Anfangsrandwertaufgaben für lineare parabolische partielle Differentialgleichungen

Prof. Dr. Peter Knabner

Friedrich–Alexander–Universität
Erlangen–Nürnberg
Wintersemester 94/95

Ausarbeitung: Bodo Igler, Holger Kappmeier, Eckhard Schneid

Realisierung in TEX: Alexander Eismann, Gerhard Summ

Inhaltsverzeichnis

0 Notation und Gegenstand der Vorlesung

Gegenstand der Vorlesung sind Anfangsrandwertaufgaben (ARWA) für parabolische partielle Differentialgleichungen (PDG oder PDE). Die auftretenden Differentialoperatoren werden i.a. linear sein.
Definitionsgebiet dieser PDG ist ein Raum–Zeit–Zylinder Q_{T_1,T_2}.

Bezeichnungen: $\Omega \subset \mathbf{R}^n$ sei ein, nicht notwendig beschränktes, Gebiet.

$$\begin{aligned} Q_{T_1,T_2} &:= \Omega \times (T_1, T_2\,] \\ Q_T &:= Q_{0,T} \\ S &:= \partial\Omega \\ S_{T_1,T_2} &:= S \times (T_1, T_2\,] \\ S_T &:= S_{0,T} \end{aligned} \tag{0.1}$$

Herkunft der parabolischen PDG:
Sie dienen dazu, den Erhalt einer extensiven Größe (z.B. Masse, Energie, Impuls) im Zeitverlauf zu beschreiben und haben folgende Gestalt:

$$\partial_t(\alpha(u)) + \operatorname{div}(J(u)) = \tilde{f}(u) \quad \text{auf } Q_T \tag{0.2}$$

u:	zur jeweiligen extensiven Größe zugehörige intensive Größe
$\alpha(u)$:	Dichte
$J(u)$:	Stromdichte
$\tilde{f}(u)$:	verteilte Quelldichte (beschreibt Zu-/Abfluß)

Die Stromdichte läßt sich in Diffusion und Konvektion aufsplitten:

$$J(u) = \underbrace{-K(x,t)\nabla u}_{\text{Diffusion}} + \underbrace{c(x,t)u}_{\text{Konvektion}}$$

Desweiteren: $\tilde{f}(u) = -d(x,t)u + f(x,t,u)$
(0.2) wird zu:

$$\partial_t(\alpha(u)) + \operatorname{div}(-K(x,t)\nabla u + c(x,t)u) + d(x,t)u = f(x,t,u) \tag{0.3}$$

$$\text{wobei:} \quad \begin{aligned} &K : Q_T \to \{K \in \mathbf{R}^{n,n} | K \text{ symm., glm. pos. def.}\} \\ &c : Q_T \to \mathbf{R}^n \\ &d : Q_T \to \mathbf{R} \end{aligned}$$

Bezeichnungen: (0.3) heißt *gleichmäßig parabolisch*, falls

$$\alpha_1 \geq \alpha(u) \geq \alpha_0 > 0 \quad \forall u \in \mathbf{R}$$

(0.3) hat einen *linearen glm. parabolischen Differentialoperator*, falls

$$\alpha(u) = \alpha_0 u\,; \quad \alpha_0 > 0 \tag{0.4}$$

(o.B.d.A. $\alpha_0 = 1$ in diesem Fall)

Bemerkungen:

1. (0.3) mit (0.4) ergibt eine semilineare parabolische PDG.
2. (0.3) mit (0.4) und $f(x,t,u) = f(x,t)$ ergibt eine lineare parabolische PDG.
3. Der Fall $\alpha'(u) = 0$ für gewisse $u \in \mathbf{R}$ hat einen Übergang vom parabolischen zum elliptischen Typ und damit eine degenerierte PDG zur Folge. Dieser Fall wird als „schnelle Diffusion" bezeichnet.
4. Der Fall $\alpha'(u) = \infty$ für gewisse $u \in \mathbf{R}$ hat einen Übergang vom parabolischen zum hyperbolischen Typ zur Folge und wird als „langsame Diffusion" bezeichnet.

Beispiel: Wärmeleitungsgleichung (bzw. Diffusionsgleichung)

$$\partial_t u - \operatorname{div}(K(x,t)\nabla u) = f(x,t)$$

Falls $K(x,t) = k(x,t)E_n$; $k(x,t) \geq k_0 > 0 \quad \forall (x,t) \in Q_T$, schreibt sich dies als:

$$\partial_t u - \operatorname{div}(k(x,t)\nabla u) = f(x,t)$$

Falls zusätzlich $k(x,t) = k_0 \quad \forall (x,t) \in Q_T$:

$$\partial_t u - k_0 \Delta u = f(x,t)$$

Um die Eindeutigkeit von Lösungen zu garantieren, muß (0.3) noch mit geeigneten Anfangsvorgaben und Randbedingungen versehen werden.
Anfangsvorgabe: $u(x,0) = u_0(x) \quad \forall x \in \Omega$.
Randbedingungen auf S_T:

$$\begin{gathered}
S = \Gamma_1 \dot\cup \Gamma_2 \dot\cup \Gamma_3 \\
S_{iT} := \Gamma_i \times (0,T] \\
u(x,t) = g_1(x,t) \text{ auf } S_{1T} \\
J(u)\cdot\eta = g_2(x,t) \text{ auf } S_{2T} \\
J(u)\cdot\eta + \gamma u = g_3(x,t) \text{ auf } S_{3T}
\end{gathered}$$

Hierbei sei η die äußere Einheitsnormale.

1 Die Wärmeleitungsgleichung

1.1 Elementare Lösungen für die Wärmeleitungsgleichung

Ausgangspunkt ist die folgende PDG:

$$\begin{aligned} \partial_t c - D\partial_{xx} c = 0\,; \qquad & x > 0,\ t > 0,\ D > 0 \\ c(0,t) = C_1 \quad & \text{für} \quad t > 0 \\ c(x,0) = C_0 \quad & \text{für} \quad x > 0 \end{aligned} \tag{1.1}$$

Entdimensionalisierung durch

$$u(x,t) = \frac{c(x,t) - C_0}{C_1 - C_0} \tag{1.2}$$

führt zu:

$$\begin{aligned} \partial_t u - D\partial_{xx} u = 0 \\ u(0,t) = 1 \\ u(x,0) = 0 \end{aligned} \tag{1.3}$$

Ziel der folgenden Analyse ist es, selbstähnliche Lösungen zu gewinnen.

$$\text{Ansatz:} \qquad u(x,t) = f(\eta); \quad \eta := \frac{x}{\sqrt{t}} \tag{1.4}$$

$$\Rightarrow \qquad \partial_t u = \partial_\eta f \partial_t \eta = -\frac{1}{2}\frac{x}{t^{3/2}} f' = -\frac{1}{t}\frac{1}{2}\eta f'$$

$$\partial_{xx} u = f''\,(\partial_x \eta)^2 = \frac{1}{t} f''$$

$$\Rightarrow \qquad \frac{1}{2}\eta f' + D f'' = 0 \quad \text{für}\, \eta > 0 \tag{1.5}$$

Für die Randbedingungen ergibt sich:

$$\begin{aligned} x = 0,\ t > 0 \Rightarrow \eta = 0 \quad &: \quad f(0) = 1 \\ x > 0,\ t = 0 \Rightarrow \eta = \infty \quad &: \quad f(\infty) = 0 \end{aligned} \tag{1.6}$$

$$\text{Setze } v := f' \qquad \Rightarrow \qquad 1/2\eta v + D v' = 0$$

$$\Rightarrow \qquad v(\eta) = A \exp\left(-\frac{\eta^2}{4D}\right)$$

$$\Rightarrow \qquad f(\eta) = A \int_0^\eta \exp\left(-\frac{s^2}{4D}\right)\, ds + B; \quad A, B \in \mathbf{R}$$

Aus (1.6) folgt: $B = 1$ und hiermit

$$-1 = A \int_0^\infty \exp\left(-\frac{s^2}{4D}\right) ds = A2\sqrt{D} \underbrace{\int_0^\infty \exp(-\tau^2)\, d\tau}_{\sqrt{\frac{\pi}{4}}}$$

$$\Rightarrow \qquad A = -\frac{1}{2\sqrt{D}\sqrt{\frac{\pi}{4}}} \tag{1.7}$$

Einschub:

Fehlerfunktion: $$\operatorname{erf}(s) := \frac{2}{\sqrt{\pi}} \int_0^s exp\left(-\tau^2\right) d\tau; \quad s \in \mathbf{R} \tag{1.8}$$

Es folgt: $\operatorname{erf}(0) = 0$ und $\operatorname{erf}(\pm\infty) = \pm 1$.
Komplementäre Fehlerfunktion:

$$\operatorname{erfc}(s) := 1 - \operatorname{erf}(s) = \frac{2}{\sqrt{\pi}} \int_s^\infty exp\left(-\tau^2\right) d\tau \tag{1.9}$$

Es folgt: $\operatorname{erfc}(0) = 1$, $\operatorname{erfc}(+\infty) = 0$ und $\operatorname{erfc}(-\infty) = 2$.

$f(\cdot)$ läßt sich durch die komplementäre Fehlerfunktion folgendermaßen ausdrücken:

$$\begin{aligned} f(\eta) &= 1 - \frac{1}{\sqrt{\pi D}} \int_0^\eta \exp\left(-\frac{s^2}{4D}\right) ds = \\ &= 1 - \frac{2}{\sqrt{\pi}} \int_0^{\frac{\eta}{2\sqrt{D}}} \exp\left(-\tau^2\right) d\tau = \operatorname{erfc}\left(\frac{\eta}{2\sqrt{D}}\right) \end{aligned}$$

Diese selbstähnliche Lösung läßt sich in der ursprünglichen Unbekannten $c(x,t)$ wie folgt schreiben:

$$\begin{aligned} c(x,t) &= (C_1 - C_0) \operatorname{erfc}\left(\frac{x}{2\sqrt{Dt}}\right) + C_0 = \qquad (1.10) \\ &= C_1 + (C_0 - C_1) \operatorname{erf}\left(\frac{x}{2\sqrt{Dt}}\right) \qquad \text{für } x > 0,\ t > 0. \end{aligned}$$

Wegen $\partial_t c$, $\partial_{xx} c \in C(Q_\infty)$ handelt es sich bei (1.10) um eine klassische Lösung. Für ihr Verhalten am Rand gilt:

$$\begin{aligned} x > 0 \text{ fest},\ t \searrow 0 &: \ c(x,t) \to C_0 \\ t > 0 \text{ fest},\ x \searrow 0 &: \ c(x,t) \to C_1 \end{aligned}$$

1.2 Ähnlichkeitslösungen auf 'R'

Wie bereits unter 1.1 für $x \in [0, \infty[$ beschrieben, können Lösungen der Wärmeleitungsgleichung mit Hilfe der Transformation $\eta = xt^{-\frac{1}{2}}$ auch für ganz $\mathbf{R}$ konstruiert werden.
Sei folgende ARWA gegeben

$$\begin{aligned}
c: \quad \mathbf{R} \times \mathbf{R}^+ &\to \mathbf{R} \\
(x,t) &\mapsto c(x,t)
\end{aligned}$$

$$x_0 \in \mathbf{R}, \; C_0, C_1 \in \mathbf{R}, \; \text{o.B.d.A. } C_1 > C_0$$

$$\partial_t c = D \partial_{xx} c \quad \text{für } x \in \mathbf{R}, t \in]0, \infty[\tag{1.11}$$

$$c(x,0) = \begin{cases} C_1 & \text{für} \quad x < x_0 \\ C_0 & \phantom{\text{für}} \quad x > x_0 \end{cases} \tag{1.12}$$

Durch die Skalierung (1.2) erhält man

$$\partial_t u = D \partial_{xx} u$$

$$u(x,0) = \begin{cases} 1 & \text{für} \quad x < x_0 \\ 0 & \phantom{\text{für}} \quad x > x_0 \end{cases}$$

Die Transformation $\eta = (x - x_0)t^{-\frac{1}{2}}$, $f(\eta) = u(x,t)$ führt die ARWA (1.11, 1.12) in eine gewöhnliche DGL über:

$$\frac{1}{2}\eta f' + D f'' = 0 \tag{1.13}$$

$$f(-\infty) = 1 \quad f(+\infty) = 0,$$

deren Lösung bereits in 1.1 hergeleitet wurde:

$$f(\eta) = \frac{1}{2} \operatorname{erfc}\left(\frac{\eta}{2D^{\frac{1}{2}}}\right)$$

Durch Rücktransformation erhält man die Lösung der ARWA (1.11, 1.12):

$$c(x,t) = C_0 + \frac{1}{2}(C_1 - C_0) \operatorname{erfc}\left(\frac{x - x_0}{2(Dt)^{\frac{1}{2}}}\right)$$

- $c(x_0, t) = \frac{1}{2}(C_0 + C_1)$
- Die Niveaulinien von c in der x–t–Halbebene haben die Form: $x = x_0 + \text{const}\,(Dt)^{\frac{1}{2}}$.

1.3 Allgemeine Anfangsrandwertaufgabe auf 'R'

Da das betrachtete Problem linear ist, kann durch Superposition von bekannten Lösungen der ARWA eine neue Lösung konstruiert werden. So erhält man für die Randvorgaben

$$c(x,0) = \begin{cases} C_1, & x < -a \\ C_0, & -a < x < a \qquad \text{(Rechteckpuls)} \\ C_1, & x > a \end{cases} \tag{1.14}$$

aus der Superposition

$$c(x,0) = c_u(x,0) + c_d(x,0) \tag{1.15}$$

$$c_u(x,0) = \begin{cases} C_1, & x < -a \\ C_0, & x > -a \end{cases} \tag{1.16}$$

$$c_d(x,0) = \begin{cases} 0, & x < a \\ C_1 - C_0, & x > a \end{cases} \tag{1.17}$$

mit den aus 1.2 bekannten Lösungen für (1.16) und (1.17)

$$\begin{aligned} c_u(x,t) &= C_0 + \frac{1}{2}(C_1 - C_0)\,\mathrm{erfc}\left(\frac{x+a}{2(Dt)^{\frac{1}{2}}}\right) \\ c_d(x,t) &= C_1 - C_0 + \frac{1}{2}(C_1 - C_0)\,\mathrm{erfc}\left(\frac{x-a}{2(Dt)^{\frac{1}{2}}}\right) \end{aligned}$$

die Lösung für die Randvorgabe (1.14)

$$\begin{aligned} c(x,t) &= C_1 + \frac{1}{2}(C_0 - C_1)\left\{\mathrm{erfc}\left(\frac{x-a}{2(Dt)^{\frac{1}{2}}}\right) - \mathrm{erfc}\left(\frac{x+a}{2(Dt)^{\frac{1}{2}}}\right)\right\} \\ &= C_1 + \frac{C_0 - C_1}{\pi^{\frac{1}{2}}} \int_{(x-a)/(2(Dt)^{\frac{1}{2}})}^{(x+a)/(2(Dt)^{\frac{1}{2}})} e^{-s^2}\, ds. \end{aligned} \tag{1.18}$$

Für $C_1 = 0$ und $C_0 = \frac{M}{2a}$ wird

$$\int_{-\infty}^{\infty} c(x,0)\, dx = M$$

und der Grenzfall $a \to 0$ läßt sich dann mit Hilfe der Dirac-Distribution δ schreiben

$$c(x,0) = M\delta(x)$$

Die Lösung für diese Anfangswertvorgabe hat folgende Form:

$$c(x,t) = \frac{M}{2(\pi Dt)^{\frac{1}{2}}} \exp\left\{-\frac{x^2}{4Dt}\right\} \tag{1.19}$$

und wird als Fundamentallösung der Diffusionsgleichung bezeichnet.
Formal läßt sich jede Anfangswertvorgabe c_0 als (kontinuierliche) Superposition von Dirac-Distributionen schreiben:

$$c_0(x) = \int_{-\infty}^{\infty} \delta(y-x)\, c_0(y)\, dy$$

Die Lösung dazu ist die entsprechende Superposition der Fundamentallösungen (1.19)

$$c(x,t) = \frac{1}{2(\pi D t)^{\frac{1}{2}}} \int_{-\infty}^{\infty} c_0(y) \exp\left\{-\frac{(y-x)^2}{4Dt}\right\} dy \quad \text{für } t>0, x \in \mathbf{R}$$

- im Gegensatz zur physikalischen Realität erhält man eine unendliche Ausbreitungsgeschwindigkeit, d.h.:

$$\begin{aligned} c(x,0) &= 0 \quad \text{für } x \notin K,\ K \subset \mathbf{R},\ K \text{ kompakt} \\ c(x,t) &> 0 \quad \text{für alle } t>0, x \in \mathbf{R} \end{aligned}$$

- Vergleichsprinzip:

$$\begin{aligned} & c_1(x,0) \geq c_2(x,0) \quad \forall x \in \mathbf{R} \quad \text{und} \quad \exists x \in \mathbf{R} :\ c_1(x,0) > c_2(x,0) \\ & \Rightarrow c_1(x,t) > c_2(x,t) \quad \forall x \in \mathbf{R}, t > 0 \end{aligned}$$

1.4 Anfangsrandwertaufgabe auf '(0,L)'

Gegeben sei folgende ARWA:

$$\begin{aligned} & \partial_t c = D\partial_{xx} c, \quad x \in (0,L),\ t \in]0,\infty[\\ & c(0,t) = C_l,\ c(L,t) = C_r, \quad t \in]0,\infty[\\ & c(x,0) = c_0(x) \end{aligned} \tag{1.20}$$

Mit folgender Transformation kann die inhomogene *Dirichlet*-Randvorgabe auf eine homogene reduziert werden:

$$\begin{aligned} \overline{c}(x) &= \frac{C_r - C_l}{L} x + C_l \\ u(x,t) &:= c(x,t) - \overline{c}(x) \end{aligned}$$

Damit läßt sich o.B.d.A. (1.20) auch in folgender Form schreiben:

$$\begin{aligned} & \partial_t u = D\partial_{xx} u, \quad x \in (0,L),\ t \in]0,\infty[\\ & u(0,t) = u(L,t) = 0, \quad t \in]0,\infty[\\ & u(x,0) = u_0(x) = c_0(x) - \overline{c}(x), \quad x \in (0,L) \end{aligned}$$

