Freie Universität Berlin

FB Mathematik und Informatik

Numerische Mathematik/Scientific Computing

Numerik von Differentialgleichungen (Numerik II)

Ralf Kornhuber und Christof Schütte

unter Mitarbeit von Heiko Berninger

2. Auflage: Wintersemester 2003/2004

Inhaltsverzeichnis

1 Woher kommen partielle Differentialgleichungen?

1.1 Variationsprinzip

1.1.1 Auslenkung einer Membran

Wir betrachten eine Membran (dünne Platte ohne Biegesteifigkeit), unter Belastung durch eine vertikal wirkende Kraft. Am Rand soll die Membran an einem ebenen Rahmen eingespannt sein. Die zur Beschreibung wesentlichen Größen sind

Auslenkung	:	u	(m)	
Kraftdichte	:	f	$(Newton/m^2)$	(wirkt in Richtung u).

Die Membran repräsentieren wir durch die Menge $\Omega \subset \mathbb{R}^2$. Im folgenden sei Ω immer ein *Gebiet*, d.h.

Ω offen und zusammenhängend

mit genügend glattem Rand (z.B. sei Ω Riemann–meßbar).
Gesucht ist die Auslenkung $u(x)$ in allen Punkten $x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2$ bei gegebener Kraftdichte f. Die Einspannung erzwingt zunächst die
Randbedingung:

$$\boxed{u(x) = 0 \qquad \forall x \in \partial\Omega \quad .}$$

Zur weiteren Charakterisierung von u verwenden wir das
Prinzip der minimalen Energie:
Die gesuchte Auslenkung u hat die Eigenschaft

$$\boxed{u \in H : \quad J(u) \leq J(v) \qquad \forall v \in H \quad .} \tag{1.1}$$

Dabei ist H die Menge der zulässigen Auslenkungen, und $J(v)$ ist die *Energie* einer Auslenkung $v \in H$.
Das *Energiefunktional* $J : H \to \mathbb{R}$ erhalten wir aus folgender *Energiebilanz*. Dabei sei vorläufig die Auslenkung v eine beliebige, genügend glatte Funktion auf Ω.

- Die *Spannungsenergie* $J_1(v)$ ist proportional zur Oberflächenänderung.

  Oberfläche von Ω : $\int_\Omega 1\, dx$

  Oberfläche nach Auslenkung v : $\int_\Omega \sqrt{1 + v_{x_1}^2 + v_{x_2}^2}\, dx$

  Insgesamt gilt also

$$\boxed{J_1(v) = \alpha \int_\Omega \left(\sqrt{1 + v_{x_1}^2 + v_{x_2}^2} - 1\right)\, dx \quad .}$$

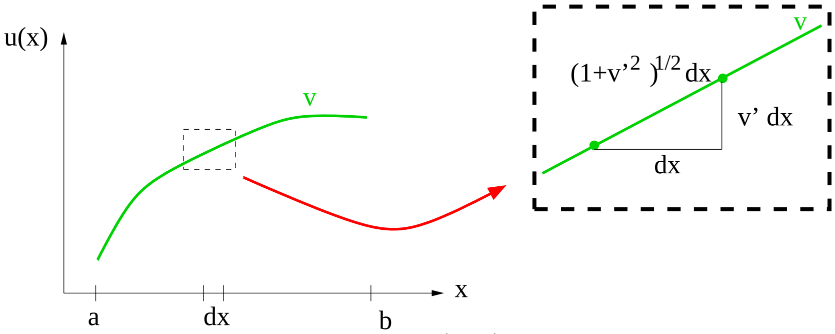


Abbildung 1: Berechnung der Bogenlänge $l(a,b)$ des Graphen einer Funktion $v = v(x)$ als $l(a,b) = \int_a^b \sqrt{1+v'(x)^2}\,dx$. v' bezeichnet die Ableitung $v' = v_x$. Zweidimensionale Verallgemeinerung ergibt obige Formel für die Oberfläche nach Auslenkung.

Die Materialkonstante $\alpha > 0$ heißt *Elastizität* von Ω .

