \log \gamma + 2) - j^{\frac{2}{3}} \left(\frac{4}{n} + 2 \log n + 2 \right) - \frac{2}{jn} + 2 < 2 \log 2 - j \log \left(\frac{\alpha}{2} |v^* - v^j|^2 \right).$$

Nach Division durch j und Logarithmierung erhält man

$$\begin{aligned} \frac{\alpha}{2} j \exp(-2 \log \gamma - 2) \exp \left(-\frac{\frac{4}{n} + 2 \log n + 2}{\sqrt[3]{j}} \right) \exp \left(\frac{2}{j^2 n} - \frac{2 - 2 \log 2}{j} \right) &< |v^* - v^j|^{-2}, \\ \frac{2}{\alpha j} \gamma^2 n^{\frac{2}{3j}} \underbrace{\exp \left(2 + \frac{\frac{4}{n} + 2}{\sqrt[3]{j}} - \frac{2}{j^2 n} + \frac{2 - 2 \log 2}{j} \right)}_{(A)} &> |v^* - v^j|^2. \end{aligned}$$

(A) ist monoton fallend in j und n . Damit kann (A) durch eine Konstante abgeschätzt werden und es folgt die Behauptung. \square

Korollar 3.10 *Sei h die Maschenweite der räumlichen und τ die Schrittweite der zeitlichen Diskretisierung. Seien v^* und v^j wie im Satz 3.3 bestimmt. Wird*

$$\tau = h^2$$

gewählt, so gilt

$$|v^* - v^j| < c(n(h)) \frac{1}{\sqrt[3]{j}} \frac{1}{\sqrt{j}}.$$

Beweis: Im Beweis vom Korollar 2.29 wurde $\alpha = \frac{1}{\tau} + \lambda_{\min}^{M^{-1}S}$ gezeigt. Sei λ_1 der kleinste Eigenwert des Gebietes Ω für das Laplace-Eigenwertproblem. Mit dem Satz 2.16 kann α nach unten abgeschätzt werden durch

$$\alpha \geq \frac{1}{\tau} + \lambda_1. \quad (3.41)$$

Für L gilt mit (2.41)

$$\begin{aligned} L &= \max_{z \in K^0} \left| \frac{z}{\tau} + M^{-1}S z - \frac{y^{k-1}}{\tau} - M^{-1}l \right| \\ &\leq \max_{z \in K^0} \left(\frac{1}{\tau} |z - y^{k-1}| + \lambda_{\max}^{M^{-1}S} |z| + |f| \right). \end{aligned}$$

Da K^0 bezüglich der M -Norm beschränkt ist, folgt mit dem Satz 2.16

$$L \leq c \left(\frac{1}{\tau} + \frac{1}{h^2} \right). \quad (3.42)$$

Aus (3.41) und (3.42) ergibt sich

$$\frac{L}{\alpha} \leq c \frac{\frac{1}{\tau} + \frac{1}{h^2}}{\frac{1}{\tau} + \lambda_1} \leq c \left(1 + \frac{\tau}{h^2} \right). \quad (3.43)$$

Mit der Wahl $\tau = h^2$ folgt die Behauptung aus (3.43) und dem Satz 3.3. \square

Bemerkung zur Komplexitätsanalyse des Algorithmus ACQCM

Der Algorithmus ACQCM arbeitet mit exakten analytischen Zentren. In einer konkreten Implementierung werden diese jedoch nur näherungsweise berechnet werden können, was einen Einfluss auf die Gesamtkomplexität des Algorithmus hat. Der Satz 3.3 macht eine Komplexitätsaussage bezüglich der Anzahl Iterationen des Algorithmus. Für eine arithmetische Komplexitätsaussage müsste die Anzahl Newtonschritte, die für die Berechnung einer Näherung des analytischen Zentrums v^{j+1} gegeben v^j nötig sind, untersucht werden.

3.2 Berechnungsbeispiele

Die aufgezeigten Beispiele basieren alle auf der Wärmeleitungsgleichung. Der elliptische Operator ist somit der Laplace-Operator. In den ersten zwei Beispielen ist die Menge Ω das Einheitsintervall $[0, 1]$. Der Funktionenraum \mathcal{U} wird durch S_h mit $h = \frac{1}{48}$ approximiert. Für die zeitliche Diskretisierung wird die Schrittweite $\tau = \frac{1}{2}h^2$ gewählt. $\bar{\psi}$ und $\underline{\psi}$ werden durch ihre Interpolierenden $\bar{\psi}_n = \Pi \bar{\psi}$, $\underline{\psi}_n = \Pi \underline{\psi}$ genähert.

In den Darstellungen zum ersten und zweiten Beispiel werden die Hindernisfunktionen gestrichelt und die approximierte Lösungsfunktion $\sum_{i=1}^n y_i^k \phi_i(x)$ der diskreten Variationsungleichung fett gezeichnet.

Im ersten Beispiel ist die Funktion f auf ganz Ω gleich -10 . Die Anfangsbedingung ist in der Abbildung 3.3(a) dargestellt. Die stationäre Lösung (Abbildung 3.3(d)) wurde durch eine

Lösung mit grossem t genähert und nicht als Lösung der elliptischen Variationsungleichung bestimmt. In diesem Beispiel hat die obere Hindernisfunktion keinen Einfluss auf die Lösung. Wird eine Hindernisfunktion weggelassen, so ist K^0 nicht mehr beschränkt und das analytische Zentrum ist nicht definiert. Existiert in jedem Zeitschritt k ein endlicher Startwert v^0 , zum Beispiel y^{k-1} , so ist K^1 beschränkt und alle gemachten Überlegungen können auf K^1 angewendet werden.