Mit Hilfe eines Produktansatzes für die Lösung

$$u(x,t) = X(x)T(t)$$

separiert die ARWA in zwei gewöhnliche Differentialgleichungen:

$$\begin{aligned}
& \partial_t u = D\partial_{xx} u \\
\Leftrightarrow\quad & XT' = DTX'' \Leftrightarrow \frac{1}{D}\frac{T'}{T} = \frac{X''}{X} := -\alpha \\
\Leftrightarrow\quad & \text{(i) } T' + \alpha DT = 0, \quad t \in]0,\infty[\qquad \text{und} && (1.21) \\
& \text{(ii) } X'' + \alpha X = 0, \quad x \in (0,L),\ X(0) = X(L) = 0 && (1.22)
\end{aligned}$$

Hier wurde als Separationskonstante $-\alpha$ gewählt. Die allgemeine Lösung von (1.21) lautet

$$T(t) = \exp\{-\alpha Dt\}$$

mit $\alpha > 0$, da u für $t \to \infty$ beschränkt bleiben soll. O.B.d.A. kann man jetzt schreiben $\alpha = \lambda^2$.
Die allgemeine Lösung von (1.22) lautet somit

$$X(x) = \eta \sin \lambda x + \xi \cos \lambda x. \tag{1.23}$$

Aus der Randbedingung $X(0) = 0$ folgt $\xi = 0$ und aus $X(L) = 0$ erhält man

$$\lambda \in \left\{ \frac{\pi}{L}, \frac{2\pi}{L}, \ldots, \frac{n\pi}{L}, \ldots \right\},\ n \in \mathbf{N}.$$

Ohne Berücksichtigung der Anfangswertvorgabe hat eine Lösung dann die Form:

$$u_n(x,t) = \sin\left(\frac{n\pi}{L}x\right) \exp\left\{ -\left(\frac{n\pi}{L}\right)^2 Dt \right\} \tag{1.24}$$

Wenn sich die Anfangswertvorgabe als Fourierreihe formulieren läßt

$$u_0(x) = \sum_{n=1}^{N} c_n \sin\left(\frac{n\pi}{L}x\right), \tag{1.25}$$

dann läßt sich die Lösung u als die entsprechende Superposition der speziellen Lösungen u_n (1.24) schreiben:

$$u(x,t) = \sum_{n=1}^{N} c_n \sin\left(\frac{n\pi}{L}x\right) \exp\left\{ -\left(\frac{n\pi}{L}\right)^2 Dt \right\}$$

und die Lösung der ursprünglichen ARWA (1.20) lautet

$$c(x,t) = \overline{c}(x,t) + u(x,t).$$

Insbesondere kann man hier erkennen, daß

$$c(x,t) \to \overline{c}(x) \quad \text{für } t \to \infty,\ x \in (0,L).$$

Für das entsprechende *Neumann*-Problem

$$\begin{aligned}
&\partial_t u = D\partial_{xx} u, \quad x \in (0,L),\ t \in]0,\infty[\\
&\partial_x u(0,t) = \partial_x u(L,t) = 0, \quad t \in]0,\infty[\\
&u(x,0) = u_0(x)
\end{aligned}$$

erhält man durch die gleiche Vorgehensweise bei (1.23) $\eta = 0$ aus $\partial_x X(0) = 0$ und aus $\partial_x X(L) = 0$ folgt

$$\lambda \in \left\{0, \frac{\pi}{L}, \frac{2\pi}{L}, \ldots, \frac{n\pi}{L}, \ldots\right\},\ n \in \mathbf{N}.$$

Läßt sich analog (1.25) die Anfangswertvorgabe als Fourierreihe schreiben

$$u_0(x) = \sum_{n=0}^{N} c_n \cos\left(\frac{n\pi}{L}x\right),$$

dann erhält man die Lösung:

$$u(x,t) = \sum_{n=0}^{N} c_n \cos\left(\frac{n\pi}{L}x\right) \exp\left\{-\left(\frac{n\pi}{L}\right)^2 Dt\right\}$$

Problem: Unklar ist, inwiefern die Anfangswertvorgabe sich als Fourierreihe in sinus bzw. cosinus schreiben läßt.
Bei der Dirichlet-Randbedingung heißt das:

1. Existieren Entwicklungskoeffizienten c_n, so daß gilt:

$$u_0(x) = \sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi}{L}x\right) \quad ?$$

2. Wie sehen die c_n aus?

Nimmt man an, daß die Entwicklungskoeffizienten existieren und

$$u(x,t) = \sum_{n=1}^{\infty} c_n \sin\left(\frac{n\pi}{L}x\right) \exp\left\{-\left(\frac{n\pi}{L}\right)^2 Dt\right\},$$

welche Konvergenzeigenschaften müssen dann gelten, damit u eine klassische Lösung ist?

Einschub: Gleichmäßige Konvergenz von Funktionenreihen
Sei $\Omega \subset \mathbf{R}^n$ beschränkt, $\|f\|_\infty := \sup\left\{|f(x)| : x \in \bar{\Omega}\right\}$, $f_n \in C(\Omega)$, Partialsumme: $s_n(x) = \sum_{k=1}^n f_k$.

1. Wenn die Folge der Partialsummen s_n gleichmäßig gegen eine Funktion f konvergiert, dann ist f stetig auf $\bar{\Omega}$.

$$\lim_{n\to\infty} s_n = f \Rightarrow f \in C(\bar{\Omega}),\ f = \sum_{n=1}^{\infty} f_n$$

2. $$\sum_{n=1}^{\infty} \|f_n\|_\infty < \infty \Rightarrow \lim_{n\to\infty} s_n = f \in C(\bar{\Omega})$$

   Bei gleichm. Konvergenz können Integration und Summation vertauscht werden.

3. $$f = \sum_{n=1}^{\infty} f_n,\ f \in C(\bar{\Omega}) \Rightarrow \sum_{n=1}^{\infty} \int_\Omega f_n(x)\,dx = \int_\Omega \sum_{n=1}^{\infty} f_n(x)\,dx$$

4. $$\left.\begin{array}{l} f_n \in C^1(\bar{\Omega}),\ \lim\limits_{n\to\infty} s_n = f \in C(\bar{\Omega}) \\ \lim\limits_{n\to\infty} \partial_{x_i} s_n = \lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \partial_{x_i} f_k = g \in C(\bar{\Omega}) \end{array}\right\} \Rightarrow \partial_{x_i} f = g$$

Für das Dirichletproblem betrachten wir den Ansatz:

$$u^D(x,t) := \sum_{n=1}^{\infty} \alpha_n^D u_n^D(x,t) \quad \text{mit} \quad u_n^D(x,t) := e^{-\left(\frac{n\pi}{L}\right)^2 Dt} \sin\left(\frac{n\pi}{L}x\right)$$

und für das Neumann-Problem:

$$u^N(x,t) := \sum_{n=0}^{\infty} \alpha_n^N u_n^N(x,t) \quad \text{mit} \quad u_n^N(x,t) := e^{-\left(\frac{n\pi}{L}\right)^2 Dt} \cos\left(\frac{n\pi}{L}x\right)$$

Die Anfangsbedingung ist erfüllt, wenn die Fourierreihe u^D bzw. u^N gleichmäßig konvergiert (dann ist auch u^D bzw. $u^N \in C(\bar{Q}_T)$) und ferner gilt:

$$\begin{array}{rll} & u_0(x) = u^D(x,0) = \sum_{n=1}^{\infty} \alpha_n^D u_n^D(x,t) & \text{für alle} \quad x \in \Omega \\ \text{bzw.} & u_0(x) = u^N(x,0) = \sum_{n=0}^{\infty} \alpha_n^N u_n^N(x,t) & \text{für alle} \quad x \in \Omega. \end{array}$$

Aus der Konvergenz der Fourierreihe folgt dann auch sofort, daß die Dirichletbedingung gilt, da

$$u_n^D(0,t) = u_n^D(L,t) = 0 \qquad \text{für alle} \quad t \in (0,T].$$

Für die Neumannbedingung muß noch die Konvergenz der ersten Ableitungen im Ort betrachtet werden: Falls die Reihe $\sum_{n=0}^{\infty} \alpha_n^N \partial_x u_n(x,t)$ gleichmäßig konvergiert, so existiert die Ableitung $\partial_x u^N$ (siehe Einschub (4)) und es gilt:

$$\begin{array}{lll} & \partial_{x_i} u^N = \sum_{n=0}^{\infty} \alpha_n^N \partial_{x_i} u_n^N & (i = 1, \ldots, n), \\ \text{also:} & \partial_{x_i} u^N(0,t) = \partial_{x_i} u^N(L,t) = 0 & \text{für alle} \quad t \in (0, T\,]. \end{array}$$

Soll der Ansatz nun in die ursprüngliche PDG eingesetzt werden, so ist es erforderlich, daß auch die zweite Ableitung im Ort, sowie die erste in der Zeit existieren, d.h.:

$$\begin{aligned} \sum_{n=\substack{1\\0}}^{\infty} \alpha_n^{\overset{D}{N}} \partial_{xx} u_n^{\overset{D}{N}} \quad &\in \quad C(\bar{Q}_T) \\ \sum_{n=\substack{1\\0}}^{\infty} \alpha_n^{\overset{D}{N}} \partial_t u_n^{\overset{D}{N}} \quad &\in \quad C(\bar{Q}_T) \end{aligned}$$

Die einzelnen Terme können wie folgt abgeschätzt werden:

$$\begin{aligned} \left| \alpha_n^{\overset{D}{N}} u_n^{\overset{D}{N}} \right| \quad &\leq \quad \left| \alpha_n^{\overset{D}{N}} \right| \cdot \left| e^{-\lambda_n^2 D t} \right| \cdot \left| \begin{matrix} \sin \\ \cos \end{matrix} \lambda_n x \right| \leq |\alpha_n| \\ \left| \alpha_n^{\overset{D}{N}} \frac{\partial^k}{\partial x^k} u_n^{\overset{D}{N}} \right| \quad &\leq \quad \left| \alpha_n^{\overset{D}{N}} \right| \cdot \lambda_n^k \\ \left| \alpha_n^{\overset{D}{N}} \frac{\partial}{\partial_t} u_n^{\overset{D}{N}}(x,t) \right| \quad &\leq \quad \left| \alpha_n^{\overset{D}{N}} \right| D \lambda n^2 \qquad \text{mit} \quad \lambda_n := \frac{n\pi}{L} \end{aligned}$$

Wenn wir also eine klassische Lösung erhalten wollen, ist gemäß Einschub (2) die Bedingung

$$\sum_{n=\substack{1\\0}}^{\infty} \left| \alpha_n^{\overset{D}{N}} \right| \lambda_n^2$$

hinreichend. Für klassische Lösungen auf $Q_{\varepsilon,T}$ für ein $\varepsilon > 0$ ist die Wachstumsbedingung an α_n wesentlich geringer, nämlich:

$$\sum_{n=\substack{1\\0}}^{\infty} \lambda_n^2 e^{-\lambda_n^2 D \varepsilon} \left| \alpha_n^{\overset{D}{N}} \right| .$$

Da der Faktor $e^{-\lambda_n^2 D\varepsilon}$ stark dämpft für große λ_n, kann in diesem Fall mit weitaus schwächeren Bedingungen für $\left\{ \alpha_n^{\overset{D}{N}} \right\}_{n=\substack{1,2,\ldots\\0,1,\ldots}}$ gerechnet werden.

1.5 Fourierreihen mit sin und cos

Die Familie der Funktionen

$$\frac{1}{\sqrt{2L}},\ \frac{1}{\sqrt{L}}\sin(\lambda_n x)\ ,\ \frac{1}{\sqrt{L}}\cos(\lambda_n x) \quad \text{für } n \in \mathbf{N}$$

liegt in $C(-L, L)$ und bildet ein Orthonormalsystem bzgl.

$$\langle f, g\rangle := \int_{-L}^{L} f(t)g(t)\,dt, \tag{1.26}$$

d.h. insbesondere für $s_N := \frac{1}{2}a_0 + \sum_{n=1}^{N} a_n \cos(\lambda_n x) + b_n \sin(\lambda_n x)$ für $n \in \mathbf{N}$ gilt:

$$\begin{aligned} a_n &= \frac{1}{L}\int_{-L}^{L} s_N(x)\cos(\lambda_n x)\,dx \\ b_n &= \frac{1}{L}\int_{-L}^{L} s_N(x)\sin(\lambda_n x)\,dx \\ a_0 &= \frac{1}{L}\int_{-L}^{L} s_N(x)\cdot 1\,dx. \end{aligned} \tag{1.27}$$

Für $s(x) := \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(\lambda_n x) + b_n \sin(\lambda_n x)$ gilt (1.27), wenn die Integral- und Summenbildung vertauschbar ist (i.Z.: $\sum_{n=1}^{\infty}\int_{-L}^{L} = \int_{-L}^{L}\sum_{n=1}^{\infty}$).
Eine reine sin– bzw. cos–Entwicklung auf $[0, L]$ kann durch ungerade bzw. gerade Fortsetzung der Funktion $f = \sum \substack{\sin \\ \cos}$ auf $[-L, L]$ erweitert werden. Ist nämlich $f(x) = f(-x)$ gerade, dann gilt: $b_n = 0$ für alle $n \in \mathbf{N}$; und analog gilt für $f(x) = -f(-x)$ ungerade: $a_n = 0$ für alle $n \in \mathbf{N}_0$.

1.6 Allgemeine Diffusionsgleichungen

Für allgemeine Lineare Diffusionsgleichungen kann die bisher dargestellte Lösungstheorie z.T. verwendet werden und zwar in den Fällen, in denen sich eine „allgemeine“ Gleichung geeignet auf den Spezialfall $\partial_t u = D\partial_{xx}u$ transformieren läßt.

1. *Lineare Diffusions-Konvektionsgleichung*

$$\partial_t - D\partial_{xx}u + q\partial_x u = 0, \qquad D, q \in \mathbf{R}, \quad D > 0,\ q \neq 0 \tag{1.28}$$

Für (1.28) betrachten wir zwei Transformationen:

(a)

$$\begin{aligned} w(x,t) &:= u(x,t)\exp\left(\frac{q^2}{4D}t - \frac{qx}{2D}\right) \\ \Rightarrow \qquad \partial_t w &= D\partial_{xx} w \end{aligned}$$

(b)
$$v(x,t) := u(x+qt,t)$$

$$\begin{aligned}\Rightarrow \quad \partial_t v(x,t) &= \partial_t u(x+qt,t) + q\partial_x u(x+qt,t)\\ &= D\partial_{xx} u(x+qt,t) - q\partial_x u(x+qt,t) + q\partial_x u(x+qt,t)\\ &= D\partial_{xx} v\end{aligned}$$

2. *Lineare Diffusionsgleichung mit Linearem Zerfall/Abbau*

$$\partial_t u - D\partial_x x u + ku = 0, \qquad k \in \mathbf{R}$$

Mit dem Ansatz
$$v(x,t) := e^{kt} u(x,t)$$
erhalten wir:
$$\partial_t v = D\partial_{xx} v$$

Für weitere Beispiele siehe z.B.:
J.R. Cannon, *The one-dimensional heat equation*, Addison-Wesley 1984, S. 15f.

2 Maximumprinzipien für lineare parabolische Anfangsrandwertaufgaben

$\Omega \subset \mathbf{R}^n$ sei ein beschränktes Gebiet. Wir betrachten:

$$Lu := \partial_t u + \tilde{L}u := \partial_t u - a_{ij}\frac{\partial^2 u}{\partial_{x_i}\partial_{x_j}} + b_i\frac{\partial u}{\partial_{x_i}} + cu \tag{2.1}$$

mit: a_{ij}, b_i, $c \in C(\bar{Q}_T)$ für alle $i, j = 1, \ldots, n$, $\tilde{L}$ elliptisch, d.h. $(a_{ij})_{ij}$ ist symmetrisch und positiv definit auf $\bar{Q}_T$.
Insbesondere existieren dann eine gleichmäßige untere und obere Schranke für $(a_{ij})_{ij}$. Auch im folgenden betrachten wir Funktionen u, die als klassische Lösungen geeignet sind, also

$$u \in C^{2,1}(Q_T) \cap C(\bar{Q}_T)$$

und $\quad Lu = f$ in Q_T für ein $f \in C(\bar{Q}_T)$.

2.1 Das schwache Maximumprinzip

Satz 2.1 (schwaches Maximumprinzip) *Sei $Lu \le 0$ ($Lu \ge 0$), $c = 0$ in Q_T, dann gilt:*

$$\max_{\bar{Q}_T} u = \max_{\Sigma} u \qquad (\min_{\bar{Q}_T} u = \min_{\Sigma} u)$$

für $\Sigma := S_T \cup (\Omega \times \{0\})$. Σ heißt parabolischer Rand.