Unter gewissen Bedingungen kann man diesen Ausdruck noch vereinfachen. Für *kleine Verzerrungen* ∇v ist nämlich

$$\sqrt{1+v_{x_1}^2+v_{x_2}^2}-1 \doteq 1+\frac{1}{2}(v_{x_1}^2+v_{x_2}^2)-1 = \frac{1}{2}(v_{x_1}^2+v_{x_2}^2) \quad .$$

Also ergibt sich

$$\boxed{J_1(v) \doteq \tfrac{1}{2}\int_\Omega \alpha\,|\nabla v|^2\,dx \quad .}$$

- Als *Potentielle Energie* (= Kraft $\times$ Weg) $J_2(v)$ erhält man

$$\boxed{J_2(v) = -\int_\Omega f v\,dx \quad .}$$

- Die Gesamtenergie $J(v)$ ist schließlich

$$\boxed{J(v) = \tfrac{1}{2}\int_\Omega \alpha\,|\nabla v|^2\,dx - \int_\Omega f v\,dx \quad .} \tag{1.2}$$

Den folgenden Funktionenräumen (lineare Räume über $\mathbb{R}$) werden wir in dieser Vorlesung häufig begegnen.

Definition 1.1 *Sei $\overline{\Omega}$ der Abschluß von Ω. Dann setzen wir:*

$$\begin{aligned}
C(\Omega) &:= \{v : \Omega \to \mathbb{R} \,|\, v \text{ stetig auf } \Omega\} \\
C(\overline{\Omega}) &:= \{v \in C(\Omega) \,|\, v \text{ stetig fortsetzbar auf } \overline{\Omega}\} \\
C^m(\Omega) &:= \{v \in C(\Omega) \,|\, \partial^\gamma v \in C(\Omega),\ |\gamma| \le m\} \\
C^m(\overline{\Omega}) &:= \{v \in C(\Omega) \,|\, \partial^\gamma v \in C(\overline{\Omega}),\ |\gamma| \le m\}
\end{aligned}$$

Dabei haben wird bequemerweise die höheren Ableitungen ∂^γ gemäß

$$\partial^\gamma = \frac{\partial^k}{\partial x_{\gamma_1} \dots \partial x_{\gamma_k}}, \quad \gamma_i = 1, 2, \quad |\gamma| = k$$

mit Hilfe von Multiindizes *γ geschrieben. Analoge Bezeichnungen verwenden wir im Falle $\Omega \subset \mathbb{R}^d$. Dann ist $\gamma_i \in 1, \dots, d$.*

An die *Daten* f machen wir die Voraussetzungen

$$f \in C(\overline{\Omega}) \quad .$$

Mit Blick auf die Gestalt des Energiefunktionals J liegt es nun nahe, als Lösungsraum H für unser Minimierungsproblem (1.0) den linearen Raum

$$\boxed{H_C := \{v \in C^1(\overline{\Omega}) \,|\, v|_{\partial\Omega} = 0\ \}}$$

zu wählen. Dann ist nämlich einerseits die Einspannbedingung erfüllt und andererseits ist

$$J(v) < \infty \qquad \forall v \in H_C \quad .$$

Es wird sich allerdings später zeigen, daß es im Sinne einer geschlossenen Existenztheorie vorteilhafter ist, Lösungen aus einem größeren Raum H zuzulassen.

Satz 1.2 (Variationsformulierung) *Eine Funktion $u \in H_C$ ist genau dann eine Lösung von* (1.0) *mit $H = H_C$, wenn sie der Variationsgleichung*

$$\boxed{\int_\Omega \alpha \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx \qquad \forall v \in H_C} \tag{1.3}$$

genügt.