In der konkreten Berechnung für den Zeitschritt k wurde als erster innere Punkt v^0 nicht das analytische Zentrum von K^0 , sondern y^{k-1} verwendet. Die weiteren Iterationsschritte erfolgten gemäss Beschreibung des Algorithmus ACQCM. Mit diesem Vorgehen nutzt der Algorithmus die Eigenschaft, dass für kleine τ y^k nahe bei y^{k-1} liegen muss, aus. Die Anzahl benötigter Iterationen für das Auffinden einer Lösung konnte damit reduziert werden. Für die Näherung eines y^k mit der Fehlertoleranz $\epsilon = 0.001$ benötigte der Algorithmus ACQCM höchstens 17 Iterationen. Das Abbruchkriterium $|v^* - v^j| < \epsilon$ wurde mit Hilfe des Lemmas 3.2 überprüft.

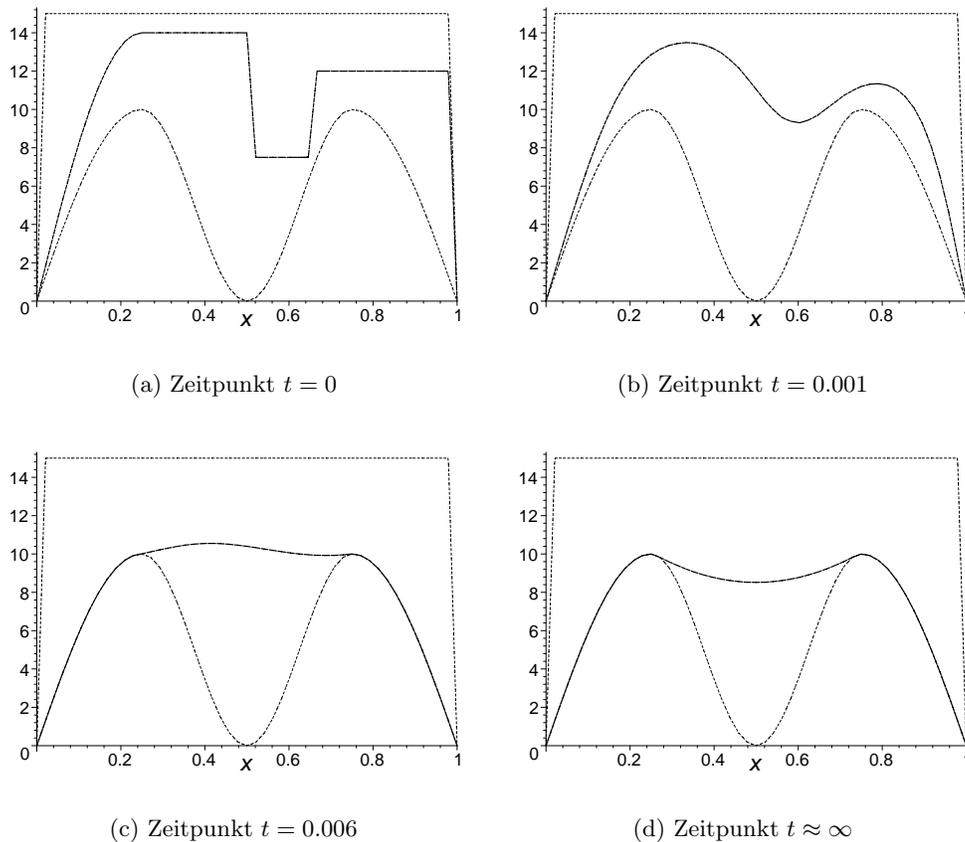


Abbildung 3.3: Erstes Beispiel zur Wärmeleitungsgleichung mit Hindernissen

Im zweiten Beispiel soll ein interessanter Aspekt der Diffusion gezeigt werden.

Lemma 3.11 Sei $u^* \in \tilde{K}$ die stationäre Lösung der Variationsungleichung (V), d.h. es gilt

$$a(u^*, \tilde{v} - u^*) - (f, \tilde{v} - u^*) \geq 0 \quad \forall \tilde{v} \in \tilde{K}. \quad (3.44)$$

Ist u die Lösung von (V), so gilt

$$|u(t) - u^*| \leq |u_0 - u^*| e^{-\alpha t}.$$

Beweis: Wir nehmen $u_0 \neq u^*$ an, da sonst nichts zu zeigen ist. Nach dem Lemma 2.10 erfüllt u für fast alle $t \in [0, T]$ die Variationsungleichung

$$(\dot{u}(t) - f, \tilde{v} - u(t)) + a(u(t), \tilde{v} - u(t)) \geq 0 \quad \forall \tilde{v} \in \tilde{K}. \quad (3.45)$$

Wird (3.45) mit $\tilde{v} = u^*$ getestet und zu (3.44) mit $\tilde{v} = u(t)$ addiert, ergibt sich

$$(\dot{u}(t), u^* - u(t)) + a(u(t) - u^*, u^* - u(t)) \geq 0.$$

Da $a(\cdot, \cdot)$ koerzitiv (vgl. (2.11)) ist, folgt zusammen mit $\dot{u}^* \equiv 0$

$$-\frac{d}{dt} |u^* - u(t)|^2 \geq 2\alpha \|u^* - u(t)\|^2 \geq 2\alpha |u^* - u(t)|^2.$$

Sei $\Phi(t) := |u^* - u(t)|^2$. Die Integration von 0 bis t führt zu

$$\int_0^t \frac{\dot{\Phi}(s)}{\Phi(s)} ds = \int_0^t \frac{d}{ds} \log \Phi(s) ds \leq -2\alpha t, \\ \Phi(t) \leq \Phi(0) e^{-2\alpha t}.$$

Damit ist die Behauptung bewiesen. □

Nach dem Lemma 3.11 strebt die L^2 -Norm über Ω von $u(t) - u^*$ monoton gegen 0. Diese Aussage gilt nicht lokal. Ein Gegenbeispiel liefert das zweite Beispiel, welches in der Abbildung 3.4 dargestellt ist.