Beweis: Betrachten wir zunächst den Fall $Lu < 0$ in Q_T, dann gilt: $u(x,t) < \max_{\bar{Q}_T} u$ für alle $(x,t) \in Q_T$. Falls nämlich ein Paar $(\overline{x}, \overline{t}) \in Q_T$ mit $u(\overline{x}, \overline{t}) = \max_{\bar{Q}_T} u$ existiert, dann muß gelten:

$$\begin{aligned} & \partial_{x_i} u(\overline{x}, \overline{t}) = 0 \quad \text{für alle } i = 1, \ldots, n, \\ & \left(\partial_{x_i}\partial_{x_j} u(\overline{x}, \overline{t})\right)_{ij} \quad \text{ist negativ semidefinit} \\ \text{und} \qquad & \partial_t u(\overline{x}, \overline{t}) \ge 0. \end{aligned}$$

Da $\left(\partial_{x_i}\partial_{x_j} u(\overline{x}, \overline{t})\right)$ negativ semidefinit und (a_{ij}) positiv definit ist, erhält man

$$-a_{ij}\partial_{x_i}\partial_{x_j} u(\overline{x}, \overline{t}) \ge 0$$

und daher

$$Lu(\overline{x}, \overline{t}) = \partial_t u(\overline{x}, \overline{t}) - \left(a_{ij}\partial_{x_i}\partial_{x_j}\right) u(\overline{x}, \overline{t}) \ge 0.$$

Das ist aber ein Widerspruch zur Annahme, also ist die Behauptung $u(x,t) < \max_{\bar{Q}_T} u$ für alle $(x,t) \in Q_T$ richtig.
Für den Fall $Lu \le 0$ betrachte man

$$u_\varepsilon := u + \varepsilon e^{-t} \quad \text{für} \quad \varepsilon > 0 \tag{2.2}$$

$$\begin{array}{ll} \text{Dann gilt:} & Lu_\varepsilon = Lu - \varepsilon e^{-t} < 0 \quad \text{in } Q_T \\ \text{und somit} & \max\limits_{\bar{Q}_T} u_\varepsilon = \max\limits_{\Sigma} u_\varepsilon. \end{array}$$

Ferner: $\lim_{\varepsilon \to 0} ||u - u_\varepsilon||_\infty = 0$, also auch punktweise Konvergenz von u_ε gegen u. Da $u_\varepsilon \geq u$ für alle $\varepsilon > 0$ erhält man nun:

$$\max_{\bar{Q}_T} u \leq \max_{\bar{Q}_T} u_\varepsilon = \max_{\Sigma} u_\varepsilon = u_\varepsilon\left(\overline{x}_\varepsilon, \overline{t}_\varepsilon\right)$$

für mindestens ein Paar $\left(\overline{x}_\varepsilon, \overline{t}_\varepsilon\right) \in \Sigma$.
Da Σ kompakt ist, existiert eine Teilfolge $\left\{\left(\overline{x}_{\varepsilon_n}, \overline{t}_{\varepsilon_n}\right)\right\}_{n \in \mathbf{N}}$, die in Σ gegen $\left(\overline{x}, \overline{t}\right)$ konvergiert. Das Anwenden der Dreiecksungleichung ergibt:

$$\left|u\left(\overline{x}, \overline{t}\right) - u_{\varepsilon_n}\left(\overline{x}_{\varepsilon_n}, \overline{t}_{\varepsilon_n}\right)\right| \leq \left|u\left(\overline{x}, \overline{t}\right) - u\left(\overline{x}_{\varepsilon_n}, \overline{t}_{\varepsilon_n}\right)\right| + \left|u\left(\overline{x}_{\varepsilon_n}, \overline{t}_{\varepsilon_n}\right) - u_{\varepsilon_n}\left(\overline{x}_{\varepsilon_n}, \overline{t}_{\varepsilon_n}\right)\right|$$

Aus der Stetigkeit von u und der punktweisen Konvergenz von u_{ε_n} gegen u folgt:

$$\left|u\left(\overline{x}, \overline{t}\right) - u_{\varepsilon_n}\left(\overline{x}_{\varepsilon_n}, \overline{t}_{\varepsilon_n}\right)\right| \stackrel{\varepsilon_n \to 0}{\longrightarrow} 0$$

und daher

$$\max_{\bar{Q}_T} u \leq u\left(\overline{x}, \overline{t}\right) \leq \max_{\Sigma} u.$$

Den Fall $Lu \geq 0$ beweist man durch Vorzeichenvertauschung. $\square$

Korollar 2.2 *Sei $c \geq 0$ in Q_T und $Lu \leq 0$ ($Lu \geq 0$).*

$$\begin{array}{ll} \textit{Dann gilt:} & 1)\ \max\limits_{\bar{Q}_T} u \leq \max\limits_{\Sigma} u^+ \\ & 2)\ Lu = 0 \quad \Rightarrow \quad \max\limits_{\bar{Q}_T} |u| = \max\limits_{\Sigma} |u| \end{array}$$

Korollar 2.3 (Vergleichsprinzip, Eindeutigkeit des Dirichletproblems)
Seien $u, v \in C^{2,1}(Q_T) \cap C(\bar{Q}_T)$.

$$\begin{array}{lll} 1) & Lu \leq Lv & \textit{in } Q_T \textit{ und } u \leq v \textit{ auf } \Sigma \\ & \Rightarrow & u \leq v \textit{ auf } Q_T \\ 2) & Lu = Lv & \textit{in } Q_T \textit{ und } u = v \textit{ auf } \Sigma \\ & \Rightarrow & u = v \textit{ auf } Q_T \end{array}$$

Beweis: 2) folgt aus 1), da: $Lu = Lv$ in Q_T, $u = v$ auf Σ

$$\begin{array}{llll} \Leftrightarrow & Lu \leq Lv,\ Lu \geq Lv \text{ in } Q_T; & u \leq v,\ u \geq v \text{ auf } \Sigma & \\ \stackrel{1)}{\Rightarrow} & u \leq v,\ u \geq v \text{ auf } Q_T & \Leftrightarrow & u = v \quad \text{auf} Q_T \end{array}$$

zu 1) Definiere $\overline{u} := ue^{-\gamma t}$, dann: $u = \overline{u}e^{\gamma t}$ und

$$Lu = e^{\gamma t} L\overline{u} + \gamma \overline{u} = e^{\gamma t} \tilde{L}\overline{u}$$

für $\tilde{L} := L + \gamma$.
Betrachte nun ein $\gamma \in \mathbf{R}_+$ mit $\gamma \geq -c$ auf $\bar{Q}_T$ und setze auch für v: $\overline{v} := ve^{-\gamma t}$, dann folgt aus $L(u-v) \leq 0$:

$$e^{\gamma t}\tilde{L}(\overline{u} - \overline{v}) \leq 0$$

$$\text{und auch} \qquad \tilde{L}(\overline{u} - \overline{v}) \leq 0.$$

Korollar 2.2 ergibt dann

$$\max_{\bar{Q}_T}(\overline{u} - \overline{v}) \leq \max_{\Sigma}(\overline{u} - \overline{v})^+ = 0,$$

da $\overline{u} - \overline{v} = u - v = 0$ auf Σ.
Also gilt $\overline{u} \leq \overline{v}$ in Q_T und somit $u \leq v$ in Q_T. $\square$

2.2 Das starke Maximumprinzip

Satz 2.4 (starkes Maximumprinzip) *Sei $Lu \leq 0$ ($Lu \geq 0$) in Q_T, $M := \max_{\bar{Q}_T} u$ ($M := \min_{\bar{Q}_T} u$).*
Wenn dann ein $(x_0, t_0) \in Q_T$ mit $u(x_0, t_0) = M$ existiert und eine der folgenden Voraussetzungen gilt

1. $c = 0$
2. $c \geq 0$, $M \geq 0$ ($M \leq 0$)
3. $M = 0$

so folgt $u = M$ in Q_{t_0}.

Um den Satz zu beweisen, werden zunächst drei Lemmata bereitgestellt.

Lemma 2.5 *Sei $B \subset \mathbf{R}^{n+1}$ eine Kugel mit $\bar{B} \subset Q_T$ und*

$$u(x,t) < M \quad \forall (x,t) \in B$$

und

$$u(x_0, t_0) = M \quad \text{für ein } (x_0, t_0) \in \partial B$$

Dann ist t_0 entweder der größte oder der kleinste Wert, den t in B annimmt.

Lemma 2.6 *Sei $u(x_0, t_0) < M$ für $x_0 \in \Omega$, $t_0 \in (0, T)$. Dann folgt*

$$u(x, t_0) < M \quad \forall x \in \Omega$$

Beweis: Annahme: In $\Omega \times \{t_0\}$ existieren Punkte mit $u = M$.
$\Rightarrow \exists x_1, x_2 \in \Omega : (u(x_1, t_0) = M) \wedge (u(x_2, t_0) < M)$
Bezeichne $L =< x_1, x_2 > \subseteq \Omega$ einen Weg zwischen x_1 und x_2. Da u stetig ist, kann x_1 o.B.d.A. so gewählt werden, daß:

$$u(x, t_0) < M \quad \forall x \in L$$

Da Ω offen ist, kann L o.B.d.A. als Strecke gewählt werden. Ferner sei definiert:

$$\delta := \min\{|x_1 - x_2|, \operatorname{dist}(L, \partial\Omega), t_0, T - t_0\} \tag{2.3}$$

Sei $x \in L$ mit $0 < |x - x_1| < \delta$. Da der Weg L ganz in Ω liegt, existiert x. Ferner: $d(x) := \operatorname{dist}((x, t_0), \operatorname{int}(Q_T) \cap \{u = M\})$. Es folgt: $d(x) \leq |x - x_1| < \delta$. Sei $B := \left\{(y, t) \,|\, (\|x - y\|^2 + |t - t_0|^2)^{\frac{1}{2}} < d(x)\right\}$. Wegen $B \subseteq \operatorname{int}(Q_T) \cap \{u < M\}$ folgt nach Lemma 2.5:

$$[(y, t) \in \operatorname{int}(Q_T) \cap \{u = M\} \cap \partial B] \Rightarrow (y = x) \wedge ((t = t_0 + d(x)) \vee (t = t_0 - d(x)))$$

Wegen (2.3) und der Wahl von x folgt $\operatorname{int}(Q_T) \cap \{u = M\} \cap \partial B \neq \emptyset$ und somit:

$$u(t_0 + d(x), x) = M \quad \vee \quad u(t_0 - d(x), x) = M \tag{2.4}$$

Sei nun n der Einheitsvektor in Richtung L. Sei $|\varepsilon|$ so gewählt, daß $(x + \varepsilon n) \in L$ und $|x + \varepsilon n - x_1| < \delta$. Aus Lemma 2.5 folgt analog zu (2.4):

$$u(t_0 + d(x + \varepsilon n), x + \varepsilon n) = M \quad \vee \quad u(t_0 - d(x + \varepsilon n), x + \varepsilon n) = M \tag{2.5}$$

O.B.d.A.

$$u(t_0 + d(x + \varepsilon n), x + \varepsilon n) = M \quad \wedge \quad u(t_0 + d(x), x) = M$$

Nach dem Satz von Pythagoras gilt für die beiden rechtwinkligen Dreiecke $\{(x, t_0), (x, t_0 + d(x)), (x + \varepsilon n, t_0)\}$ und $\{(x + \varepsilon n, t_0), (x + \varepsilon n, t_0 + d(x + \varepsilon n)), (x, t_0)\}$:

$$\begin{aligned} d(x + \varepsilon n)^2 + \varepsilon^2 &= H^2 \\ d(x)^2 + \varepsilon^2 &= h^2 \end{aligned}$$

wobei:

$$\begin{aligned} H &:= |(x + \varepsilon n, t_0 + d(x + \varepsilon n)) - (x, t_0)| \\ h &:= |(x, t_0 + d(x)) - (x + \varepsilon n, t_0)| \end{aligned}$$

Es folgt:

$$\begin{gathered} d(x)^2 + \varepsilon^2 \geq d(x + \varepsilon n)^2 \\ d(x)^2 \leq \varepsilon^2 + d(x + \varepsilon n)^2, \end{gathered}$$

da wegen $(x+\varepsilon n, t_0+d(x+\varepsilon n)) \in \mathrm{int}(Q_T) \cap \{u=M\}$, $(x, t_0+d(x)) \in \mathrm{int}(Q_T) \cap \{u=M\}$ $H \geq d(x)$, $h \geq d(x+\varepsilon n)$ gilt. Also:

$$\begin{aligned}
& -\varepsilon^2 \leq d(x+\varepsilon n)^2 - d(x)^2 \leq \varepsilon^2 \\
\Rightarrow \quad & -\varepsilon \leq \frac{d(x+\varepsilon n)-d(x)}{\varepsilon}(d(x+\varepsilon n)+d(x)) \leq \varepsilon \\
\Rightarrow \quad & \lim_{\varepsilon \to 0} \frac{d(x+\varepsilon n)-d(x)}{\varepsilon} = 0 \\
\text{d.h.} \quad & \frac{\partial}{\partial \varepsilon} d(x+\varepsilon n) \mid_{\varepsilon=0} = 0
\end{aligned}$$

Da $x \in L \cap \{0 < |x-x_1| < \delta\}$ beliebig ist, ist d auf dieser Menge konstant. Für $x \to x_1$: $d(x) \to 0$, Widerspruch zu $d(x) \neq 0$. $\square$

Lemma 2.7 *Sei* $0 \leq t_0 \leq t_1 \leq T$ *mit* $u(x,t) < M \quad \forall (x,t) \in \mathrm{int}(Q_{t_1,t_2})$.
Dann ist auch $u(x,t_1) < M \quad \forall x \in \mathbf{R}$.

Beweis: Annahme: $\exists x_1 \in \Omega : u(x_1,t_1) = M$. Sei

$$v(x,t) := \exp(-|x-x_1|^2 - \alpha(t-t_1)) - 1 \quad \text{für großes } \alpha > 0$$

$$\begin{aligned}
\Rightarrow Lv \quad = \quad & \exp(-|x-x_1|^2 - \alpha(t-t_1)) \cdot \\
& \cdot(-4a_{ij}(x_i - x_{1_i})(x_j - x_{1_j}) - 2(-a_{ii} + b_i(x_i - x_{1_i})) - \alpha) + cv
\end{aligned}$$

Für hinreichend großes α existiert dann eine Umgebung N von (x_1,t_1) mit

$$Lv < 0 \quad \text{in } N$$

Sei nun $A \subset Q_T$ als eine durch $|x-x_1|^2 + \alpha(t-t_1) = 0$ und durch eine ε-Umgebung $U_\varepsilon(x_1,t_1)$ berandete Menge gewählt.
Für hinreichend kleines ε gilt:

$$\begin{aligned}
& u + \varepsilon v - M \leq 0 \quad \text{auf } \partial A \\
\wedge \quad & L(u+\varepsilon v - M) \leq -cM \leq 0 \quad \text{in } A
\end{aligned}$$

1. Fall: $c \geq 0$
Nach Korollar 2.2 folgt:

$$u + \varepsilon v - M \leq 0 \quad \text{in } A$$

Da $(u+\varepsilon v - M)(x_1,t_1) = 0$, $u + \varepsilon v - M$ beim Randpunkt (bzgl. A) (x_1,t_1) also sein Maximum annimmt, muß gelten:

$$\begin{aligned}
& 0 \leq (\partial_t u + \varepsilon \partial_t v - \partial_t M)(x_1,t_1) = \partial_t u(x_1,t_1) - \varepsilon\alpha \\
\Rightarrow \quad & \partial_t u(x_1,t_1) > 0 \qquad (2.6)
\end{aligned}$$

Wegen $Lu \leq 0$ und da (x_1, t_1) Maximum bzgl. x ist, gilt andererseits:

$$0 \geq L(u) \geq (\partial_t u + cu) \geq \partial_t u(x_1, t_1)$$

Dies ist ein Widerspruch zu (2.6).
2. Fall: $M = 0$
Führe den Beweis für $\tilde{L} := (L - c_-)$ analog durch. $\square$

Beweis von Satz 2.4: Nach Lemma 2.6 folgt für $t > 0$:

$$(\, u(x,t) = m \quad \forall x \in \Omega\,) \quad \vee \quad (\, u(x,t) < M \quad \forall x \in \Omega\,)$$

Da $\mathbf{Q}$ dicht in $\mathbf{R}$ liegt und u stetig ist, ist

$$Z := \{t > 0 \,|\, u(x,t) < M \quad \forall x \in \Omega\}$$

eine Vereinigung von abzählbar vielen Intervallen.
Nach Lemma 2.7 enthält jedes dieser Intervalle seinen rechten Endpunkt. Aufgrund der Stetigkeit von u folgt somit:

$$Z = (t_0, T] \quad \text{für ein } t_0 \geq 0 \qquad \square$$

Bemerkung: Sei $\partial\Omega \in C^2$, $u \in C^1(\Omega)$ und u sei weiter C^1 bis in den Rand hinein.
Nach dem Lemma von Hopf (elliptischer Fall) gilt dann unter den Voraussetzungen von Satz 2.4:

$$\begin{aligned}&[x_0 \in \partial\Omega \,\wedge\, t \in (0,T) \,\wedge\, u(x_0,t) = M \,\wedge\, u(x,t) < M \,\forall x \in \Omega]\\ &\Rightarrow \partial_\eta u(x_0,t) > 0\end{aligned}$$

Ferner gilt für $\partial\Omega \in C^2$, $u, v \in C^2(Q_T) \cap C^1(\bar{Q}_T)$ das folgende Vergleichsprinzip:

$$\begin{aligned}&\left[Lu \leq Lv \text{ auf } Q_T \,\wedge\, u(x,0) \geq v(x,0) \,\forall x \in \Omega \,\wedge\, \frac{\partial u}{\partial \eta} \geq \frac{\partial v}{\partial \eta}\right]\\ &\Rightarrow \quad u \geq v \text{ in } Q_T\end{aligned}$$

3 Eigenwertprobleme und Elliptische RWA 2.Ordnung

3.1 Exkurs: Spektraltheorie

Seien V, H Hilberträume mit

$$V \hookrightarrow H \quad \text{(d.h. } V \text{ stetig eingebettet in } H\text{)} \tag{3.1}$$

$$V \text{ dicht in } H \tag{3.2}$$

Sei ferner a eine stetige Bilinearform auf $V \times V$, symmetrisch und V–elliptisch (d.h. $a(v,v) \geq \alpha \|v\|^2$). $\langle \cdot\,, \cdot \rangle$ bezeichne das Skalarprodukt auf H und $|\cdot|$ die zugehörige Norm. $\|\cdot\|$ sei die Norm auf V.
Betrachtet wird das folgende Eigenwertproblem:
Finde $\lambda \in \mathbf{R}\,(\mathbf{C})$, $u \in V$, $|u| = 1$, so daß

$$a(u,v) = \lambda \langle u, v \rangle \quad \forall v \in V \tag{3.3}$$

Beispiel:

$$V = H_0^1(\Omega), \quad H = L^2(\Omega)$$

$$a(u,v) = \int_\Omega \nabla u \nabla v \, dx$$

Nach dem Satz von Lax–Milgram existiert $T \in L[H,V]$, so daß

$$a(Tf, v) = \langle f, v \rangle \quad \forall v \in V$$

Somit ist (3.3) äquivalent zu folgender Formulierung:
Finde $\lambda \in \mathbf{R}\,(\mathbf{C})$, $u \in V$, $|u| = 1$ mit

$$u = \lambda T u$$

Die Eigenschaften von T werden in dem folgenden Lemma zusammengefaßt:

Lemma 3.1 *V werde mit Skalarprodukt $a(\cdot\,,\cdot)$ betrachtet. Dann gilt:*

1. *T ist symmetrisch: $a(Tu,v) = a(u,Tv)$*
2. *T ist positiv: $a(Tv,v) > 0$ für $v \neq 0$.*

Ferner ist $T \in L[H,V]$ linear beschränkt.

Begründung:

1. $a(Tu,v) = \langle u,v\rangle = a(Tv,u) = a(u,Tv)$

2. $a(Tv,v) = \langle v,v\rangle > 0$ für $v \neq 0$

$\alpha\|Tf\|^2 \leq a(Tf,Tf) = \langle f,Tf\rangle \leq |f||Tf| \leq C|f|\|Tf\| \Rightarrow T$ ist linear beschränkt. $\square$

Ab jetzt sei $V \hookrightarrow\hookrightarrow H$, d.h. V sei kompakt und stetig in H eingebettet.