Beweis:

Sei zunächst $u \in H_C$ eine Lösung des Minimierungsproblems (1.0), $v \in H_C$ beliebig aber fest gewählt und $t > 0$. Nun gilt $J(u) \le J(u + tv)$, da u das Funktional J minimiert, und daher

$$0 \le J(u + tv) - J(u) = t\Big(\int_\Omega \alpha \nabla u \cdot \nabla v \, dx - \int_\Omega f v \, dx\Big) + t^2 \frac{1}{2} \int_\Omega \alpha |\nabla v|^2 \, dx \quad .$$

Division durch t und Grenzübergang $t \to 0$ liefert

$$\int_\Omega \alpha \nabla u \cdot \nabla v \, dx \geq \int_\Omega f v \, dx \quad .$$

Die umgekehrte Abschätzung folgt analog durch Wahl von $t < 0$.
Genügt umgekehrt $u \in H_C$ der Variationsgleichung (1.2), so folgt für jedes beliebige $v \in H_C$

$$J(u+v) - J(u) = \int_\Omega \alpha \nabla u \cdot \nabla v \, dx - \int_\Omega f v \, dx + \frac{1}{2} \int_\Omega \alpha |\nabla v|^2 \, dx \geq 0 \quad .$$

□

Bemerkung: In Satz 1.2 ist nichts über Existenz oder Eindeutigkeit einer Lösung u gesagt.

Bemerkung: Für jedes feste $u \in H_C$ ist durch

$$J'(u)(v) = \int_\Omega (\alpha \nabla u \nabla v - f v) dx, \quad v \in H_C$$

eine lineare Abbbildung von H_C nach $\mathbb{R}$ definiert. Setzt man noch

$$\|v\|_1 = \left(\int_\Omega (v^2 + |\nabla v|^2) \, dx \right)^{1/2}$$

so folgt offenbar

$$|J(u+v) - J(v) - J'(u)(v)| = \frac{1}{2} \int_\Omega \alpha |\nabla v|^2 \, dx \leq \frac{1}{2} \alpha \|v\|_1^2 \quad .$$

Damit ist $J'(u)$ gerade die *Fréchèt-Ableitung* von J in u. Die Variationsformulierung (1.2) besagt gerade, daß das Minimum von J genau dann in u angenommen wird, wenn $J'(u)$ verschwindet.

Als nächstes wollen wir einen Zusammenhang zwischen der Variationsformulierung (1.2) und einer partiellen Differentialgleichung herleiten. Dazu brauchen wir

Satz 1.3 (Greensche Formel, partielle Integration) *Seien $v, w \in C^1(\overline{\Omega})$ und $\partial\Omega$ glatt (hinreichend: $\partial\Omega$ hat eine stetig differenzierbare Parametrisierung). Dann gilt*

$$\boxed{\int_\Omega v_{x_i} w \, dx = - \int_\Omega v w_{x_i} \, dx + \int_{\partial\Omega} v w \, n_i \, d\sigma \qquad i = 1, 2 \quad ,} \tag{1.4}$$

wobei $n = (n_1, n_2)$ die nach außen orientierte Normale von $\partial\Omega$ bedeutet.

Der Einfachheit halber nennen wir ein Gebiet, auf dem die Greensche Formel (1.3) gilt, *Greenschen Bereich.*

Bemerkung: Vergleiche die Produktregel

$$\int_a^b vw'\,dx = -\int_a^b v'w\,dx + vw|_a^b$$

für $v, w \in C^1[a,b]$.

Satz 1.4 *Sei $\partial\Omega$ ein Greenscher Bereich. Eine Funktion $u \in C^2(\overline{\Omega})$ ist genau dann Lösung des Minimierungsproblems* (1.0)*, wenn sie eine Lösung des Randwertproblems*

$$\begin{aligned} -\alpha\Delta u(x) &= f(x) \qquad \forall x \in \Omega \\ u(x) &= 0 \qquad \forall x \in \partial\Omega \end{aligned} \tag{1.5}$$

ist.