Es ist $f(x) \equiv -30$ in ganz Ω . Über $\Omega' = [0.55, 0.8]$ nimmt die L^2 -Norm nicht monoton ab. u^* stimmt in Ω' mit der unteren Hindernisfunktion überein. In der Abbildung 3.4(c) ist zu erkennen, dass $u(t)$ für $t = 0.0016$ in Ω' mit der stationären Lösung zusammenfällt und im Zeitpunkt $t = 0.006$ (Abbildung 3.4(d)) sich wegen der Diffusion wieder teilweise von u^* entfernt.

Ein Lösungsalgorithmus kann somit nicht dadurch vereinfacht und beschleunigt werden, dass man sich auf die Teilgebiete von Ω mit $u(t) \neq u^*$ beschränkt.

Im letzten Beispiel ist die Menge Ω das Einheitsquadrat im \mathbb{R}^2 . Es ist $f(x) \equiv -1$ in ganz Ω . Für die Bildung des Raumes S_h wurde eine reguläre Triangulation mit Maschenweite $h = \frac{1}{21}$ verwendet. Die zeitliche Diskretisierung wurde mit $\tau = 0.00125$ durchgeführt. Das Ergebnis ist in der Bildfolge 3.5 dargestellt. Da die obere Hindernisfunktion $\bar{\psi}(x) \equiv 15$ keinen Einfluss auf die Lösung hat, wurde auf die Darstellung verzichtet.

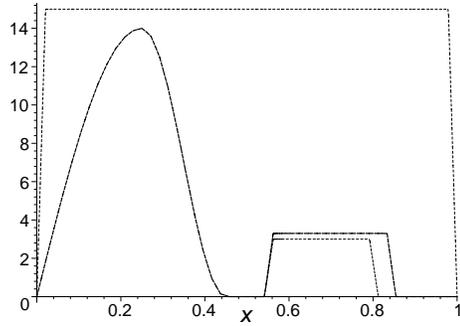
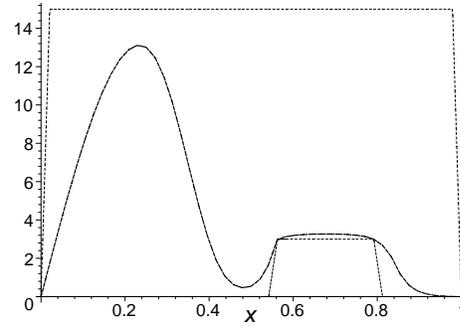
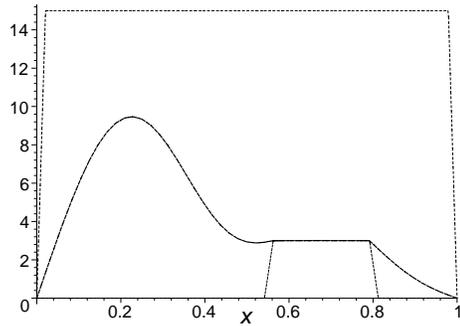
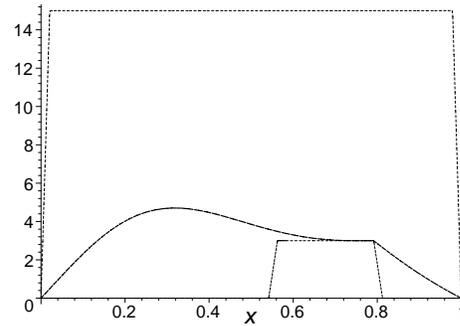
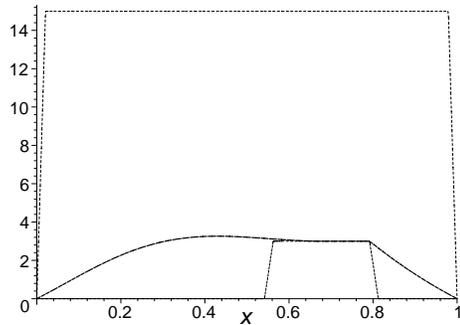
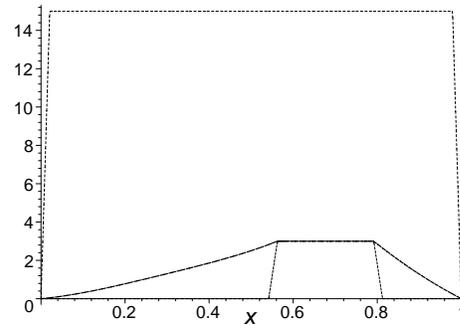
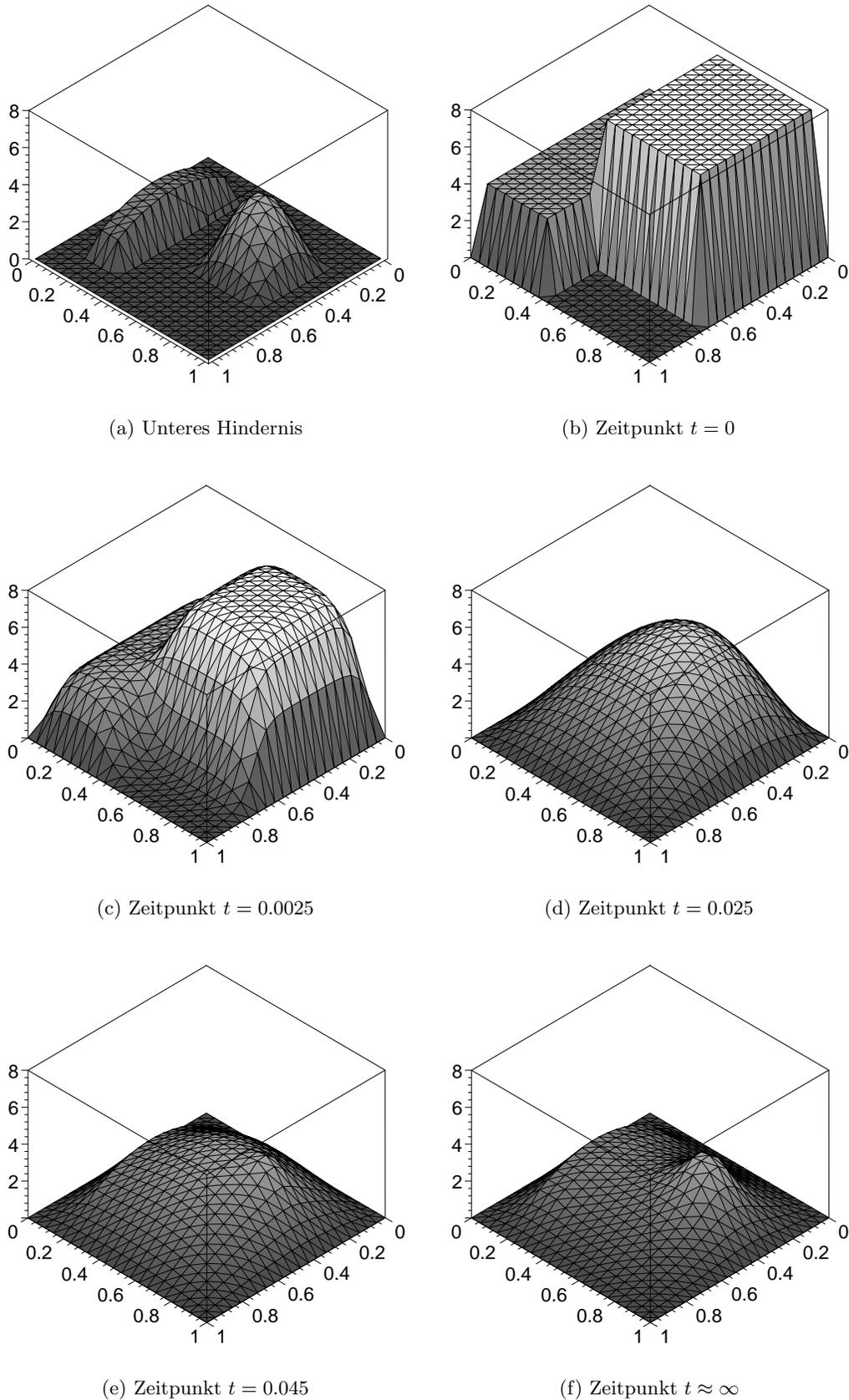
(a) Zeitpunkt $t = 0$ (b) Zeitpunkt $t = 0.0002$ (c) Zeitpunkt $t = 0.0016$ (d) Zeitpunkt $t = 0.006$ (e) Zeitpunkt $t = 0.01$ (f) Zeitpunkt $t \approx \infty$