Folgerung: $T : V \to V$ ist eine kompakte Abbildung.

Satz 3.2 *Die Eigenwerte aus (3.3) erfüllen die folgende Beziehung:* $0 < \lambda_1 \leq \lambda_2 \leq \dots$ *und es gibt eine in* H *orthonormale Hilbertbasis* $\{w_m\}$*, wobei* w_m *Eigenvektor zu* λ_m *ist. (Dabei sei* $\{\lambda_m^{-\frac{1}{2}} w_m\}$ *orthonormale Hilbertbasis in* V *bzgl.* $a(\cdot\,,\cdot)$*)*

Beweis: Aus der Funktionalanalysis ist bekannt, daß für die kompakte, symmetrische Abbildung $T \in L(V,V)$ gilt: Die Eigenwerte μ_m von T sind strikt positiv und es existiert eine orthonormale Hilbertbasis aus Eigenvektoren $\{v_m\}$, d.h. $v_m \neq 0$ und $Tv_m = \mu_m v_m$. Mit $\lambda_m = \mu_m^{-1}$ gilt $a(v_m,v) = \lambda_m a(Tv_m,v) = \lambda_m\langle v_m,v\rangle \quad \forall v \in V$, wobei für das zweite Gleichheitszeichen der Satz von Lax–Milgram verwendet wurde.
Wählt man nun $w_m = \lambda_m^{\frac{1}{2}} v_m$, dann gilt:

$$\langle w_m,w_n\rangle = \lambda_m^{-1} a(w_m,w_n) = \delta_{mn} \quad \text{(Orthonormalität)}$$

Sei nun $f \in H$ mit $\langle f,w_m\rangle = 0 \quad \forall m \in \mathbf{N}$. Dann gilt:

$$\begin{aligned}
& \langle f,v_n\rangle = 0 \quad \forall n \in \mathbf{N} \\
\Rightarrow\ & a(Tf,v_n) = 0 \quad \forall n \in \mathbf{N} \\
\Rightarrow\ & a(Tf,v) = 0 \quad \forall v \in V \\
\Rightarrow\ & \langle f,v\rangle = 0 \quad \forall v \in V \\
\Rightarrow\ & f = 0 \quad \text{(Vollständigkeit)} \qquad \square
\end{aligned}$$

Bemerkung 3.3 *Ist* a *nur koerziv, d.h.* $\exists \alpha > 0, \lambda \in \mathbf{R} : a(v,v) + \lambda\|v\|^2 \geq \alpha\|v\|^2 \quad \forall v \in V$*, dann ist* $\tilde{a}(u,v) = a(u,v) + \lambda\langle u,v\rangle$ V*–elliptisch und die Eigenwerte von* a *sind gerade die Eigenwerte von* $\tilde{a} - \lambda$*. Nach Satz 3.2 gilt dann die Behauptung mit* $-\lambda < \lambda_1 \leq \lambda_2 \leq \dots$

$$f = \sum_{m\in\mathbf{N}} \langle f,w_m\rangle w_m, \qquad |f| = \left(\sum_{m\in\mathbf{N}} |\langle f,w_m\rangle|^2\right)^{\frac{1}{2}}$$

da

$$f = \sum_{m\in\mathbf{N}} \alpha_m w_m \Rightarrow \langle f, w_k\rangle = \sum_{m\in\mathbf{N}} \alpha_m \langle w_m, w_k\rangle = \alpha_k$$

und

$$\begin{aligned} |f| &= \lim_{n\to\infty} \left|\sum_{m=1}^{n} \langle f, w_m\rangle w_m\right| = \lim_{n\to\infty} \left(\sum_{m=1}^{n} |\langle f, w_m\rangle|^2 \langle w_m, w_m\rangle\right)^{\frac{1}{2}} = \\ &= \left(\sum_{m\in\mathbf{N}} |\langle f, w_m\rangle|^2\right)^{\frac{1}{2}}. \end{aligned}$$

Für $v_m := \lambda_m^{-\frac{1}{2}} w_m$ *folgt:*

$$a(v,v) = \sum_{m\in\mathbf{N}} |a(v, v_m)|^2 = \sum_{m\in\mathbf{N}} \lambda_m^{-1} |a(v, w_m)|^2 = \sum_{m\in\mathbf{N}} \lambda_m |\langle v, w_m\rangle|^2$$

3.2 Anwendung auf elliptische RWA 2. Ordnung

Sei $\Omega \subset \mathbf{R}^n$ ein Gebiet, dessen Rand $\partial\Omega$ Lipschitz-stetig ist. Der Differentialoperator L habe die Form:

$$(Lu)(x) := -\mathrm{div}(K(x)\nabla u(x)) + c(x)\nabla u(x) + d(x)u(x)$$

mit $K_{ij},\ c_i,\ d \in L^\infty(\Omega)$; $\sum_{ij} K_{ij}(x)\xi_i\xi_j \geq K_0|\xi|^2\ \forall \xi \in \mathbf{R}^n, x \in \Omega$; $d(x) \geq d_0$ für ein $d_0 \in \mathbf{R}$; $\Gamma := \partial\Omega = \Gamma_1 \dot\cup \Gamma_2 \dot\cup \Gamma_3$;

$$\begin{aligned} \partial_{\eta_K} u(x) := K(x)\nabla u(x)\eta(x) &= g_1(x) \quad x \in \Gamma_1 \\ \partial_{\eta_K} u(x) + \alpha(x)u(x) &= g_2(x) \quad x \in \Gamma_2 \\ u(x) &= g_3(x) \quad x \in \Gamma_3 \end{aligned}$$

Variationelle Formulierung:

Sei
$$\begin{aligned} & w \in H^1(\Omega) \text{ mit } w|_{\partial\Omega} = g_3, \\ & V := \left\{ v \in H^1(\Omega) \mid v|_{\Gamma_3} = 0 \right\}, \\ & a(u,v) := \int_\Omega (K\nabla u\nabla v + c\nabla uv + duv)\, dx + \int_{\Gamma_2} \alpha uv\, d\sigma, \\ & \tilde{b}(v) := b(v) - a(w,v), \\ & b(v) := \int_\Omega fv\, dx + \int_{\Gamma_1} g_1 v\, d\sigma + \int_{\Gamma_2} g_2 v\, d\sigma, \\ & H := L^2(\Omega). \end{aligned}$$

Voraussetzungen aus Abschnitt 3.1:

1. $V \hookrightarrow H$: es gilt sogar $V \hookrightarrow\hookrightarrow H$
2. V liegt dicht in H, da nach dem Satz von Lusin gilt: $C_0^\infty(\Omega)$ liegt dicht in $L^2(\Omega)$
3. a ist stetig und koerziv
4. a ist symmetrisch, falls K symmetrisch und $c = 0$
5. a ist V-elliptisch, falls einer der folgenden Fälle vorliegt:
   (a) $c = 0$
      i. $d_0 > 0$
      ii. $\text{meas}_{n-1}(\Gamma_3) \neq 0$ und $d_0 > -\frac{K_0}{C_p^2}$,
      C_p: Konstante aus der Poincaré-Ungleichung

$$\|v\|_0 \leq C_p |v|_1 = C_p \left(\sum_{i \in \mathbf{N}} \|\partial_{x_i} v\|_0^2 \right)^{\frac{1}{2}}$$

   (b) $c \neq 0$, c ist konstant, $\Gamma = \Gamma_3$, $d_0 > -\frac{K_0}{C_p^2}$

$\Rightarrow$ Alle abstrakten Aussagen gelten, insbesondere

$$\begin{array}{l} -\lambda < \lambda_1 \leq \lambda_2 \leq \ldots \text{ für die Eigenwerte} \\ (\lambda = 0, \text{ falls } L\ V\text{-elliptisch ist}) \end{array}$$

Die zugehörigen Eigenfunktionen bilden eine orthonormale Hilbertbasis in $L^2(\Omega)$, d.h. $w_m \in V$ lösen variationell:

$$\begin{array}{l} Lw_m = \lambda_m w_m \quad \text{in } \Omega \\ \partial_{\eta_K} w_m = 0 \quad \text{auf } \Gamma_1 \\ \partial_{\eta_K} w_m + \alpha w_m = 0 \quad \text{auf } \Gamma_2 \\ w_m = 0 \quad \text{auf } \Gamma_3 \end{array}$$

3.3 Approximation von Spektralproblemen

Es werden nun endlichdimensionale Teilräume $V_h \subset V$ betrachtet mit $\dim V_h = n = n(h)$, wobei $\lim_{h \to 0} n(h) \to \infty$.
Gesucht sind hierbei die Eigenfunktionen $u_h \in V_h$ zu den Eigenwerten $\lambda_h \in \mathbf{R}$, so daß gilt:

$$a(u_h, v_h) = \lambda_h \langle u_h, v_h \rangle \quad \forall v_h \in V_h \tag{3.4}$$

Satz 3.4 *Es existieren Eigenwerte* $\lambda_{m,h}$ *von (3.4) mit* $-\lambda < \lambda_{1,h} \leq \lambda_{2,h} \leq \ldots \leq \lambda_{n,h}$ *und es existiert eine Orthonormalbasis* $\{w_{m,h}\}$ *mit* $w_{m,h} \in V_h$ *von zugehörigen Eigenfunktionen.*

Beweis: O.B.d.A. sei a V-elliptisch. Es handelt sich um einen Spezialfall von Satz 3.2 mit $V = (V_h, \|\cdot\|)$, $H = (V_h, |\cdot|)$.
Da V endlichdimensional ist: $V \hookrightarrow\hookrightarrow H$
Direkter Beweis: $\{\varphi_i\}_{i=1,\ldots,n}$ sei Basis von V_h

$$
\begin{aligned}
& u_h = \sum_{j=1}^{n} \xi_j \varphi_j \text{ und (3.4)} \\
\Leftrightarrow\quad & \sum_{j=1}^{n} a(\varphi_j, \varphi_i)\xi_j = \lambda_h \sum_{j=1}^{n} \langle \varphi_j, \varphi_i \rangle \xi_j \\
& \text{wobei} \begin{array}{ll} a(\varphi_j, \varphi_i) = S_{ji} : & \text{Steifigkeitsmatrix} \\ \langle \varphi_j, \varphi_i \rangle = B_{ij} : & \text{Massenmatrix} \end{array} \\
\Leftrightarrow\quad & S\xi = \lambda B \xi
\end{aligned}
$$

$$
\left[\begin{array}{l} B \text{ ist symmetrisch und positiv definit} \\ \Rightarrow\ \exists \text{ Cholesky-Zerlegung } B = LL^T, \text{ wobei } L \text{ eine untere Dreiecksmatrix} \\ \Rightarrow\ L^{-1} S L^{-T} L^T \xi = L^T \xi \lambda \end{array} \right]
$$

$$
\begin{aligned}
\Leftrightarrow\quad & A\eta = \lambda\eta \quad \text{mit } A := L^{-1} S L^{-T}, \eta = L^T \xi \\
\Rightarrow\quad & \text{Behauptung} \qquad \square
\end{aligned}
$$

4 Parabolische Randwertaufgaben – abstrakte Theorie

4.1 Beispiel: Wärmeleitungsgleichung im $\mathbf{R}^n$

Sei $\Omega \subset \mathbf{R}^n$ ein beschränktes offenes Gebiet mit hinreichend glattem Rand $\partial\Omega$, $Q_T := \Omega \times (0,T]$, $S_T := \partial\Omega \times (0,T]$.
Gesucht ist ein $u : Q_T \to \mathbf{R}$, so daß für vorgegebene $u_0 : \Omega \to \mathbf{R}$, $f : Q_T \to \mathbf{R}$ gilt:

$$\begin{aligned} \partial_t u - \Delta u &= f && \text{in } Q_T \\ u &= 0 && \text{auf } S_T \\ u &= u_0 && \text{für } t = 0 \end{aligned} \tag{4.1}$$

Mit Testfunktionen $v \in H_0^1(\Omega)$ ergibt sich die schwache Formulierung:

$$\int_\Omega \partial_t u(x,t)v(x)\,dx - \int_\Omega \Delta u(x,t)v(x)\,dx = \int_\Omega f(x,t)v(x)\,dx$$

Aus
$$\int_\Omega \partial_t u(x,t)v(x)\,dx = \frac{d}{dt}\int_\Omega u(x,t)v(x)\,dx$$
und mit partieller Integration folgt:

$$\frac{d}{dt}\int_\Omega u(x,t)v(x)\,dx + \int_\Omega \nabla u(x,t)\nabla v(x)\,dx = \int_\Omega f(x,t)v(x)dx \qquad \forall t \in (0,T]$$

(Bei dieser Variationsgleichung wurde — wie bei parabolischen Partiellen Differentialgleichungen üblich — nur die x–Variable in die schwache Formulierung aufgenommen. Rein prinzipiell könnte jedoch die t–Variable genauso behandelt werden.)
Gesucht ist nun eine Funktion u mit der Eigenschaft, daß für jedes t eine Funktion $x \mapsto (u(t))(x) := u(x,t)$ gegeben ist. Analog dazu ist auch für f eine funktionswertige Abbildung gesucht.
Für die folgenden Betrachtungen vereinbaren wir als Schreibweisen:

$$\begin{aligned} \langle\varphi,\psi\rangle &:= \int_\Omega \varphi(x)\psi(x)\,dx && \text{für alle} \quad \varphi,\psi \in L^2(\Omega) \\ a(\varphi,\psi) &:= \int_\Omega \nabla\varphi\nabla\psi\,dx && \text{für alle} \quad \varphi,\psi \in H^1(\Omega) \end{aligned}$$

Definition: (schwache Formulierung von (4.1))
Finde eine Funktion $u : [0,T] \to H_0^1(\Omega)$, so daß

$$\frac{d}{dt}\langle u(t),v\rangle + a(u(t),v) = \langle f(t),v\rangle$$

für alle $v \in H_0^1(\Omega), t \in (0,T]$ gilt.

Die Zeitableitung ist dabei im disributionellen Sinne zu verstehen (s.u.).

4.2 Abstrakte parabolische Gleichungen

Mit V, H bezeichnen wir separable Hilberträume, die folgende Eigenschaften erfüllen:

$$\begin{array}{ll} V \hookrightarrow H & \text{(stetig eingebettet)} \\ \overline{V} = H & \text{(V liegt dicht in H bzgl. der H–Norm)} \end{array}$$

a bezeichne stets eine stetige Bilinearform auf $V \times V$, $\|\,.\,\|$ die Norm auf V, $\langle .,. \rangle$ das Skalarprodukt in H und $|\,.\,|$ die davon erzeugte Norm.

Banachraumwertige Abbildungen

$(X, \|.\|_X)$ sei ein Banachraum; dann definiert man:

$$\begin{aligned} C^m([0,T];X) &:= \{v : [0,t] \to X \mid v \text{ ist } m\text{-fach stetig differenzierbar}\}, \\ \|v\|_{C^m([0,T];X)} &:= \max_{0 \le l \le m} \left\{ \sup_{0 \le t \le T} \| \frac{d^l}{dt^l} v(t) \|_X \right\}, \\ L^p(0,T;X) &:= \{v : [0,T] \to X \mid v \text{ stark meßbar, } \|v\|_{L^p(0,T;X)} < \infty\} \quad \text{mit} \\ \|v\|_{L^p(0,T;X)} &:= \left(\int_0^T \|v(t)\|_X^p \, dt \right)^{1/p}. \end{aligned}$$

Definition: $u : [a,b] \to X$ heißt *stark meßbar*, wenn es eine Folge $(u_n)_{n \in \mathbf{N}}$ einfacher Funktionen gibt, so daß

$$\lim_{n \to \infty} \|(u_n - u)(t)\|_X = 0 \qquad \text{für alle } t \in [a,b].$$

Dabei heißt eine Funktion $v : [a,b] \to X$ *einfach*, wenn es eine Zerlegung von $[a,b]$ in endlich viele disjunkte und meßbare Mengen A_i, $i = 1, 2, \ldots, l$, gibt, so daß $u|_{A_i}$ konstant ist für alle $i = 1, 2, \ldots, l$.
u heißt *schwach meßbar*, wenn für alle $\varphi \in X'$ die Funktion $\varphi(u) : [a,b] \to \mathbf{R}$ meßbar ist.

Satz 4.1 *Sei u stark meßbar. Dann gilt:*
i) $t \mapsto \|u(t)\|_X$ *ist meßbar,*
ii) u *ist schwach meßbar.*

Beweis: Wegen der Stetigkeit der Abbildungen $\|.\|_X$ und $\varphi \in X'$ überträgt sich die starke Meßbarkeit von u auf $\|u\|_X$ und $\varphi(u)$. □

Sei X ein Hilbertraum und $t \mapsto u(t)$ eine $L^2(0,T;X)$–Abbildung;
dann ist $t \mapsto \langle u(t), v \rangle_X$ eine $L^2(0,T)$–Abbildung für alle $v \in X$. Dabei gilt (durch Anwendung der Cauchy-Schwarz-Ungleichung):

$$|\langle u(t), v \rangle_X|^2 \le \|u(t)\|_X^2 \, \|v\|_X^2.$$

Ferner ist $L^2(0,T;X)$ mit dem Skalarprodukt

$$\langle u, v\rangle_{L^2(0,T;X)} := \int_0^T \langle u(t), v(t)\rangle_X \, dt \quad \text{für } u, v \in L^2(0,T;X)$$

ein Hilbertraum.

Definition: Sei $u_0 \in H$, $f \in L^2(0,T;H)$. Eine Lösung des abstrakten parabolischen Problems ist gegeben durch ein $u \in L^2(0,T;V) \cap C(0,T;H)$, das die Variationsgleichung

$$\frac{d}{dt}\langle u(t), v\rangle + a(u(t), v) = \langle f(t), v\rangle \quad \text{für alle } v \in V \tag{4.2}$$

im Sinne der Distributionen auf $(0,T]$ und

$$u(0) = u_0$$

erfüllt.

(Die Forderung $u \in C(0,T;H)$ beruht auf der Anfangsbedingung $u(0) = u_0$. Alle Funktionen $t \mapsto \langle u(t), v\rangle$, $t \mapsto a(u(t), v)$, $t \mapsto \langle f(t), v\rangle$ liegen in $L^2(0,T)$.)

Beispiel: Wärmeleitungsgleichung mit homogener Dirichletbedingung
$V := H_0^1(\Omega)$, $H := L^2(\Omega)$, $a(u,v) := \int_\Omega \nabla u \nabla v \, dx$
Nach dem Satz von Fubini kann $L^2(0,T;L^2(\Omega))$ mit $L^2(Q_T)$ identifiziert werden.