Beweis:

Sei $u \in C^2(\overline{\Omega})$ Lösung des Minimierungsproblems. Nach Satz 1.2 genügt dann u der Variationsgleichung (1.2). Anwendung der Greenschen Formel liefert

$$\int_\Omega (-\alpha\Delta u - f)v\,dx = 0 \qquad \forall v \in H_C \quad . \tag{1.6}$$

Im Widerspruch zur Behauptung nehmen wir an, daß es ein $x_0 \in \Omega$ gibt, so daß

$$-\alpha\Delta u(x_0) - f(x_0) > 0$$

gilt. Da Ω offen und $-\alpha\Delta u - f$ in Ω stetig ist, können wir dann eine offene Umgebung $U(x_0)$ von x_0 finden, so daß

$$-\alpha\Delta u(x) - f(x) > 0 \qquad \forall x \in U(x_0)$$

richtig ist. Nun lässt sich aber eine nichtnegative Funktion $v \in H_C$ mit den Eigenschaften

$$v(x) > 0 \quad \forall x \in \bar{K}_\varepsilon(x_0) \subset U(x_0), \quad v(x) = 0 \quad \forall x \in \Omega \setminus U(x_0)$$

konstruieren. Dabei ist $\bar{K}_\varepsilon(x_0)$ die abgeschlossene Kugel um x_0 mit Radius $\varepsilon > 0$. Einsetzen einer solchen Funktion v in (1.5) führt auf einen Widerspruch. Analog kann man im Fall $-\alpha\Delta u(x_0) - f(x_0) < 0$ argumentieren.

Ist umgekehrt $u \in C^2(\overline{\Omega})$ eine Lösung des Randwertproblems, so folgt durch Multiplikation der Differentialgleichung mit einem beliebigen $v \in H_C$, Integration über Ω und Anwendung der Greenschen Formel, daß u der Variationsgleichung (1.2) genügt.

□

Bemerkung: Die Differentialgleichung $-\Delta u = f$ heißt *Poisson–Gleichung.* In der Variationsrechnung nennt man (1.5) die *Eulersche Differentialgleichung* des Minimierungsproblems (1.0).

Bemerkung: Beachte, daß die Äquivalenz von Minimierungs- und Randwertproblem nur unter einer zusätzlichen *Regularitätsbedingung* vorliegt. Es kann also insbesondere sein, daß (1.0) eine (physikalisch sinnvolle!) Lösung hat, während (1.5) nicht lösbar ist.

1.1.2 Potentialgleichung

Auf einer beschränkten Teilmenge $\Omega \subset \mathbb{R}^2$ wollen wir zu einer gegeben *Ladungsdichte* f das *Potential* u ermitteln.

Elektrisches Potential	:	u	$(Volt)$
Ladungsdichte	:	f	(As/m^3)
angelegte Spannung	:	$u\vert_{\partial\Omega} = g$	$(Volt)$
Dielektrizitätskonstante	:	ε	$(Asm^2/Volt)$

Nach dem Prinzip der minimalen Energie und einer entsprechenden Energiebilanz ist die Lösung durch das **Minimierungsproblem**

$$\boxed{u \in H : \qquad J(u) = \int_\Omega \varepsilon \, |\nabla u|^2 \, dx - \int_\Omega f u \, dx \leq J(v) \qquad \forall v \in H}$$

charakterisiert. Unter geeigneten Regularitätsbedingungen an die angelegte Spannung g käme die Lösungsmenge $H = H_C$,

$$H_C := \{v \in C^1(\overline{\Omega}) \,|\, v|_{\partial\Omega} = g\}$$

in Frage. Beachte, daß diese Lösungsmenge kein linearer Raum ist. Ähnlich wie oben kann man eine **Variationsformulierung** herleiten (Wie?). Aus dieser erhält man unter einer zusätzlichen **Regularitätsbedingung** die zugehörige **Eulersche Differentialgleichung**

$$\boxed{-\operatorname{div}(\varepsilon \nabla u) = f \quad \text{ in } \Omega}$$

nebst Randbedingungen (Welche?). Im Falle $f = 0$ spricht man von der *Potentialgleichung.*

Bemerkung: Auch die Potentialgleichung haben wir nur unter gewissen Glattheitsbedingungen *a posteriori* hergeleitet. Die ursprüngliche Aufgabe war wieder ein Minimierungsproblem.