Abbildung 3.4: Zweites Beispiel zur Wärmeleitungsgleichung mit Hindernissen

Abbildung 3.5: Beispiel zur Wärmeleitungsgleichung mit $\Omega \subset \mathbb{R}^2$

Kapitel 4

Zusatzbetrachtungen und Ausblick

4.1 Kollokationsverfahren bei Variationsungleichungen

Für den Übergang von der semidiskreten zur diskreten Variationsungleichung können verschiedene Differenzenschemen verwendet werden. Bei parabolischen Gleichungen kann mit dem Crank-Nicolson Schema die Fehlerordnung bezüglich der Zeit gegenüber dem Euler Schema verbessert werden. Mit dem Kollokationsverfahren lassen sich die verschiedenen Schemen einheitlich betrachten.

Wird die Idee des Kollokationsverfahrens für parabolische Gleichungen (vgl. [2]) in gleicher Art und Weise auf parabolische Variationsungleichungen übertragen, so erhält man mit dem Kollokationsparameter $c \in [0, 1]$ die diskrete Variationsungleichung

$$\left(\frac{1}{\tau} M(y^k - y^{k-1}) + S \left((1-c)y^{k-1} + cy^k \right) - l, v - \left((1-c)y^{k-1} + cy^k \right) \right) \geq 0 \quad \forall v \in \tilde{K}_n. \quad (4.1)$$

Für $c = 1$ entspricht dies dem rückwärts Euler, für $c = \frac{1}{2}$ dem Crank-Nicolson und für $c = 0$ dem vorwärts Euler Schema. Sei $y_c^k := (1-c)y^{k-1} + cy^k$. Mit $y^k - y^{k-1} = \frac{y_c^k - y^{k-1}}{c}$ kann (4.1) geschrieben werden als

$$\left(\frac{1}{\tau c} M(y_c^k - y^{k-1}) + S y_c^k - l, v - y_c^k \right) \geq 0 \quad \forall v \in \tilde{K}_n. \quad (4.2)$$

Für jedes k und $c \in (0, 1]$ lässt sich die Existenz und Eindeutigkeit einer Lösung y_c^k von (4.2) in \tilde{K}_n wie im Fall $c = 1$ (Korollar 2.29) beweisen. Für $c < 1$ muss $y^k = y^{k-1} + \frac{y_c^k - y^{k-1}}{c}$ aber nicht in \tilde{K}_n liegen.