Im folgenden sei a *koerziv*, d.h. es existieren Zahlen $\alpha > 0$, $\lambda \in \mathbf{R}$, so daß

$$a(u,v) + \lambda |v|^2 \geq \alpha \|v\|^2 \qquad \text{für alle} \quad v \in V. \tag{4.3}$$

Darüberhinaus sei

$$V \hookrightarrow\hookrightarrow H \qquad \text{und} \qquad a \quad \text{symmetrisch.} \tag{4.4}$$

Dann existieren Eigenwerte $-\lambda < \lambda_1 \leq \lambda_2 \leq \ldots$ und eine vollständige orthonormale Hilbert–Basis $(w_i)_i$ von H mit $a(w_i, v) = \lambda_i \langle w_i, v\rangle$ für alle $v \in V$.

Lemma 4.2 *Sei u Lösung des abstrakten Problems* (4.2)*, dann gilt:*

$$u(t) = \sum_{i=1}^{\infty} \left\{ \langle u_0, w_i\rangle e^{-\lambda_i t} + \int_0^t \langle f(s), w_i\rangle e^{-\lambda_i (t-s)} \, ds \right\} w_i. \tag{4.5}$$

Beweis: Sei u Lösung von (4.2), dann gilt: $u \in C(0,T;H)$. Für festes t betrachte man nun die Entwicklung von $u(t) \in H$ bzgl. $(w_i)_i$:

$$u(t) = \sum_{i=1}^{\infty} \langle u(t), w_i\rangle w_i = \sum_{i=1}^{\infty} \alpha_i(t) w_i \quad \text{mit} \quad \alpha_i(t) := \langle u(t), w_i\rangle.$$

Andererseits gilt: $a(u(t), w_i) = a(w_i, u(t)) = \lambda_i \langle u(t), w_i \rangle$ für alle $t \in (0,T]$. (4.2) und $v := w_i$ liefern:

$$\frac{d}{dt}\alpha_i(t) + \lambda_i \alpha_i(t) = \frac{d}{dt}\langle u(t), w_i\rangle + a(u(t), w_i) = \langle f(t), w_i\rangle, \qquad (4.6)$$
$$\text{wobei} \quad \alpha_i(0) = \langle u(0), w_i\rangle = \langle u_0, w_i\rangle.$$

Die Lösung dieser Gewöhnlichen Differentialgleichung, bzw. dieses Systems Gewöhnlicher Differentialgleichungen für $i = 1, 2, \ldots, \infty$, ergibt die Behauptung. □

Satz 4.3 (4.2) *besitzt genau eine Lösung.*

Beweis: Wegen Lemma 4.2 genügt es, die Existenz zu zeigen. Der Beweis gliedert sich im wesentlichen in zwei Teile:

1. (4.5) konvergiert in H.
2. Der Grenzwert der Reihe ist die Lösung.

A) Ohne Einschränkung können wir annehmen, daß $\lambda = 0$ in (4.3). Dazu betrachte man die Transformation $u := e^{\lambda t} w$. Dann liegt w in $L^2(0,T;V) \cap C(0,T;H)$ und für alle $v \in V$ gilt:

$$\tfrac{d}{dt}\langle w(t), v\rangle + \lambda \langle w(t), v\rangle + a(w(t), v) = \langle e^{-\lambda t} f(t), v\rangle$$
$$\text{und} \quad w(0) = w_0$$

Setze nun $\tilde{a}(u,v) := \lambda\langle u, v\rangle + a(u,v)$, dann gilt $\tilde{a}(u,v) \geq \alpha\|v\|^2$ für alle $v \in V$ und somit ist $\tilde{a}$ V–elliptisch.
B) Wir setzen nun $\lambda = 0$ und konstruieren Näherungsprobleme:

$$V_m := \operatorname{span}\{w_1, \ldots, w_m\}, \qquad v_{0,m} := \sum_{i=1}^{m} \langle u_0, w_i\rangle w_i \in V_m,$$

Gesucht: $u_m : [0,T] \to V_m$ mit

$$\tfrac{d}{dt}\langle u_m(t), w_i\rangle + a(u_m(t), w_i) = \langle f(t), w_i\rangle \quad \text{für alle} \quad i = 1, \ldots, m$$
$$\text{und} \quad u_m(0) = u_{0,m}$$

Wir erhalten so eine Gewöhnliche Differentialgleichung aus

$$u_m(t) = \sum_{i=1}^{m} \alpha_i(t) w_i, \qquad \alpha_i(t) := \langle u_m(t), w_i\rangle,$$

da wie in Beweis von Lemma 4.2, (4.6), für alle $i = 1, \ldots, m$ gilt:

$$\frac{d}{dt}\alpha_i(t) + \lambda_i \alpha_i(t) = \langle f(t), w_i\rangle.$$

Somit existiert eine Lösung u_m mit:

$$u_m(t) = \sum_{i=1}^{m} \left\{ \langle u_0, w_i \rangle e^{-\lambda_i t} + \int_0^t \langle f(s), w_i \rangle e^{-\lambda_i (t-s)}\, ds \right\} w_i. \tag{4.7}$$

C) Z. zg.: $(u_m)_{m=1}^{\infty}$ ist Cauchyfolge in $L^2(0,T,V)$ und in $C(0,T,H)$.
Seien $p, m \in \mathbf{N}$, $p > m > 1$:

$$|u_p(t) - u_m(t)| = \left(\sum_{i=m+1}^{p} \left[\langle u_0, w_i \rangle e^{-\lambda_i t} + \int_0^t \langle f(s), w_i \rangle e^{-\lambda_i (t-s)}\, ds \right]^2 \right)^{\frac{1}{2}} \leq$$

$$\leq \underbrace{\left(\sum_{i=m+1}^{p} \left| \langle u_0, w_i \rangle e^{-\lambda_i t} \right|^2 \right)^{\frac{1}{2}}}_{A_1} + \underbrace{\left(\sum_{i=m+1}^{p} \left| \int_0^t \langle f(s), w_i \rangle e^{-\lambda_i (t-s)}\, ds \right|^2 \right)^{\frac{1}{2}}}_{A_2}.$$

Aufgrund der Cauchy-Schwarzschen Ungleichung folgt:

$$\begin{aligned} \left| \int_0^t \langle f(s), w_i \rangle e^{-\lambda_i (t-s)}\, ds \right|^2 &\leq \int_0^t \langle f(s), w_i \rangle^2\, ds \underbrace{\int_0^t e^{-2\lambda_i (t-s)}\, ds}_{=\frac{1}{2\lambda_i}\left(1 - e^{-2\lambda_i t}\right)} \leq \\ &\leq \frac{1}{2\lambda_1} \int_0^T \langle f(s), w_i \rangle^2\, ds. \end{aligned}$$

Damit erhält man eine von $t \in [0,T]$ unabhängige obere Schranke und folglich:

$$\sup_{t \in [0,T]} |u_p(t) - u_m(t)| \leq \left(\sum_{i=m+1}^{p} \langle u_0, w_i \rangle^2 \right)^{\frac{1}{2}} + \left(\frac{1}{2\lambda_1} \sum_{i=m+1}^{p} \int_0^T \langle f(s), w_i \rangle^2\, dt \right)^{\frac{1}{2}} \tag{4.8}$$

Um eine Abschätzung bzgl. der V-Norm für den ersten Teil der Behauptung in C) zu bekommen, wird die V-Elliptizität der Bilinearform ausgenutzt. Es gilt:

$$\begin{aligned} &a\left(u_p(t) - u_m(t), u_p(t) - u_m(t)\right) = \\ &\quad = \sum_{i=1}^{p} \lambda_i \langle u_p(t) - u_m(t), w_i \rangle^2 = \\ &\quad = \sum_{i=m+1}^{p} \lambda_i \left(\langle u_0, w_i \rangle e^{-\lambda_i t} + \int_0^t \langle f(s), w_i \rangle e^{-\lambda_i (t-s)}\, ds \right)^2 \end{aligned}$$

Aufgrund der V-Elliptizität von $a(\cdot\,,\cdot)$ und wegen $(a+b)^2 \leq 2(a^2 + b^2)$ folgt somit:

$$\|u_p(t) - u_m(t)\|^2 \leq \frac{2}{\alpha} \sum_{i=m+1}^{p} \lambda_i \left(\langle u_0, w_i \rangle e^{-2\lambda_i t} + \left(\int_0^t \langle f(s), w_i \rangle e^{-\lambda_i (t-s)}\, ds \right)^2 \right)$$

Mit
$$\lambda_i \int_0^T e^{-2\lambda_i t}\, dt = \frac{1}{2}\left(1 - e^{-2\lambda_i T}\right) < \frac{1}{2}$$
und
$$\begin{aligned}\lambda_i \int_0^T \left(\int_0^t \langle f(s), w_i\rangle e^{-\lambda_i(t-s)}\, ds\right)^2\, dt &\le \\ &\le \lambda_i \int_0^T \left(\int_0^t \langle f(s), w_i\rangle^2\, ds\; e^{-2\lambda_i t} \int_0^t e^{2\lambda_i s}\, ds\right)\, dt \le \\ &\le \frac{1}{2}\int_0^T \left(\int_0^t \langle f(s), w_i\rangle^2\, ds\right)\, dt \le \frac{T}{2}\int_0^T \langle f(t), w_i\rangle^2\, dt\end{aligned}$$
ergibt sich die Ungleichung
$$\int_0^T \|u_p(t) - u_m(t)\|^2\, dt \le \frac{1}{\alpha}\sum_{i=m+1}^{p}\left(\langle u_0, w_i\rangle^2 + T\int_0^T \langle f(t), w_i\rangle^2\, dt\right). \tag{4.9}$$
Ferner wegen
$$\begin{aligned} u_0 \in H &\Rightarrow |u_0|^2 = \sum_{i\in\mathbf{N}} \langle u_0, w_i\rangle^2 < \infty \\ f \in L^2(0,T,H) &\Rightarrow \|f\|_{L^2(0,T,H)} = \sum_{i\in\mathbf{N}} \int_0^T \langle f(t), w_i\rangle^2\, dt < \infty \end{aligned}$$
gilt:
$$\begin{aligned} \lim_{m,p\to\infty} \sum_{i=m+1}^{p} \langle u_0, w_i\rangle^2 = 0 \\ \lim_{m,p\to\infty} \sum_{i=m+1}^{p} \int_0^T \langle f(t), w_i\rangle^2\, dt = 0 \end{aligned}$$
Daraus folgt in Verbindung mit (4.8):
$$\lim_{m,p\to\infty} \|u_p - u_m\|_{C(0,T,H)} = 0$$
und zusammen mit (4.9):
$$\lim_{m,p\to\infty} \|u_p - u_m\|_{L^2(0,T,V)} = 0.$$
D) Zu zeigen: Die Folge $(u_m)_{m=1}^{\infty}$ konvergiert in $L^2(0,T,V)$ und in $C(0,T,H)$ gegen den gleichen Grenzwert u und dieser löst das betrachtete Problem (4.2).

$L^2(0,T,V)$, $L^2(0,T,H)$ sind vollständig und $L^2(0,T,V) \hookrightarrow L^2(0,T,H)$, somit:

$$u_p \overset{p\to\infty}{\longrightarrow} u_V \quad \text{in } L^2(0,T,V) \qquad \Rightarrow \qquad u_p \overset{p\to\infty}{\longrightarrow} u_V \quad \text{in } L^2(0,T,H)$$

Wegen $C(0,T,H) \hookrightarrow L^2(0,T,H)$ gilt gleiches für den Grenzwert u_H von $(u_m)_{m=1}^{\infty}$ in $C(0,T,H)$.
Es folgt:

$$u_H = u_V \qquad \text{bzw.} \qquad u_m \overset{m\to\infty}{\longrightarrow} u \quad \text{in } L^2(0,T,V) \cap C(0,T,H)$$

Nachzuweisen bleibt, daß u Lösung von (4.2) ist.
Sei $\mu \geq 1$, $m > \mu$. Dann gilt für die Lösung $u_m \in V_m$ im distributionellen Sinn:

$$\frac{d}{dt}\langle u_m(t), v\rangle + a(u_m(t), v) = \langle f(t), v\rangle \qquad \forall\, v \in V_\mu,$$

d. h.

$$\begin{aligned} -\int_0^T \langle u_m(t), v\rangle \frac{d\varphi(t)}{dt} + \int_0^T a(u_m(t), v)\varphi(t)\, dt = \\ = \int_0^T \langle f(t), v\rangle \varphi(t)\, dt \qquad \forall\, v \in V_\mu \quad \forall\, \varphi \in \mathcal{D}(0,T) \end{aligned}$$

Für festes μ gilt für den Grenzwertübergang $m \to \infty$

- $\langle u_m(t), v\rangle \to \langle u(t), v\rangle$ gleichmäßig in t (vgl. (4.8))
- die Abbildung $t \mapsto a(u_m(t), v)$ konvergiert in $L^2(0,T)$ gegen
$$t \mapsto a(u(t), v)$$

Es folgt:

$$\begin{aligned} -\int_0^T \langle u(t), v\rangle \frac{d\varphi(t)}{dt} + \int_0^T a(u(t), v)\varphi(t)\, dt = \\ = \int_0^T f(t)\varphi(t)\, dt \qquad \forall\, v \in V_\mu \quad \forall\, \varphi \in \mathcal{D}(0,T) \end{aligned} \tag{4.10}$$

$\bigcup_{\mu\in\mathbf{N}} V_\mu$ ist dicht in V, folglich gilt (4.10) für alle $v \in V$.
Zu prüfen bleibt lediglich noch, ob u auch die Anfangsbedingung erfüllt.
Wegen $u_m \to u$ in $C(0,T,H)$ folgt:

$$u_m(0) \to u(0) \quad \text{in } H$$

Mit

$$u_m(0) = \sum_{i=1}^{m} \langle u_0, w_i\rangle w_i \overset{m\to\infty}{\longrightarrow} \sum_{i=1}^{\infty} \langle u_0, w_i\rangle w_i = u_0 \quad \text{in } H$$

erhält man dann:

$$u(0) = u_0 \qquad \square$$

Satz 4.4 *Unter den obigen Voraussetzungen gilt:*

$$|u(t)| \leq |u_0|e^{-\lambda_1 t} + \int_0^t |f(s)|e^{-\lambda_1(t-s)}\,ds$$

Beweis: Aus der Lösungsdarstellung von Lemma 4.2 und analog zur Herleitung von (4.8). □

Beispiele für parabolische ARWA 2. Ordnung

Sei $\Omega \subset \mathbf{R}^n$ offen, beschränkt mit Rand $\Gamma = \partial\Omega \in C^1$.
V sei ein abgeschlossener Teilraum von $H^1(\Omega)$ mit

$$H_0^1(\Omega) \subset V \subset H^1(\Omega).$$

$H := L^2(\Omega)$
Seien weiter $a_{ij} \in L^\infty(\Omega)$ mit

$$a_{ij} = a_{ji} \qquad 1 \leq i,j \leq n$$

und

$$\exists\, \alpha > 0 : \sum_{i,j=1}^n a_{ij}(x)\xi_j\xi_i \geq \alpha|\xi|^2 \quad \forall\, \xi \in \mathbf{R}^n,\ \text{f. ü. in } \Omega$$

und $a_0 \in L^\infty(\Omega) \geq \text{const} \geq 0$ gegeben.
Mit Hilfe dieser Koeffizientenfunktionen wird die folgende Bilinearform definiert:

$$a(u,v) := \int_\Omega \sum_{i,j=1}^n a_{ij}\partial_{x_i}u\partial_{x_j}v + a_0 uv\,dx$$

Nach Satz 4.3 existiert dann zu jedem $u_0 \in L^2(\Omega)$, $f \in L^2(0,T,L^2(\Omega))$ eine eindeutige Lösung $u \in L^2(0,T,V) \cap C(0,T,L^2(\Omega))$ des Problems:

$$\frac{d}{dt}\langle u(t),v\rangle + a(u(t),v) = \langle f(t),v\rangle \quad \forall\, v \in V$$

Für $v \in \mathcal{D}(\Omega)$, $\varphi \in \mathcal{D}(0,T)$ gilt:

$$-\int_0^T \langle u(t),v\rangle\frac{d\varphi(t)}{dt} + \int_0^T a(u(t),v)\varphi(t)\,dt = \int_0^T \langle f(t),v\rangle\varphi(t)\,dt$$

Da span $\{v(x)\varphi(t) \,|\, v \in \mathcal{D}(\Omega), \varphi \in \mathcal{D}(0,T)\}$ dicht in $\mathcal{D}(Q_T)$ ist, kann obige Gleichung im distributionellen Sinn bzgl. Q_T wie folgt notiert werden:

$$\frac{du}{dt} + Au = f \qquad \text{in } \mathcal{D}'(Q_T),$$

wobei

$$Au = -\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + a_0 u.$$

Die Randbedingungen sind entweder in der Definition von V enthalten (homogene Dirichletrandbedingung) oder ergeben sich durch partielle Integration der aus

$$\frac{d}{dt}\langle u(t), v\rangle + a(u(t), v) = \int_\Omega \left(\frac{\partial u}{\partial t} + Au \right) v \, dx \quad \forall\, v \in V$$

folgenden Identität

$$a(u(t), v) = \int_\Omega Au(t) v \, dx \quad \forall\, v \in V.$$

5 Die Methode der Semidiskretisierung

Sei $V_h \subset V$ ein endlich-dimensionaler Teilraum mit $\dim V_h = M = M(h)$. Betrachtet wird ein Näherungsproblem:

$$
\begin{aligned}
\text{Gegeben:} \quad & u_{0,h} \in V_h \\
\text{Gesucht:} \quad & u_h : [0,T] \to V_h \\
\text{mit} \quad & \frac{d}{dt}\langle u_h(t), v_h\rangle + a(u_h(t), v_h) = \langle f(t), v_h\rangle \qquad \forall v_h \in V_h \qquad (5.1) \\
\text{und} \quad & u_{0,h} = u_h(0) \qquad (5.2)
\end{aligned}
$$

Aus der Spektraltheorie folgt: Es existieren Eigenwerte $-\lambda < \lambda_{1,h} \leq \ldots \leq \lambda_{M,h}$ und eine Basis $\{w_{i,h}\}$ von V_h aus ortonormalen Eigenfunktionen, d.h. $a(w_{i,h}, v_h) = \lambda_{i,h}\langle w_{i,h}, v_h\rangle \quad \forall v_h \in V_h$.