1.2 Erhaltungsprinzip

1.2.1 Massenerhaltung, Diffusion und Konvektion

Wir betrachten die Bewegung eines Fluids in einem gegebenen Rechengebiet $\Omega \subset \mathbb{R}^3$.

Dichte	:	ϱ	(kg/m^3)
Geschwindigkeit	:	v	(m/s)
Quelldichte	:	f	$(kg/(s\,m^3))$

Als Grundlage für eine mathematische Beschreibung dieser Situation verwenden wir das **Prinzip der Massenerhaltung:**
In jedem raumfesten *Kontrollvolumen* $\Omega' \subset \Omega$ gilt die *Massenbilanz*

$$\text{zeitliche Massenänderung} = -\text{ Massenabfluß} + \text{Quelle.}$$

Fließt Masse aus Ω' hinein, so haben wir einen negativen Massenabfluß, also Abfluß = – Zufluß. Analog wird eine Senke als negative Quelle interpretiert.
Sei also $\Omega' \subset \Omega$ ein solches Kontrollvolumen mit genügend glattem Rand $\partial\Omega'$. Wir wollen Massenänderung, Eintrag und Abfluß im Zeitraum von t bis $t+\Delta t$ durch die Funktionen ϱ, v und f ausdrücken.

- Zeitliche Massenänderung:

$$\int\limits_{\Omega'} \varrho(x, t + \Delta t)\; dx - \int\limits_{\Omega'} \varrho(x, t)\; dx \quad .$$

- Massenabfluß :

  Durch ein kleines Teilstück von $\partial\Omega'$ mit der Fläche $b = \Delta\sigma$ tritt in Normalenrichtung n (nach außen gerichtet!) Masse mit einer Geschwindigkeit $(n \cdot v)n$ aus. Die Geschwindigkeit v wird auf diesem Teilstück als konstant angenommen. Dann legt die austretende Masse in der Zeit Δt die Strecke

$$a = n \cdot v\; \Delta t$$

  zurück. Damit verlässt in diesem Zeitraum ein Quader mit dem Volumen

$$ab = n \cdot v\; \Delta\sigma\; \Delta t$$

  das Kontrollvolumen Ω' (siehe Abbildung (2).

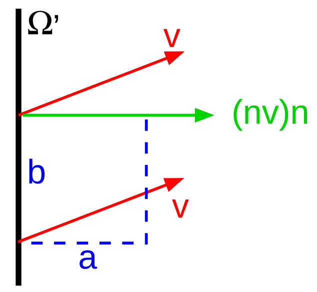


Abbildung 2: Fluß durch den Rand von Ω

Über dieses Oberflächenstück erfolgt also der Massenabfluß

$$\boxed{\varrho n \cdot v \, \Delta\sigma \, \Delta t \quad .}$$

Durch Summation und Grenzübergang ergibt sich der Massenabfluß über ganz $\partial\Omega'$ zu dem Oberflächenintegral

$$\Delta t \int\limits_{\partial\Omega'} \varrho n \cdot v \, d\sigma \quad .$$

- Quelle

  Ist die Quelldichte f über das Zeitintervall Δt konstant, so ist die in diesem Zeitraum bewirkte Massenänderung gegeben durch

$$\int\limits_{\Omega} f(x,t)\Delta t \, dx \quad .$$

Daraus ergibt sich für die Massenänderung in der Zeit Δt die Gesamtbilanz

$$\frac{1}{\Delta t}\left(\int\limits_{\Omega'} \varrho(x,t+\Delta t)\, dx - \int\limits_{\Omega'} \varrho(x,t)\, dx\right) = -\int\limits_{\partial\Omega'} \varrho v \cdot n \, d\sigma + \int\limits_{\Omega'} f \, dx \quad .$$

Der Grenzübergang $\Delta t \to 0$ führt auf die
Massenerhaltung in Integralform:

$$\boxed{\frac{\partial}{\partial t}\int_{\Omega'} \varrho \, dx + \int_{\partial\Omega'} \varrho v \cdot n \, d\sigma = \int_{\Omega'} f \, dx \qquad \forall \Omega' \subset \Omega \quad .} \tag{1.7}$$