Die Abbildung 4.1 liefert ein Gegenbeispiel zur Existenz einer Lösung $y^k \in \tilde{K}_n$. Das Beispiel basiert auf der Wärmeleitungsgleichung über dem Intervall $\Omega = [0, 1]$ mit $f(x) \equiv 0$, $\bar{\psi}(x) \equiv 10$ und $\underline{\psi}(x) \equiv 1$. Der Funktionenraum \mathcal{U} wird durch den eindimensionalen Raum S_h mit $h = 0.5$ approximiert. Der Koeffizientenvektor y einer Funktion aus S_h besteht aus einer Komponente, die den Funktionswert an der Stelle $x^1 = 0.5$ angibt. Die zeitliche Schrittweite beträgt $\tau = 0.05$. In der Abbildung 4.1 sind y^k $k = 0, \dots, 6$ für $c = 0, \frac{1}{2}$ und 1 dargestellt. Als Anfangsbedingung wurde $y^0 = 9$ verwendet. Zum Vergleich ist die Lösung der semidiskreten Variationsungleichung $y(t)$ durchgehend eingezeichnet. Für $c = 0$ bzw. $c = \frac{1}{2}$ ist deutlich zu erkennen, dass y^3 bzw. y^4 nicht in $\tilde{K}_n = [1, 9]$ liegen.

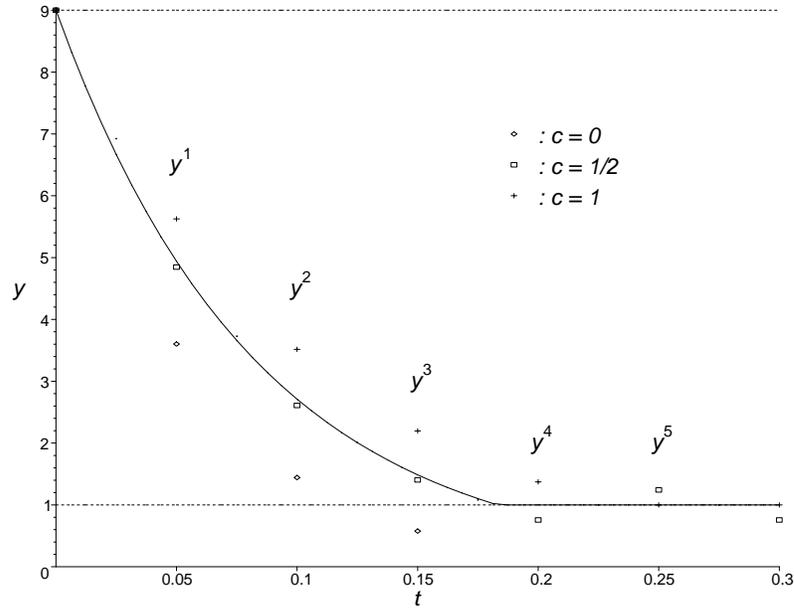


Abbildung 4.1: Gegenbeispiel zur Existenz einer Lösung in \tilde{K}_n für $c < 1$

Wie kann (4.1) geändert werden, um die Existenz einer Lösung in \tilde{K}_n zu sichern? Es drängen sich zwei Methoden auf:

1. Wegen $y^k = y^{k-1} + \frac{y_c^k - y^{k-1}}{c}$ muss y_c^k in (4.2) auf die Menge

$$\tilde{K}_n^{c,k} := \left\{ v \in \mathbb{R}^n : y^{k-1} + \frac{v - y^{k-1}}{c} \in \tilde{K}_n \right\}$$

eingeschränkt werden. Damit ist $y^k \in \tilde{K}_n$ gesichert. $\tilde{K}_n^{c,k}$ entsteht aus \tilde{K}_n durch zentrische Streckung mit dem Faktor c um das Zentrum y^{k-1} .

2. In der Variationsungleichung (4.1) ist der Teil $v - y_c^k$ dafür verantwortlich, dass y_c^k in \tilde{K}_n liegt. $y^k \in \tilde{K}_n$ kann daher erreicht werden mit der Variationsungleichung

$$\left(\frac{1}{\tau} M(y^k - y^{k-1}) + S y_c^k - l, v - y^k \right) \geq 0 \quad \forall v \in \tilde{K}_n. \quad (4.3)$$

Lemma 4.1 Für $c \in (0, 1]$ seien $\tilde{y}_c^k \in \tilde{K}_n^{c,k}$ die Lösung von

$$\left(\frac{1}{\tau c} M(\tilde{y}_c^k - y^{k-1}) + S \tilde{y}_c^k - l, \tilde{v} - \tilde{y}_c^k \right) \geq 0 \quad \forall \tilde{v} \in \tilde{K}_n^{c,k} \quad (4.4)$$

und $\tilde{y}^k = y^{k-1} + \frac{\tilde{y}_c^k - y^{k-1}}{c}$. Ist $y^k \in \tilde{K}_n$ die Lösung von (4.3) mit $y_c^k = (1-c)y^{k-1} + cy^k$, so gilt

$$\tilde{y}^k = y^k,$$

d.h. (4.3) und (4.4) sind äquivalent.