Satz 5.1 *Das Problem* (5.1),(5.2) *hat eine eindeutig bestimmte Lösung und*

$$
u_h(t) = \sum_{i=1}^{M} \left\{ \langle u_{0,h}, w_{i,h}\rangle e^{-\lambda_{i,h} t} + \int_0^t \langle f(s), w_{i,h}\rangle e^{-\lambda_{i,h}(t-s)}\, ds \right\} w_{i,h} \tag{5.3}
$$

Beweis: Sei $\{\varphi_i\}_{i=1,\ldots,M}$ eine Basis von V_h, $u_h(t) = \sum_{i=1}^{M} \xi_i(t)\, \varphi_i$ und $u_{0,h} = \sum_{i=1}^{M} \xi_{0,i}\, \varphi_i$. Dann gilt:

$$
\begin{aligned}
&(5.1), (5.2) \Leftrightarrow \\
&\sum_{j=1}^{M} \langle \varphi_j, \varphi_i\rangle \frac{d\xi_j(t)}{dt} + \sum_{j=1}^{M} a(\varphi_j, \varphi_i)\, \xi_j(t) = \langle f(t), \varphi_i\rangle \qquad \forall i \in \{1, \ldots, M\}.
\end{aligned}
$$

Mit der Steifigkeitsmatrix $S := (a(\varphi_j, \varphi_i))_{ij}$, der Massenmatrix $B := (\langle \varphi_j, \varphi_i\rangle)_{ij}$ und $\beta(t) := (\, \langle f(t), \varphi_i\rangle\,)_i$ erhält man:

$$
B\, \frac{d}{dt}\xi(t) + S\, \xi(t) = \beta(t) \tag{5.4}
$$

$$
\xi(0) = \xi_0. \tag{5.5}
$$

□

Voraussetzungen: a sei V-elliptisch und symmetrisch, und $V \hookrightarrow\hookrightarrow H$.

Definition: *Elliptische oder Ritz–Projektion R_h*
$R_h \in L(V, V_h)$ ist definiert durch

$$
a(R_h u - u, v_h) = 0 \qquad \forall v_h \in V_h.
$$

Dies ist äquivalent mit

$$a(R_h u - u, R_h u - u) = \min_{v_h \in V_h} a(v_h - u, v_h - u) \tag{5.6}$$

(Zur Beschränktheit von R_h siehe Beweis zu Satz 5.2.)

Satz 5.2 *Sei $u \in C^1(0,T;V)$. Dann gilt:*

$$\begin{aligned} |u_h(t) - u(t)| \quad &\leq \quad |u_{0,h} - R_h u_0| e^{-\lambda_1 t} + |(I - R_h) u(t)| \\ &+ \int_0^t \left| (I - R_h) \frac{du}{dt} \right| e^{-\lambda_1 (t-s)}\, ds, \end{aligned} \tag{5.7}$$

wobei $\lambda_1 > 0$ der kleinste Eigenwert von a sei.

Beweis:

$$\begin{aligned} &\frac{d}{dt}\langle u_h(t) - R_h u(t), v_h\rangle + a(u_h(t) - R_h u(t), v_h) \\ &+\frac{d}{dt}\langle R_h u(t) - u(t), v_h\rangle + a(R_h u(t) - u(t), v_h) \quad = \quad 0 \qquad \forall v_h \in V_h \end{aligned} \tag{5.8}$$

Der 4. Term verschwindet per definitionem. Der 3. Term wird wie folgt umgeformt:

$$\begin{array}{rc} & \alpha \,\|R_h u\|^2 \leq a(R_h u, R_h u) = a(u, R_h u) \leq C \,\|R_h u\| \,\|u\| \\ \Rightarrow & R_h \in L(V, V_h) \quad \text{mit} \quad \|R_h u\| \leq \frac{C}{\alpha}\|u\|. \end{array}$$

$$\|R_h u(t+\varepsilon) - R_h u(t) + \varepsilon R_h \frac{d}{dt} u(t)\| \leq \frac{C}{\alpha}\|u(t+\varepsilon) - u(t) + \varepsilon \frac{d}{dt} u(t)\| = O(\varepsilon),$$

da $u \in C^1(0,T;V)$.
Damit gilt:

$$\begin{array}{lcl} u \in C^1(0,T;V) & \Rightarrow & R_h u \in C^1(0,T;V) \\ \text{und} \qquad \frac{d}{dt}(R_h u) = R_h \frac{d}{dt} u & & \end{array}$$

Somit erhält man für den 3. Term:

$$\frac{d}{dt}\left\langle u(t) - R_h u(t), v_h \right\rangle = \left\langle \frac{du}{dt} - R_h \frac{du}{dt}, v_h \right\rangle$$

Aus (5.8) wird somit:

$$\frac{d}{dt}\langle u_h(t) - R_h u(t), v_h\rangle + a(u_h(t) - R_h u(t), v_h) = \left\langle \frac{du}{dt} - R_h \frac{du}{dt}, v_h \right\rangle$$

Damit ergibt sich:

$$\begin{aligned} u_h(t) - R_h u(t) &= \sum_{i=1}^{M} \Bigg\{ \langle u_{0,h} - R_h u_0, w_{i,h} \rangle e^{-\lambda_{i,h} t} \\ &\quad + \int_0^t \left\langle (I - R_h) \frac{du}{dt}, w_{i,h} \right\rangle e^{-\lambda_{i,h}(t-s)}\, ds \Bigg\} w_{i,h} \end{aligned}$$

$$\Rightarrow \quad |u_h(t) - R_h u(t)| \le |u_{0,h} - R_h u_0| e^{-\lambda_{1,h} t} + \int_0^t \left| (I - R_h) \frac{du}{dt}(s) \right| e^{-\lambda_{1,h}(t-s)}\, ds$$

Aus der Beziehung $\lambda_1 \le \lambda_{1,h}$ (ohne Beweis) ergibt sich $e^{-\lambda_{1,h} t} \le e^{-\lambda_1 t}$, und damit folgt die Behauptung aus:

$$|u_h(t) - u(t)| \le |u_h(t) - R_h u(t)| + |R_h u(t) - u(t)|. \qquad \square$$

Korollar 5.3 *Wenn gilt:* $\lim_{h \to 0} |u_{0,h} - u_0| = 0$, *dann folgt:* $\lim_{h \to 0} |u_h(t) - u(t)| = 0$.

Zum *Beweis* beachte Arzela–Ascoli.

Interpretation von Satz 5.2 Fehlerabschätzung von $|u_h(t) - u(t)|$ durch

- Ritzprojektionsfehler von u_0 (exponentiell fallend in t)
- Ritzprojektionsfehler von $u(t)$
- Ritzprojektionsfehler von $\frac{du}{dt}$ auf $[0, T]$ (gewichtet mit $e^{-\lambda_1(t-s)}$)

Sei $f \in V'$ gegeben. $Tf \in V$ sei die (nach Lax–Milgram) gegebene Lösung von $a(Tf, v) = \langle f, v \rangle \quad \forall v \in V$. $T \in L(V', V)$.
Die Galerkin-Approximation hierzu liefert: $T_h \in L(V', V_h)$ mit $a(T_h f, v_h) = \langle f, v_h \rangle \quad \forall v_h \in V_h$.

$$\begin{aligned} &\Rightarrow \quad a(Tf - T_h f, v_h) = 0 \quad \forall f_h \in V_h, \quad \text{d.h. } T_h f = R_h T f \\ &\Rightarrow \quad u - R_h u = Tf - T_h f \quad \text{für} \quad f := T^{-1} u \end{aligned}$$

$$\begin{aligned} \alpha \|u - R_h u\|^2 &\le a(u - R_h u, u - R_h u) = a(u - R_h u, u - v_h) \\ &\le C \|u - R_h u\| \, \|u - v_h\| \quad \forall v_h \in V_h \\ \Rightarrow \qquad & \|u - R_h u\| \le \frac{C}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|, \end{aligned}$$

d.h. $\|u - R_h u\|$ ist (asymptotisch) optimal. Beispielsweise gilt für V_h = lineare Elemente mit regulärer Triangulierung: $\|u - R_h u\| \le Ch \|u\|_2$ für $u \in H^2(\Omega)$.

Allgemeiner: Sei $\Omega \subset \mathbf{R}^n$ beschränkt und polyedrisch. $(\mathcal{T}_h)$ sei eine reguläre Zerlegung in Simplices, d.h. es existiert ein $\sigma > 0$, so daß $h_K < \sigma\rho_K \ \forall K \in \mathcal{T}_h$, wobei $h_K = \mathrm{diam}(K)$ und ρ_K gleich dem Maximum der Durchmesser der in K enthaltenen Kugeln sei. Nicht erlaubt ist also ein beliebig spitz werdender Winkel in $K \in \mathcal{T}_h$.
Ferner gelte $h := \max\{h_K \,|\, K \in \mathcal{T}_h\} \to 0$.
Die finiten Elemente (K, P_K, Σ_K) seien affin-äquivalent bzgl. des Referenzelementes $(\hat{K}, \hat{P}, \hat{\Sigma})$
und es sei $V_h \subset C(\bar{\Omega})$, d.h C^0–Elemente.

Satz 5.4 *$(\hat{K}, \hat{P}, \hat{\Sigma})$ sei Lagrange–Element.*

$$\begin{aligned} \textit{Gilt:} \quad & P^k(\hat{K}) \subset \hat{P} \subset H^1(\hat{K})\,, H^{k+1}(\hat{K}) \hookrightarrow C(\hat{K}), \\ \textit{so folgt:} \quad & \|u - R_h u\| \le C h^k |u|_{k+1} \quad \textit{für}\, u \in H^{k+1}(\Omega). \end{aligned}$$

Die Frage nach einer Verbesserung dieser Abschätzung beantwortet der folgende Satz:

Satz 5.5 (Aubin-Nitsche-Lemma) *Sei a V-elliptisch, $V \hookrightarrow H$, $\bar{V} = H$. Dann gilt*

$$|u - R_h u| \le C\,\|u - R_h u\| \sup_{g \in H} \frac{1}{|g|} \inf_{\varphi_h \in V_h} \|\varphi_g - \varphi_h\|,$$

wobei für $g \in H$ $\varphi_g \in V$ als die eindeutige Lösung von

$$a^*(\varphi_g, v) := a(v, \varphi_g) = \langle g, v\rangle \quad \textit{für alle}\ v \in V \qquad \textit{(adjungiertes Problem)}$$

gegeben ist.

Beweis: $u_h := R_h u$, dann

$$|u - u_h| = \sup_{g \in H, g \neq 0} |\langle g, u - u_h\rangle| / |g|$$

Nach Lax-Milgram existiert φ_g eindeutig, somit

$$\begin{aligned} & a(u - u_h, \varphi_g) = \langle g, u - u_h\rangle \\ \Rightarrow\ & \langle g, u - u_h\rangle = a(u - u_h, \varphi_g - \varphi_h) \quad \text{für alle } \varphi_h \in V_h \\ \Rightarrow\ & |\langle g, u - u_h\rangle| \le C\|u - u_h\| \cdot \|\varphi_g - \varphi_h\| \\ \Rightarrow\ & |\langle g, u - u_h\rangle| \le C\|u - u_h\| \inf_{\varphi_h \in V_h} \|\varphi_g - \varphi_h\|. \qquad \square \end{aligned}$$

Eine Anwendung von Satz 5.4 wäre:
Eine elliptische Randwert-Anfangs-Aufgabe heißt *regulär*, wenn zu $f \in L^2(\Omega)$ die Lösung von $a(u, v) = \langle f, v\rangle$ für alle $v \in V$, $H_0^1 \subset V \subset H^1(\Omega)$ folgende Bedingung erfüllt:

$$u \in H^2(\Omega), \quad \|u\|_2 \le C\,\|f\|_0.$$

Korollar 5.6 *Unter den Voraussetzungen von Satz 5.5 gilt*

$$u \in H^2(\Omega), \quad \|u\|_2 \leq C\,\|f\|_0,$$

wenn das adjungierte Problem regulär ist.

Beweis: $\inf\limits_{\varphi_h \in V_h} \|\varphi_g - \varphi_h\|_1 \leq C\,\|\varphi_g - R_h\varphi_h\|$
Aus Satz 5.5 folgt (für das reguläre adjungierte Problem)

$$\begin{aligned} \inf_{\varphi_h \in V_h} \|\varphi_g - \varphi_h\|_1 &\leq C \cdot h \cdot \|\varphi_g\|_1 \\ &\leq C \cdot h \cdot \|g\|_0 \qquad \square \end{aligned}$$

Satz 5.7 *Sei* $H_0^1(\Omega) \subset V \subset H^1(\Omega)$, $H = L^2(\Omega)$ *und* $V \hookrightarrow\hookrightarrow H$. *Die Lösung* u *von (5.1, 5.2) erfülle* $u \in C^1(0,T;V)$. a *sei symmetrisch und* V*-elliptisch. Es gelten ferner die Voraussetzungen von Satz 5.4.*
Sei

$$\sigma(k) := \begin{cases} k+1, & \text{falls die elliptische RWA zu } a \text{ regulär ist} \\ k & \text{sonst.} \end{cases}$$

Ist $u \in C^1(0,T;H^{k+1}(\Omega))$, *dann gilt:*

$$\begin{aligned} \|u_h(t) - u(t)\|_0 \;\leq\; & \|u_{0,h} - u_0\|\, e^{-\lambda_1 t} + C \cdot h^{\sigma(k)} \Bigg(\|u_0\|_{k+1} e^{-\lambda_1 t} + \\ & + \|u(t)\|_{k+1} + \int_0^t \left\| \frac{d}{dt} u(s) \right\|_{k+1} e^{-\lambda_1 (t-s)}\, ds \Bigg). \end{aligned}$$

Beweis: Folgt aus Satz 5.2, Satz 5.5 und Korollar 5.6. $\square$

Satz 5.8 *1.* $\|u(t)\| \leq \|u_0\| e^{-\lambda_1 t} + \int_0^t \|f(s)\| e^{-\lambda_1 (t-s)}\, ds$

2. $\|u(t)\| \leq C \cdot \left\{ \frac{1}{\sqrt{t}} e^{-\lambda_1 t} |u_0| + \int_0^t \frac{1}{\sqrt{t-s}} e^{-\lambda_1 (t-s)} |f(s)|\, ds \right\}$ *für* $t > 0$
(parabolische Glättung)
Dabei gelten die Voraussetzungen von Satz 4.4.

6 Volldiskrete Schemata für parabolische Probleme

Nach erfolgter Semidiskretisierung ist noch ein AWP in $\mathbf{R}^{M(h)}$ zu lösen. Dazu betrachten wir Differenzenverfahren für

$$\begin{aligned} y'(t) &= \varphi(t, y(t)), \quad 0 < t \leq T \\ y(0) &= y_0 \end{aligned} \tag{6.1}$$

Dabei sei eine Zeitdiskretisierung $0 = t_0 < t_1 < \ldots < t_k = T$, $\Delta t_\ell := t_\ell - t_{\ell-1}$ gegeben (z. B. äquidistant: $t_k := k \cdot \Delta t$).
Für $r \in [0,1]$ werde (6.1) durch das folgende Einschrittverfahren approximiert:

$$\frac{1}{\Delta t_{k+1}}(y_{k+1} - y_k) = r \cdot \varphi(t_{k+1}, y_{k+1}) + (1-r)\varphi(t_k, y_k) \tag{6.2}$$

Man unterscheidet verschiedene Verfahren abhängig von r:
$r = 1$: nur Auswertung an der alten Zeitschicht, *explizit*
$r = \frac{1}{2}$: *Crank-Nicolson*
$r < 1$: (6.2) ist ein nichtlineares Gleichungssystem für bekanntes y_k und unbekanntes y_{k+1}, *implizit.*
Um die Rechengenauigkeit zu betrachten, beschränken wir uns zunächst auf den eindimensionalen Fall ($M(h) = 1$). Es sei ferner $y \in C^3[0,T]$, $\Delta t_k = \Delta t$ und y Lösung von (6.1).

Konsistenz-Fehler

Mit Taylor-Formel:

$$\begin{aligned} & y(t_{k+1}) - y(t_k) = \Delta t\, y'(t_k) + \frac{\Delta t^2}{2} y''(t_k) + O(\Delta t^3) \\ \Rightarrow\quad & r \cdot \varphi(t_{k+1}, y(t_{k+1})) + (1-r)\varphi(t_k, y(t_k)) = r \cdot y'(t_{k+1}) + (1-r)y'(t_k) = \\ & \qquad = y'(t_k) + r \cdot \Delta t\, y''(t_k) + O(\Delta t^2) \\ \Rightarrow\quad & \frac{1}{\Delta t}(y(t_{k+1}) - y(t_k)) = y'(t_k) + \frac{\Delta t}{2} y''(t_k) + O(\Delta t^2) = \\ & \qquad = r \cdot \varphi(t_{k+1}, y(t_{k+1})) + (1-r)\varphi(t_k, y(t_k)) + \varepsilon_k \end{aligned}$$

mit

$$\varepsilon_k := \left(\frac{1}{2} - r\right) \Delta t\, y''(t_k) + O(\Delta t^2) = \begin{cases} O(\Delta t) & \text{für } r \neq \frac{1}{2} \\ O(\Delta t^2) & \text{für } r = \frac{1}{2} \end{cases} \tag{6.3}$$

Der *Konsistenz-Fehler* ist also für $r = \frac{1}{2}$ minimal.

Stabilität

Zur Bestimmung der Stabilität betrachtet man das Modellproblem:

$$\begin{aligned} y' &= -\lambda y, \quad t \geq 0 \\ y(0) &= y_0 \end{aligned} \tag{6.4}$$

Die Lösung von (6.2) lautet:

$$y_{k+1} = \frac{1-(1-r)\lambda\Delta t}{1+r\lambda\Delta t} y_k$$

Mit $g(x) := \dfrac{1-(1-r)x}{1+rx}$ gilt

$$y_{k+1} = g(\lambda\Delta t) y_k = g(\lambda\Delta t)^k y_0.$$

Für $\lambda > 0$ sollte (y_k) beschränkt sein, am besten:

$$\lim_{k\to\infty} y_k = 0, \quad \text{d. h.: } |g(\lambda\Delta t)| \leq 1. \tag{6.5}$$

Für $r > 0$ gilt:

$$g(x) = \frac{\frac{1}{x} - (1-r)}{\frac{1}{x} + r}$$

und somit

$$\lim_{x\to\infty} g(x) = 1 - \frac{1}{r}.$$

Also folgt aus $|g(x)| \leq 1$ für alle $x \geq 0$:

$$1 - \frac{1}{r} \geq -1 \quad \Leftrightarrow \quad r \geq \frac{1}{2}$$

(6.5) gilt somit für alle $\Delta t > 0$, falls $r \in [\frac{1}{2}, 1]$.

Stabilitätsanforderungen:

$H_A(0) := \{z \in \mathbf{C} \,|\, |g(-z)| < 1\}$ *Bereich absoluter Stabilität*
Ein Verfahren heißt *absolut stabil (A–stabil)*, wenn gilt:

$$z \in \mathbf{C},\ \Re z < 0 \quad \Rightarrow \quad z \in H_A(0)$$

Für das explizite Euler–Verfahren (d.h. $r = 0$) gilt: $H_A(0) = B_1(-1)$,
für das implizite Euler–Verfahren (d.h. $r = 1$): $H_A(0) = \mathbf{C} \setminus B_1(1)$.

Ein Verfahren heißt *stark absolut stabil (L–stabil)*, wenn gilt:

$$g(-z) \to 0 \quad \text{für} \quad \Re z \to -\infty.$$

Satz 6.1

$$\begin{aligned} A\text{–}stabil &\Leftrightarrow r \geq \frac{1}{2} \\ L\text{–}stabil &\Leftrightarrow r = 1 \end{aligned}$$

Ein explizites Runge–Kutta–Verfahren der Stufe l mit Konsistenzordnung l ist nicht A–stabil.

Definition: Sei Φ in einer Umgebung von $z = 0$ analytisch. $R_{jk}(z) = \frac{P_{jk}(z)}{Q_{jk}(z)}$, $z \in \mathbf{C}$, heißt *Padé–Approximation* von Φ zum Index (j,k), wenn P_{jk}, Q_{jk} Polynome sind mit $\text{grad}P_{jk} \leq k$, $\text{grad}Q_{jk} \leq j$ und $\Phi(z)Q_{jk}(z) - P_{jk}(z) = O(|z|^{j+k+1})$.

Satz 6.2 *Für $\Phi = \exp$ lautet die Padé–Approximation folgendermaßen:*

j \ k	0	1	2
0	1	$1+x$	$1+x+\frac{1}{2}x^2$
1	$\frac{1}{1-x}$	$\frac{1+\frac{1}{2}x}{1-\frac{1}{2}x}$	$\frac{1+\frac{2}{3}x+\frac{1}{6}x^2}{1-\frac{1}{3}x}$
2	$\frac{1}{1-x+\frac{1}{2}x^2}$	$\frac{1+\frac{1}{3}x}{1-\frac{2}{3}x+\frac{1}{6}x^2}$	$\frac{1+\frac{1}{2}x+\frac{1}{12}x^2}{1-\frac{1}{2}x+\frac{1}{2}x^2}$

(j,k) (1,0) (0,1) (1,1)
impliz. Euler expliz. Euler Crank–Nicholson
Genau für $j = k$, $j = k+1$, $j = k+2$ gilt: $|R_{jk}(z)| < 1$ für $z \to \infty$, $\Re z < 0$.
Für $j = k$ gilt: $|R_{jk}(z)| \to 1$ für $|z| \to \infty$, $\Re z < 0$.
Für $j = k+1$, $j = k+2$ gilt: $|R_{jk}(z)| \to 0$ für $|z| \to \infty$, $\Re z < 0$.

Satz 6.3 1. *Bei implizitem Runge–Kutta–Verfahren folgt: g existiert und ist rational mit Zähler-, Nennergrad $\leq$ Stufenzahl l.*

2. *Ist g diagonale oder eine der ersten beiden subdiagonalen Padé–Approximationen von* exp*, dann folgt A–Stabilität (subdiagonal $\Rightarrow$ L–Stabilität).*

3. *Bei l Stufen und Konsistenzordnung $2l$ folgt A–Stabilitaet.*

Sei $u_h^1, u_k^2, \ldots, u_h^N \in V_h$ die Lösung des volldiskreten Schemas mit Parameter $r \in [0,1]$.
Das volldiskrete Schema lautet wie folgt:

$$\begin{aligned} \frac{1}{\Delta t_k}\left\langle \left(u_h^{k+1} - u_h^k\right), v_h \right\rangle + a\left(r u_h^{k+1} + (1-r)u_k^h, v_h\right) &= \\ \langle r f(t_{k+1}) + (1-r) f(t_k), v_h \rangle \qquad &\forall v_h \in V_h \end{aligned} \tag{6.6}$$

Für die Abweichung von der exakten Lösung gilt:

$$u(t_k) - u_h^k = u(t_k) - R_h u(t_k) + \underbrace{R_h u(t_k) - u_h^k}_{=:-e_h^k \in V_h} \tag{6.7}$$

Mit Hilfe des volldiskreten Schemas und der Definition der Ritz-Projektion folgt:

$$\begin{aligned} \frac{1}{\Delta t_k}\left\langle e_h^{k+1} - e_h^k, v_h\right\rangle + a\left(r e_h^{k+1} + (1-r)e_h^k, v_h\right) \quad &= \qquad\qquad (6.8) \\ \langle r f(t_k) + (1-r) f(t_{k+1}), v_h\rangle - \frac{1}{\Delta t_k}\langle R_h u(t_{k+1}) - R_h u(t_k), v_h\rangle \quad &- \\ a\left(r u(t_{k+1}) + (1-r)u(t_k), v_h\right) \quad &=: \quad \left\langle \varepsilon_h^k, v_h\right\rangle . \end{aligned}$$

ε_h^k beschreibt also den lokalen Diskretisierungsfehler „in der Zeit“.
e_h^k beschreibt den gesamten Diskretisierungsfehler „in der Zeit“.
In den folgenden Sätzen 6.4 und 6.5 wird die Stabilität bzw. Konsistenz der Zeitdiskretisierung gezeigt, und die Fehler e_h^k bzw. ε_h^k werden nach oben abgeschätzt. Zusammen mit der bereits bekannten Abschätzung des Interpolationsfehlers der Ritzprojektion führen diese Abschätzungen zu einer Schranke für den Diskretisierungsfehler $|u(t_k) - u_h^k|$ (vergl. die Sätze 6.6 und 6.7).

Satz 6.4 *a sei V-elliptisch, stetig und $V \hookrightarrow\hookrightarrow H$. Dann gilt:*

1. *Wenn $\frac{1}{2} < r \le 1$, dann gibt es zu jedem $\Delta t_0 > 0$ Konstanten $\mu > 0$, $C > 0$ (nur abhängig von λ_1, r und Δt_0), so daß für alle $\Delta t \le \Delta t_0$ gilt:*

$$|e_h^k| \le |e_h^0| e^{-\mu t_k} + C\Delta t \sum_{j=0}^{k-1} e^{-\mu(t_k - t_j)} |\varepsilon_h^j| \tag{6.9}$$

2. *Für $r = \frac{1}{2}$ gilt für $\Delta t > 0$:*

$$|e_h^k| \le |e_h^0| + \Delta t \sum_{j=0}^{k-1} |\varepsilon_h^j| \tag{6.10}$$

3. *für $0 \le r < \frac{1}{2}$ gilt die Abschätzung* (6.10)*, falls die folgende Stabilitätsbedingung erfüllt ist:*

$$\Delta t\, \lambda_{\mu,h} \le \frac{2}{1-2r}, \tag{6.11}$$

wobei $\lambda_{\mu,h}$ der größte Eigenwert des diskreten Eigenwertproblems 3.4 *ist.*

Beweis: Betrachte das Eigenwertproblem in V_h, mit Eigenvektoren $w_{i,h}$:

$$e_i^k := \langle e_h^k, w_{i,h}\rangle, \qquad \varepsilon_i^k := \langle \varepsilon_h^k, w_{i,h}\rangle \qquad \text{(für festes } h\text{)}$$

Da die $w_{i,h}$ bzgl. $\langle.,.\rangle$ orthonormal sind, folgt:

$$|e_h^k| = \left(\sum_i (e_i^k)^2\right)^{1/2}, \qquad |\varepsilon_h^k| = \left(\sum_i (\varepsilon_i^k)^2\right)^{1/2}$$

Damit folgt unter Beachtung von (6.8):

$$\begin{array}{rl} & \frac{1}{\Delta t}(e_i^{k+1} - e_i^k) + \lambda_{i,h}(re_i^{k+1} + (1-r)e_i^k) = \varepsilon_i^k \quad \text{für } i = 1, \ldots, M(h) \\ \Rightarrow & e_i^{k+1} = g(\Delta t\,\lambda_{i,h})e_i^k + \frac{\Delta t}{1+r\Delta t\,\lambda_{i,h}}\varepsilon_i^k \end{array} \tag{6.12}$$

Sei $|g(\Delta t\,\lambda_{i,h})| \leq 1$, d.h. beliebiges $\Delta t > 0$ für $r \geq \frac{1}{2}$ bzw. $\Delta t\,\lambda_{i,h} \leq \frac{2}{1-2r}$ für $r < \frac{1}{2}$, dann gilt:

$$\begin{array}{rrcl} & |e_i^{k+1}| & \leq & |e_i^k| + \Delta t|\varepsilon_i^k| \\ \Rightarrow & |e_i^k| & \leq & |e_i^0| + \Delta t \sum_{j=0}^{k-1} |\varepsilon_i^j| \end{array}$$

Es gilt folgende Abschätzung (Dreiecksungl. für $\|.\|_2$ im $\mathbf{R}^2$):

$$\begin{array}{rrcl} & \left(\sum_{j=1}^n \left(\sum_{i=1}^m a_{ij}\right)^2\right)^{1/2} & \leq & \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2\right)^{1/2} \\ \Rightarrow & |e_h^k| = \left(\sum_{i=1}^M |e_i^k|^2\right)^{1/2} & \leq & \left(\sum_{i=1}^M \left(|e_i^0| + \Delta t \sum_{j=0}^{k-1} |\varepsilon_i^j|\right)^2\right)^{1/2} \\ & & \leq & |e_h^0| + \Delta t \sum_{j=0}^{k-1} |\varepsilon_h^j| \end{array}$$

Damit sind 2) & 3) gezeigt. $\square$

Satz 6.5 *Es existiert eine nur von r abhängige Konstante $C > 0$, so daß:*

1. *Für $r \neq \frac{1}{2}$ und $u \in C^1(0,T;V) \cap C^2(0,T;H)$ gilt:*

$$|\varepsilon_h^k| \leq \frac{1}{\Delta t_k} \int_{t_k}^{t_{k+1}} |(I - R_h)\frac{du}{dt}(s)|\,ds + C \int_{t_k}^{t_{k+1}} |\frac{d^2u}{dt^2}(s)|\,ds \tag{6.13}$$

2. *Für $r = \frac{1}{2}$ und $u \in C^1(0,T;V) \cap C^3(0,T;H)$ gilt:*

$$|\varepsilon_h^k| \leq \frac{1}{\Delta t_k} \int_{t_k}^{t_{k+1}} |(I - R_h)\frac{du}{dt}(s)|\,ds + C\,\Delta t \int_{t_k}^{t_{k+1}} |\frac{d^3u}{dt^3}(s)|\,ds \tag{6.14}$$

Beweis: Sei $u \in C^1(0,T;V)$ die Lösung des parabolischen Problems, d.h.:

$$\begin{aligned} \frac{d}{dt}\langle u(t), v\rangle + a\left(u(t), v\right) &= \langle f(t), v\rangle \qquad \forall v \in V \quad \forall t \in [0,T) \\ u(0) &= u_0 \end{aligned}$$

Durch Einsetzen in (6.8) folgt:

$$\begin{aligned} \langle \varepsilon_h^k, v_h\rangle &= \left\langle r\frac{du}{dt}(t_{k+1}) + (1-r)\frac{du}{dt}(t_k), v_h\right\rangle - \frac{1}{\Delta t}\left\langle R_h\left(u(t_{k+1} - u(t_k)\right), v_h\right\rangle \\ &= \left\langle r\frac{du}{dt}(t_{k+1}) + (1-r)\frac{du}{dt}(t_k) - \frac{1}{\Delta t}\left(u(t_{k+1} - u(t_k)\right), v_h\right\rangle \\ &\quad + \frac{1}{\Delta t}\left\langle (I - R_h)\left(u(t_{k+1} - u(t_k)\right), v_h\right\rangle \\ &= \left\langle r\frac{du}{dt}(t_{k+1}) + (1-r)\frac{du}{dt}(t_k) - \frac{1}{\Delta t}\left(u(t_{k+1} - u(t_k)\right), v_h\right\rangle \\ &\quad + \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} \left\langle (I - R_h)\frac{du}{dt}(s), v_h\right\rangle \, ds \end{aligned}$$

Für $v_h := \varepsilon_h^k$ folgt

$$|\varepsilon_h^k| \leq |\eta(t_k)| + \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} |(I - R_h)\frac{du}{dt}(s)| \, ds, \tag{6.15}$$

wobei

$$\eta(t_k) = \frac{1}{\Delta t}\left(u(t_{k+1}) - u(t_k)\right) - r\frac{du}{dt}(t_{k+1}) - (1-r)\frac{du}{dt}(t_k)$$

$\eta(t_k)$ wird durch Taylorentwicklung von u bei t_k abgeschätzt.
1. Fall: $r \neq \frac{1}{2}$ und $u \in C^2(0,T;H)$

$$u(t_{k+1}) = u(t_k) + \Delta t\frac{du}{dt}(t_k) + \int_{t_k}^{t_{k+1}} (t_{k+1} - s)\frac{d^2u}{dt^2}(s) \, ds$$

$$\begin{aligned} \Rightarrow \quad \eta(t_k) &= \frac{1}{\Delta t}\left(u(t_{k+1}) - u(t_k)\right) - r\left(\frac{du}{dt}(t_{k+1}) - \frac{du}{dt}(t_k)\right) - \frac{du}{dt}(t_k) \\ &= \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} (t_{k+1} - s)\frac{d^2u}{dt^2} \, ds - r\int_{t_k}^{t_{k+1}} \frac{d^2u}{dt^2}(s) \, ds \\ &= \frac{1}{\Delta t}\int_{t_k}^{t_{k+1}} \left(t_{k+1} - s - r(t_{k+1} - t_k)\right)\frac{d^2u}{dt^2} \, ds \end{aligned}$$

$t_{k+1} - s - r(t_{k+1} - t_k)$ nimmt sein Maximum in $s = t_k$ oder $s = t_{k+1}$ an, d.h.:

$$|\eta(t_k)| \leq max\{r, 1-r\}\int_{t_k}^{t_{k+1}} |\frac{d^2u}{dt^2}(s) \, ds. \tag{6.16}$$

2. Fall: $r = \frac{1}{2}$ und $u \in C^3(0,T;H)$: Analoges Vorgehen wie in Fall 1 liefert

$$|\eta(t_k)| \leq \frac{\Delta t}{8} \int_{t_k}^{t_{k+1}} |\frac{d^3 u}{dt^3}(s)\, ds. \tag{6.17}$$

Aus (6.15) und (6.16) folgt (6.13).
Aus (6.15) und (6.17) folgt (6.14). $\square$

Aus den Sätzen 6.4 und 6.5 erhält man nun die folgende Abschätzung des gesamten Diskretisierungsfehlers.

Satz 6.6 *a sei stetig und V-ellptisch, sowie $V \hookrightarrow\hookrightarrow H$. Dann gilt für die Näherungslösung $\{u_h^k \in V_h : \; 0 \leq k \leq K\}$:*

1. *Für $\frac{1}{2} < r \leq 1$ und $u \in C^1(0,T;V) \cap C^2(0,T;H)$ gibt es zu jedem $\Delta t_0 > 0$ zwei Konstanten $\mu > 0$, $C > 0$, die nur von $\lambda_{1,h}$ und Δt_0 abhängen, so daß für jedes $\Delta t \leq \Delta t_0$ gilt:*

$$\begin{aligned} |u_h^k - u(t_k)| \quad &\leq \quad |u_{0,h} - R_h u_0| e^{-\mu t_k} + |(I - R_h) u(t_k)| \\ &+ \int_0^{t_k} \left\{ \left| (I - R_h) \frac{du}{dt}(s) \right| + C \Delta t \left| \frac{d^2 u}{dt^2}(s) \right| \right\} e^{-\mu(t_k - s)}\, ds. \end{aligned} \tag{6.18}$$

2. *Für $r = \frac{1}{2}$ und $u \in C^1(0,T;V) \cap C^3(0,T;H)$ gilt:*

$$\begin{aligned} |u_h^k - u(t_k)| \quad &\leq \quad |u_{0,h} - R_h u_0| + |(I - R_h) u(t_k)| \\ &+ \int_0^{t_k} \left\{ \left| (I - R_h) \frac{du}{dt}(s) \right| + C \Delta t^2 \left| \frac{d^3 u}{dt^3}(s) \right| \right\} ds, \end{aligned} \tag{6.19}$$

   wobei C unabhängig von h und Δt ist.

3. *Für $0 \leq r < \frac{1}{2}$ und $u \in C^1(0,T;V) \cap C^3(0,T;H)$ gilt, wenn zusätzlich* (6.11) *erfüllt ist:*

$$\begin{aligned} |u_h^k - u(t_k)| \quad &\leq \quad |u_{0,h} - R_h u_0| + |(I - R_h) u(t_k)| \\ &+ \int_0^{t_k} \left\{ \left| (I - R_h) \frac{du}{dt}(s) \right| + C \Delta t \left| \frac{d^2 u}{dt^2}(s) \right| \right\} ds, \end{aligned} \tag{6.20}$$

   wobei C unabhängig von h und Δt ist.