Beweis: Mit $y^k - y^{k-1} = \frac{y_c^k - y_c^{k-1}}{c}$ folgt nach Multiplikation mit c aus (4.3)

$$\left(\frac{1}{\tau c} M(y_c^k - y_c^{k-1}) + S y_c^k - l, \underbrace{cv + (1-c)y^{k-1}}_{\in \tilde{K}_n^{c,k}} - \underbrace{(1-c)y^{k-1} - cy^k}_{=-y_c^k} \right) \geq 0 \quad \forall v \in \tilde{K}_n.$$

Da zu jedem $\tilde{v} \in \tilde{K}_n^{c,k}$ ein $v \in \tilde{K}_n$ mit $\tilde{v} = cv + (1-c)y^{k-1}$ existiert und $y_c^k \in \tilde{K}_n^{c,k}$ aus $y^k \in \tilde{K}_n$ folgt, ist y_c^k auch Lösung der Variationsungleichung (4.4).

Umgekehrt schliessen wir mit $\tilde{v} = cv + (1-c)y^{k-1}$ und $\tilde{y}^k = y^{k-1} + \frac{\tilde{y}_c^k - y_c^{k-1}}{c}$ aus (4.4)

$$\left(\frac{1}{\tau} M(\tilde{y}^k - y^{k-1}) + S \tilde{y}_c^k - l, cv + (1-c)y^{k-1} - c\tilde{y}^k - (1-c)y^{k-1} \right) \geq 0 \quad \forall v \in \tilde{K}_n.$$

Nach Division durch c folgt, dass \tilde{y}^k auch Lösung von (4.3) ist. \square

Es bleibt die Frage offen, wie die Fehlerabschätzung aus dem Abschnitt 2.4.2 auf den Fall der Kollokation übertragen werden kann.

4.2 Semilineare Variationsungleichungen

Der vorgeschlagene Algorithmus ACQCM kann auch zur Lösung von semilinearen Problemen verwendet werden. Hängt f nicht linear von u ab, so gibt es zwei Möglichkeiten. Entweder ist $F(u) = \dot{u} + Au - f(u)$ auf K stark monoton oder nicht. Beide Fälle werden in einem Beispiel dargestellt.

Die diskrete Variationsungleichung wird wie im Abschnitt 2.3 hergeleitet. Die Semilinearität $f(u)$ wird durch ihre Interpolierende $\Pi f(u)$ genähert.

Das erste Beispiel basiert auf der Gleichung

$$\dot{\tilde{u}} - \Delta \tilde{u} = -c\sqrt{\tilde{u}} \quad \text{in } \Omega = [0, 1]$$

mit der Randbedingung $\tilde{u} = 1$ auf $\partial\Omega$. Die Randbedingung kann mit der Transformation $u = \tilde{u} - 1$ zu 0 geändert werden. F hat nun die Form

$$F(u) = \dot{u} - \Delta u + c\sqrt{u+1}.$$

Es gilt für $u, v \in L^2(0, T; \mathcal{U})$

$$\begin{aligned} \int_0^T (F(u) - F(v), u - v) dt &= \int_0^T (\dot{u} - \dot{v}, u - v) dt + \int_0^T (\nabla(u - v), \nabla(u - v)) dt \\ &\quad + c \int_0^T \underbrace{(\sqrt{u+1} - \sqrt{v+1}, u - v)}_A dt. \end{aligned}$$

A kann geschrieben werden als

$$\begin{aligned} A &= \int_{\Omega} (\sqrt{u+1} - \sqrt{v+1}) (u+1 - (v+1)) dx \\ &= \int_{\Omega} (\sqrt{u+1} - \sqrt{v+1})^2 (\sqrt{u+1} + \sqrt{v+1}) dx \geq 0. \end{aligned}$$

Damit ist $F(u)$ stark monoton auf $L^2(0, T; \mathcal{U})$. Für alle $v \in K$ gelte $v(x) \geq -1$ fast überall Ω . Sei $W : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $(W(y))_i := \sqrt{y_i}$ für $i = 1, \dots, n$. Die diskrete Variationsungleichung lautet

$$\left(\left(\frac{M}{\tau} + S \right) y^k - \frac{M}{\tau} y^{k-1} + cM W(y^k + 1), v - y^k \right) \geq 0 \quad \forall v \in \tilde{K}_n.$$

Da F auch mit der Semilinearität stark monoton ist, kann der Term $MW(y^k + 1)$ in impliziter Form belassen werden und in die Definition der quadratischen Schnitte integriert werden. Im Zeitschritt k gilt für beliebiges $v \in \tilde{K}_n$

$$C_v = \{z \in \mathbb{R}^n : t_v(z) = \left(\left(\frac{M}{\tau} + S \right) v - \frac{M}{\tau} y^{k-1} + cM W(v + 1), v - z \right) - \alpha |v - z|^2 \geq 0\}.$$

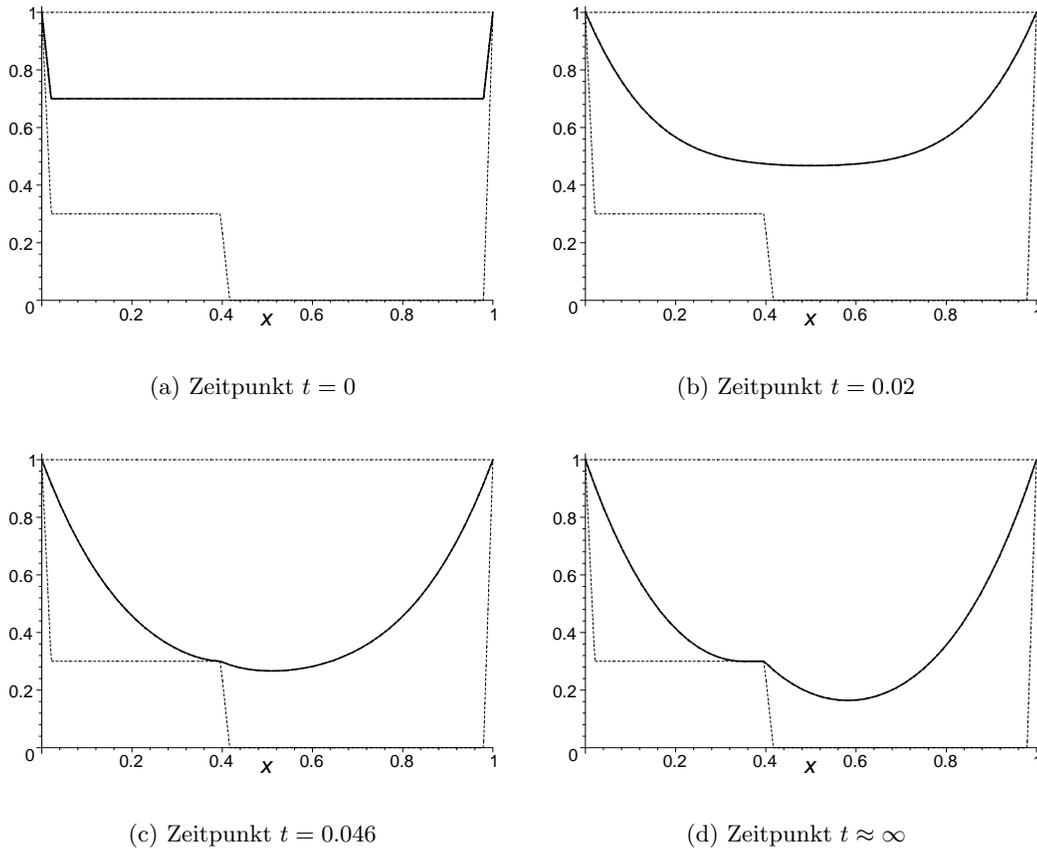


Abbildung 4.2: Semilineares Beispiel mit monotoner Semilinearität

Die Hindernisfunktionen und die Anfangsbedingung sind in der Abbildung 4.2(a) dargestellt. Es wurden $h = \frac{1}{48}$, $\tau = 0.002$, $c = 20$ und $\epsilon = 0.0005$ verwendet. Das Abbruchkriterium wurde mit dem Lemma 3.2 bestimmt. Der Algorithmus ACQCM fand in jedem Zeitschritt k eine Lösung nach höchstens 17 Iterationen. Das Ergebnis der Berechnungen ist in der Abbildung 4.2 gezeigt.

Das nächste Beispiel basiert auf der Gleichung

$$i - \Delta u - u^2 = 0 \quad \text{in } \Omega \text{ und } u = 0 \text{ auf } \partial\Omega,$$

welche bei genügend grosser Anfangbedingung zu Blow up führt. In diesem Beispiel ist $F(u) = \dot{u} - \Delta u - u^2$ nicht mehr stark monoton. Sei $Q(y^k) \in \mathbb{R}^n$ definiert durch $(Q(y^k))_i := (y^k)_i^2$ für $i = 1, \dots, n$. Die diskrete Variationsungleichung

$$\left(\left(\frac{M}{\tau} + S \right) y^k - \frac{M}{\tau} y^{k-1} - M Q(y^k), v - y^k \right) \geq 0 \quad \forall v \in \tilde{K}_n$$

kann nicht mit dem Algorithmus ACQCM gelöst werden. Da F nicht stark monoton ist, muss die Lösung v^* nicht im quadratischen Schnitt C_{v^j} enthalten sein. Wird die Semilinearität in expliziter Form in die diskrete Variationsungleichung übertragen, so erhält man

$$\left(\left(\frac{M}{\tau} + S \right) y^k - \frac{M}{\tau} y^{k-1} - M Q(y^{k-1}), v - y^k \right) \geq 0 \quad \forall v \in \tilde{K}_n. \quad (4.5)$$

$F(y) = \left(\frac{M}{\tau} + S \right) y - \frac{M}{\tau} y^{k-1} - M Q(y^{k-1})$ ist stark monoton und (4.5) kann mit dem Algorithmus ACQCM numerisch gelöst werden. Für die Berechnungen wurden $h = \frac{1}{48}$, $\tau = 0.001$ und $\epsilon = 0.05$ gewählt. Das Ergebnis ist in der Bildfolge 4.3 dargestellt.