Beweis: (nur für 1.; 2. und 3. folgen analog)
Aus (6.9) und (6.13) folgt:

$$\begin{aligned} \Delta t \sum_{j=0}^{k-1} e^{-\mu(t_k - t_j)} |\varepsilon_h^j| \quad &\leq \quad \sum_{j=0}^{k-1} e^{-\mu(t_k - t_j)} \int_{t_j}^{t_{j+1}} \left| (I - R_h) \frac{du}{dt}(s) \right| ds \\ &+ C \Delta t \sum_{j=0}^{k-1} e^{-\mu(t_k - t_j)} \int_{t_j}^{t_{j+1}} \left| \frac{d^2 u}{dt^2}(s) \right| ds \end{aligned} \tag{6.21}$$

Ferner:

$$\begin{aligned}
|u_h^k - u(t_k)| &\leq |u_h^k - R_h u(t_k)| + |R_h u(t_k) - u(t_k)| \\
&= |e_h^k| + |R_h u(t_k) - u(t_k)| \leq |e_h^0| e^{-\mu t_k} \\
&\quad + |R_h u(t_k) - u(t_k)| + C\Delta t \sum_{j=0}^{k-1} e^{-\mu(t_k - t_j)} |\varepsilon_h^j|
\end{aligned}$$

Aus (6.21) folgt nun zusammen mit

$$e^{-\mu(t_k - t_j)} \leq e^{-\mu(t_k - s)} \qquad \text{für } t_j \leq s, \quad 0 \leq j \leq K$$

die Behauptung. $\square$

Der räumliche Diskretisierungsfehler läßt sich unter den üblichen Konvergenzvoraussetzungen für FE-Diskretisierungen noch weiter abschätzen.
Die folgende Liste umfaßt die wichtigsten Voraussetzungen bei linearen Finiten Elementen:

1. $H_0^1(\Omega) \subset V \subset H^1(\Omega)$, $H = L^2(\Omega)$, $V \hookrightarrow\hookrightarrow H$
2. a sei V-elliptisch und symmetrisch
3. Die Triangulierung $\mathcal{T}_h$ sei regulär
4. Die finiten Elemente (K, P_K, Σ_h) seien affin-äquivalent zu $(\hat{K}, \hat{P}, \hat{\Sigma})$
5. Sei $l \in \mathbf{N}$ mit $P_l(\hat{K}) \subset \hat{P} \subset H^1(\hat{K})$
6. Die finiten Elemente seien aus der Klasse C^0

Satz 6.7 *1. Für $\frac{1}{2} < r \leq 1$ und $u \in C^1(0,T; H^{l+1}(\Omega)) \cap C^2(0,T; L^2(\Omega))$ für ein $l \in \mathbf{N}$ gibt es Konstanten $\mu > 0$, $C > 0$, die unabhängig von h und Δt_0 sind, so daß für alle $\Delta t \leq \Delta t_0$ gilt:*

$$\begin{aligned}
\|u_h^k - u(t_k)\|_{0,\Omega} &\leq \|u_{0,h} - u_0\|_{0,\Omega} e^{-\mu t_k} + C \Bigg\{ h^{\sigma(l)} \Bigg(\|u_0\|_{l+1,\Omega}\, e^{-\mu t_k} \\
&\quad + \|u(t_k)\|_{l+1,\Omega} + \int_0^{t_k} \|\frac{du}{dt}(s)\|_{l+1,\Omega}\, e^{-\mu(t_k - s)}\, ds \Bigg) \\
&\quad + \Delta t \int_0^{t_k} \|\frac{d^2u}{dt^2}(s)\|_{0,\Omega} e^{-\mu(t_k - s)}\, ds \Bigg\}.
\end{aligned}$$

2. Für $r = \frac{1}{2}$ und $u \in C^1(0,T; H^{l+1}(\Omega)) \cap C^3(0,T; L^2(\Omega))$ gilt:

$$\begin{aligned}
\|u_h^k - u(t_k)\|_{0,\Omega} &\leq \|u_{0,h} - u_0\|_{0,\Omega} + C \Bigg\{ h^{\sigma(l)} \Bigg(\|u_0\|_{l+1,\Omega} + \|u(t_k)\|_{l+1,\Omega} + \\
&\quad + \int_0^{t_k} \|\frac{du}{dt}(s)\|_{l+1,\Omega}\, ds \Bigg) + \Delta t^2 \int_0^{t_k} \|\frac{d^3u}{dt^3}(s)\|_{0,\Omega}\, ds \Bigg\}.
\end{aligned}$$

3. Für $0 \leq r < \frac{1}{2}$ *und* $u \in C^1(0,T;H^{l+1}(\Omega)) \cap C^3(0,T;L^2(\Omega))$ *gilt, wenn zusätzlich* (6.11) *erfüllt ist:*

$$\begin{aligned} \|u_h^k - u(t_k)\|_{0,\Omega} \;\leq\; & \|u_{0,h} - u_0\|_{0,\Omega} + C\Bigg\{ h^{\sigma(l)} \Bigg(\|u_0\|_{l+1,\Omega} + \|u(t_k)\|_{l+1,\Omega} + \\ & + \int_0^{t_k} \|\frac{du}{dt}(s)\|_{l+1,\Omega}\, ds \Bigg) + \Delta t \int_0^{t_k} \|\frac{d^2u}{dt^2}(s)\|_{0,\Omega}\, ds \Bigg\}, \end{aligned}$$

$$\textit{wobei} \quad \sigma(l) = \begin{cases} l+1 & , \ \textit{falls die zugehörige elliptische RWA regulär} \\ & \ \ \textit{(vergl. Aubin–Nitsche Lemma)} \\ l & , \ \textit{sonst} \end{cases} .$$

7 Bemerkungen zu konvektionsdominierten parabolischen partiellen Differentialgleichungen

7.1 Existenz von Lösungen

Wir betrachten eine parabolische PDG

$$\partial_t u + \nabla(qu) - \nabla(D\nabla u) = 0, \quad x \in [0, L].$$

In der zugehörigen schwachen Formulierung ergebe der diffusive Anteil $-\nabla(D\nabla u)$ eine symmetrische V-elliptische Bilinearform mit

$$\int_\Omega \nabla u D \nabla u \, dx \geq \alpha \|u\|_1^2 \quad \text{für ein } \alpha > 0 \text{ und alle } u \in V.$$

Der konvektive Anteil ist gegeben durch

$$\int_\Omega \nabla(qu) v \, dx, \quad u, v \in V.$$

Um diese PDG unserer Lösungstheorie zugänglich zu machen, müssen wir die V-Koerzivität nachweisen. Dazu schätzen wir ab:

$$\begin{aligned}
&\int_\Omega \nabla(q \cdot u) u \, dx = \\
&\quad = \int_\Omega \nabla q \cdot u \cdot u \, dx + \int_\Omega q \nabla u \cdot u \, dx \leq \\
&\quad \leq \int_\Omega \nabla q \cdot u \cdot u \, dx + \varepsilon \int_\Omega \|q\|_\infty |\nabla u|^2 \, dx + \frac{1}{\varepsilon} \int_\Omega \|q\|_\infty u^2 \, dx \qquad \forall \varepsilon > 0
\end{aligned}$$

$$\left[\text{Trick: } \left(\sqrt{\varepsilon}\, a - \frac{1}{\sqrt{\varepsilon}} b\right)^2 \geq 0 \Rightarrow ab \leq \varepsilon a^2 + \frac{1}{\varepsilon} b^2\right]$$

und erhalten so

$$\begin{aligned}
&\int_\Omega \nabla u D \nabla u \, dx - \int_\Omega \nabla(q \cdot u) u \, dx \geq \\
&\quad \geq -\int_\Omega \nabla q \, u^2 \, dx - \frac{1}{\varepsilon} \int_\Omega \|q\|_\infty u^2 \, dx + \int_\Omega \nabla u D \nabla u \, dx - \varepsilon \int_\Omega \|q\|_\infty |\nabla u|^2 \, dx \geq \\
&\quad \geq -\int_\Omega \nabla q \, u^2 \, dx - \frac{1}{\varepsilon} \int_\Omega \|q\|_\infty u^2 \, dx + \alpha \|u\|_1^2 - \varepsilon \int_\Omega \|q\|_\infty |\nabla u|^2 \, dx.
\end{aligned}$$

Für $\varepsilon < \alpha$ gilt $\tilde{\alpha} := \alpha - \varepsilon > 0$.Für ein

$$\lambda > \left| \alpha - \frac{1}{\varepsilon} \|\nabla q\|_\infty + \|q\|_\infty \right| + \tilde{\alpha}$$

können wir nun die V-Koerzivität nachweisen:

$$\int_\Omega (\nabla u D \nabla u - \nabla(q \cdot u) u)\, dx + \lambda |u|_0^2 \geq \tilde{\alpha} \|u\|_1^2.$$

Für geeignete ε können wir also unser Problem lösbar machen. Wir verlieren dabei aber für zu kleine ε an numerischer Stabilität:

7.2 Eindimensionales Modellproblem

Es sei

$$\partial_t u + \mathrm{Pe}\partial_x u - \partial_{xx} u = 0 \quad \text{für } x \in [0,1]$$

gegeben, mit der *Peclet-Zahl* $\mathrm{Pe} := \frac{qL}{D}$. (Wir haben so obiges Problem auf $[0,1]$ skaliert und die Diffusion auf 1 normiert.)
Im stationären Fall gilt:

$$\mathrm{Pe}\partial_x u = \partial_{xx} u.$$

Betrachte nun die Randbedingungen $u(0) = 1$ und $u(1) = 0$, dann kann die exakte Lösung als

$$u(x) = \frac{e^{\mathrm{Pe}} - e^{\mathrm{Pe}x}}{e^{\mathrm{Pe}} - 1}$$

angegeben werden. Approximieren wir das Problem mit linearen finiten Elementen oder mit der Finite-Differenzen-Methode und zentralen Differenzenquotienten für $\partial_x u$ (Finite-Volumen-Methode geht natürlich auch), dann gilt:

$$\begin{aligned} & \mathrm{Pe}\frac{u_{i+1} - i_{i-1}}{2\Delta x} = \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} \qquad (\mathbf{LGS}) \\ \Leftrightarrow \quad & -(\mu+1)u_{i-1} + 2u_i + (\mu - 1)u_{i+1} = 0 \\ & \text{mit } \mu := \frac{\mathrm{Pe}\Delta x}{2} \end{aligned}$$

$-(\mu+1)$ und $(\mu-1)$ sind nur dann zusammen kleiner null, wenn $\mu \leq 1$.
Die Lösung des linearen Gleichungssystems **(LGS)** ergibt sich als

$$\begin{aligned} u_i &= A + B\left(\frac{1+\mu}{1-\mu}\right)^{i-1} \qquad \text{für } i = 0, \ldots, N+1 \ (N \text{ Knoten}) \\ & A := -\left(\frac{1+\mu}{1-\mu}\right)^N \left(1 - \left(\frac{1+\mu}{1-\mu}\right)^N\right)^{-1} \\ & B := \left(1 - \left(\frac{1+\mu}{1-\mu}\right)^N\right)^{-1} \end{aligned}$$

Es sind zwei Fälle zu unterscheiden:

(a) $\mu < 1$: Lösung fällt mit wachsendem i, was dem Verhalten der exakten Lösung entspricht.

(b) $\mu > 1$: Die Lösung oszilliert, da $B < 0$ (sog. *Über- bzw. Unterschießen*) und somit B^n alterniert.

Um die Stabilität des numerischen Verfahrens auch für $\mu < 1$ zu gewährleisten gibt es verschiedene Ansätze:

7.2.1 Upwind-Ansatz in einer Raumdimension

Wir nehmen anstatt des symmetrischen Differenzenquotienten den „nach links gerichteten" für ∂_x:

$$\begin{aligned}
&\mathrm{Pe}\frac{u_i - u_{i-1}}{\Delta x} = \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} \\
\Leftrightarrow\quad & -(1 + \mathrm{Pe}\Delta x)u_{i-1} + (\mathrm{Pe}\Delta x + 2)u_i - u_{i+1} = 0 \\
\Rightarrow\quad & u_i = A + B(1 + \mathrm{Pe}\Delta x)^{i-1}, \quad i = 0, \ldots, N+1 \\
& \text{mit } A = \frac{-(1 + \mathrm{Pe}\Delta x)^N}{1 - (1 + \mathrm{Pe}\Delta x)^N}, \quad B = \frac{1}{1 - (1 + \mathrm{Pe}\Delta x)^N}
\end{aligned}$$

In diesem Fall ist die numerische Lösung stets monoton fallend (gut!). Leider ist sie etwas zu glatt:

Was ist der entscheidende Unterschied zum vorherigen Ansatz? Es gilt:

$$\mathrm{Pe}\frac{u_i - u_{i-1}}{\Delta x} = \mathrm{Pe}\frac{u_{i+1} - u_{i-1}}{2\Delta x} - \mathrm{Pe}\left(\frac{u_{i+1}}{2\Delta x} - \frac{u_i}{\Delta x} + \frac{u_{i-1}}{2\Delta x}\right)$$

Im Vergleich zum zentralen Differenzenquotienten haben wir also den Term

$$\frac{\mathrm{Pe}\Delta x}{2}\left(\frac{-u_{i-1} + 2u_i - u_{i+1}}{\Delta x^2}\right)$$

addiert. Es wird eine Art künstliche Diffusion oder Viskosität hinzugefügt. Dabei integrieren wir $(-u_{i-1} + 2u_i - u_{i+1})/\Delta x^2$ als Approximation für $-\partial_{xx}u$. Dies erklärt auch, warum das Problem nun stabil gelöst wird und warum die numerische Lösung glatter bzw. verschmierter ist. Eine schlichte Fortsetzung dieses Ansatzes im $\mathbf{R}^n$ durch Addition eines Termes vom Typ $(c \cdot \mathrm{Pe} \cdot \Delta x)\Delta u$ führt zu unerwünschter Diffusion bzw. Viskosität senkrecht zur Strömungsrichtung und verändert bzw. verschlechtert so das ursprüngliche Verfahren längs einer Richtung, in der keine Stabilitätsprobleme zu befürchten sind (*cross-wind diffusion*).

7.2.2 Analoge Ansätze für höhere Dimensionen

Petrov-Galerkin-Verfahren Bei diesem Verfahrenstyp wählt man verschiedene Ansatz- und Testräume $V_h \neq W_h$.

Eindimensionales Beispiel Lineare Elemente für V_h
Die Basisfunktionen in W_h seien lokal auf zwei Referenzelementen $[-1,0] \cup [0,1]$ gegeben durch

$$w(x) = N(x) \begin{cases} +\alpha F(x) & \text{für } x < 0 \\ -\alpha F(x) & \text{für } x > 0 \end{cases}$$

mit $N(x)$ und $F(x)$ wie folgt:

$$N(x) = \begin{cases} \frac{x - x_{i-1}}{\Delta x} & x_{i-1} < x < x_i \\ \frac{x_i - x}{\Delta x} & x_i < x < x_{i+1} \end{cases}$$

$$F(x) = \begin{cases} (x - x_{i-1})(x - x_i) & x_{i-1} < x < x_i \\ (x - x_i)(x - x_{i+1}) & x_i < x < x_{i+1} \end{cases}$$

(vgl. auch H.-G. Roos; 'Ten ways to generate the Il'in and related schemes'; Journal of Comp. Appl. Math 53(1994) 43-59)

Unsere Diskretisierung lautet dann:

$$\frac{\text{Pe}}{2\Delta x}\left(-(1+\alpha)u_{i-1} + 2\alpha u_i + (1-\alpha)u_{i+1}\right) = \frac{1}{\Delta x^2}\left(u_{i-1} + 2u_i - u_{i+1}\right)$$

Für $\alpha = 0$ erhalten wir den alten Differenzenquotienten, also ganz normale FE. $\alpha = 1$ ergibt das Upwind-Verfahren.
Es läßt sich zeigen, daß die numerische Lösung in den Knotenpunkten exakt(!) ist für

$$\alpha := \coth\left(\text{Pe}\frac{\Delta x}{2}\right) - \frac{1}{\left(\text{Pe}\frac{\Delta x}{2}\right)}.$$

Um Stabilität zu gewährleisten, müssen wir

$$\text{Pe}\Delta x \leq \frac{2}{1-\alpha}$$

fordern. Eine mögliche Approximation von $\coth(\mu) - \frac{1}{\mu}$ wäre zum Beispiel

$$\tilde{\alpha}(\mu) := \begin{cases} \frac{1}{3}\mu & \text{für } 0 \leq \mu \leq 1 \\ 1 & \text{sonst.} \end{cases}$$

Bemerkung Die optimale Wahl von α entspricht dem sog. *exponentiell angepaßten Upwind-Schema.* Für konstante Koeffizienten und eine Dimension heißt es auch *Il'in-Schema.*

Stromliniendiffusion (streamline diffusion) Dieses Verfahren läßt sich als Spezialfall eines Petrov-Galerkin-Ansatzes interpretieren:
Zu einem Ansatzraum V_h seien die Elemente des Testraums W_h durch

$$w_h := v_h + \delta \frac{q}{||q||} \nabla v_h, \quad v_h \in V_h$$

gegeben. Im eindimensionalen Fall erhält man:

$$w_h = v_h + \delta \operatorname{sgn}(q) \partial_x v_h.$$

Charakteristiken

Beispiel (instationär, $D = 0$)
Für $\partial_t u + q\partial_x u = 0$ ergeben sich die Lösungen als laufende Wellen

$$u(x,t) = \tilde{u}(x - qt),$$

denn

$$\partial_t \tilde{u} = -q\tilde{u}', \qquad \partial_x u = \tilde{u}'$$

und somit

$$\partial_t u + q\partial_x u = \partial_t \tilde{u} + q\partial_x \tilde{u} = -q\tilde{u}' + q\tilde{u}' = 0.$$

Setze $\eta := x - qt$, dann gilt:

$$\partial_\eta u = \partial_x u\, \eta_x + \partial_t u\, \eta_t = 0$$

und somit $u(\eta) = \text{const}$, also $u(x - qt) = \text{const}$.
Die Kurven $x = qt$ mit konstanten u-Werten nennt man *Charakteristiken.* Sie geben an, wie u-Werte transportiert werden:

Lagrange–Galerkin–Verfahren Betrachten wir nun auch Diffusion:

$$\partial_t u + q\partial_x u - D\partial_{xx} u = 0,$$

dann kann die Gleichung in einen Teil, der sich längs der Charakteristik bewegt, und einen diffusiven Anteil aufgespaltet werden:

$$\underbrace{\partial_\eta u}_{\text{Differenzenquotient}} - \underbrace{D\partial_{xx} u}_{\text{FE}} = 0$$

Der prinzipielle Ansatz besteht nun darin, den Transport längs einer Charakteristik durchzuführen (von der alten zur neuen Zeitschicht) und anschließend die Diffusion das transportierte u verschmieren zu lassen. Ein Problem ist z. B. dann gegeben, wenn die Punkte auf der alten und neuen Zeitschicht nicht durch Charakteristiken verbunden sind, d. h. der Transport erfolgt nicht genau von einem Diskretisierungspunkt zum nächsten, sondern trifft auf die Verbindungslinie zwischen zwei Diskretisierungspunkten.