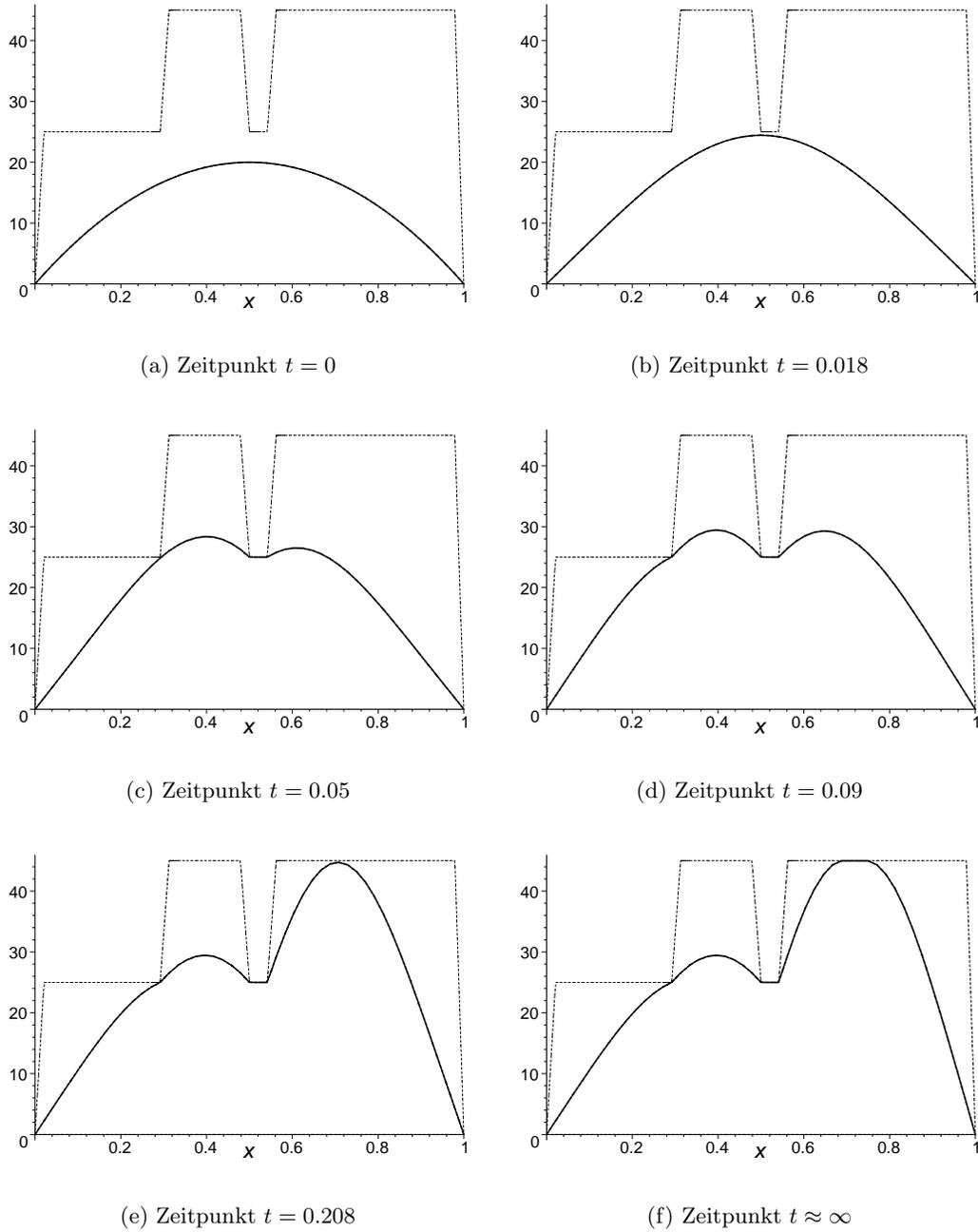


Abbildung 4.3: Semilineares Beispiel mit nicht monotoner Semilinearität

Literaturverzeichnis

- [1] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.
- [2] C. Bandle and H. Brunner. Numerical analysis of semilinear parabolic problems with blow-up solutions. *Rev. Real Acad. Cienc. Exact. Fis. Natur. Madrid*, 88(2-3):203–222, 1994.
- [3] Alain Bensoussan and Jacques-Louis Lions. *Applications of variational inequalities in stochastic control*, volume 12 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1982. Translated from the French.
- [4] Dietrich Braess. *Finite Elemente*. Springer-Verlag, Berlin, 1991.
- [5] Haïm Brézis. Problèmes unilatéraux. *J. Math. Pures Appl. (9)*, 51:1–168, 1972.
- [6] Philippe G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.
- [7] J. N. Dewynne, S. D. Howison, I. Ruf, and P. Wilmott. Some mathematical results in the pricing of American options. *European J. Appl. Math.*, 4(4):381–398, 1993.
- [8] G. Duvaut and J.-L. Lions. *Les inéquations en mécanique et en physique*. Dunod, Paris, 1972. Travaux et Recherches Mathématiques, No. 21.
- [9] Gerald B. Folland. *Real analysis*. John Wiley & Sons Inc., New York, 1984. Modern techniques and their applications, A Wiley-Interscience Publication.
- [10] Isaac Fried. Bounds on the spectral and maximum norms of the finite element stiffness, flexibility and mass matrices. *Internat. J. Solids and Structures*, 9:1013–1034, 1973.
- [11] Jean-Louis Goffin, Patrice Marcotte, and Daoli Zhu. An analytic center cutting plane method for pseudomonotone variational inequalities. *Oper. Res. Lett.*, 20(1):1–6, 1997.
- [12] David Kinderlehrer and Guido Stampacchia. *An introduction to variational inequalities and their applications*, volume 88 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980.
- [13] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod, 1969.

- [14] J.-L. Lions and G. Stampacchia. Variational inequalities. *Comm. Pure Appl. Math.*, 20:493–519, 1967.
- [15] Hans-Jakob Lüthi and Benno Büeler. The analytic center quadratic cut method (acqcm) for strongly monotone variational inequality problems. To appear in *SIAM Journal on Optimization*, 1998.
- [16] Thomas L. Magnanti and Georgia Perakis. A unifying geometric solution framework and complexity analysis for variational inequalities. *Math. Programming*, 71(3, Ser. A):327–351, 1995.
- [17] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.
- [18] Reinhard Scholz. Numerical solution of the obstacle problem by the penalty method. *Computing*, 32(4):297–306, 1984.
- [19] Reinhard Scholz. Numerical solution of the obstacle problem by the penalty method. II. Time-dependent problems. *Numer. Math.*, 49(2-3):255–268, 1986.
- [20] Larry L. Schumaker. *Spline functions: basic theory*. John Wiley & Sons Inc., New York, 1981. Pure and Applied Mathematics, A Wiley-Interscience Publication.
- [21] Gy. Sonnevend. An “analytical centre” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In *System modelling and optimization (Budapest, 1985)*, pages 866–875. Springer, Berlin, 1986.