Insbesondere gilt im stationären Fall

$$\int\limits_{\partial\Omega'} \varrho v \cdot n \, d\sigma = \int\limits_{\Omega} f \, dx \qquad \forall \Omega' \subset \Omega \quad . \tag{1.8}$$

Wir wollen nun die Massenerhaltungsgleichung in Differentialform, also als partielle Differentialgleichung schreiben. Entscheidendes Hilfsmittel ist wieder die Greensche Formel (Satz 1.3 auf Seite 4). Eine direkte Folgerung ist nämlich

Satz 1.5 (Divergenztheorem) *Sei Ω' ein Greenscher Bereich und $v = (v_1, v_2, v_3)^T$ ein Vektorfeld auf $\Omega' \subset \mathbb{R}^d$ mit $v_i \in C^1(\overline{\Omega'})$, $i = 1, 2, 3$, kurz $v \in C^1(\overline{\Omega'})$. Dann gilt*

$$\boxed{\quad \int_{\Omega'} \operatorname{div} v \, dx = \int_{\partial\Omega'} (v \cdot n) \, d\sigma \quad . \quad}$$

Satz 1.6 (Kontinuitätsgleichung) *Unter den* Regularitätsvoraussetzungen

$$(\varrho v)_i \in C^1(\Omega), \quad i = 1, 2, 3, \qquad \varrho_t \in C(\Omega), \qquad f \in C(\Omega)$$

genügen die Funktionen ϱ, v und f genau dann für alle Greenschen Bereiche $\Omega' \subset \Omega$ der Erhaltungsgleichung in Integralform (1.6)*, wenn sie die Kontinuitätsgleichung*

$$\boxed{\quad \varrho_t + \operatorname{div}(\varrho v) = f \qquad \text{in } \Omega \quad} \tag{1.9}$$

erfüllen.

Beweis:

Es gelte zunächst die Integralform (1.6). Im Widerspruch zur Behauptung sei für ein gewisses $x_0 \in \Omega$

$$\varrho_t(x_0) + \operatorname{div}(\varrho v)(x_0) - f(x_0) > 0 \quad .$$

Da Ω offen und $\varrho_t + \operatorname{div}(\varrho v) - f$ auf Ω stetig ist, existiert nun eine offene Kugel $K_\varepsilon(x_0)$, so daß $\overline{K_\varepsilon(x_0)} \subset \Omega$ und

$$\varrho_t(x) + \operatorname{div}(\varrho v)(x) - f(x) > 0 \qquad \forall x \in K_\varepsilon(x_0) \quad .$$

Wir wählen das Kontrollvolumen $\Omega' = K_\varepsilon(x_0)$. Anwendung des Divergenztheorems auf (1.6) liefert dann

$$\frac{\partial}{\partial t} \int_{\Omega'} \varrho \, dx + \int_{\Omega'} (\operatorname{div}(\varrho v) - f) \; dx = 0 \quad . \tag{1.10}$$

Nach Voraussetzung ist ϱ_t gleichmäßig stetig auf Ω' und damit die Vertauschung von Differentiation $\frac{\partial}{\partial t}$ und Integration erlaubt. Es ist also

$$\int_{\Omega'} (\varrho_t + \operatorname{div}(\varrho v) - f) \; dx = 0 \quad .$$

Nach Konstruktion ist der Integrand aber auf $\Omega' = K_\varepsilon(x_0)$ größer als 0. Das ist ein Widerspruch.

Die Umkehrung folgt direkt aus dem Divergenztheorem.

□

Bemerkung: Man beachte, daß Integral- und Differentialform nur unter *zusätzlichen Regularitätsbedingungen* äquivalent sind. Die Verwendung von (1.9) an Stelle von (1.6) bedeutet also eine Einschränkung. Es kann insbesondere (physikalisch sinnvolle!) Lösungen von (1.6) geben, obwohl die Differentialgleichung (1.9) keine Lösung hat.

In einer Vielzahl von Problemstellungen ist die Quelldichte f bekannt. Leider bleiben dann immer noch die unbekannten Funktionen v und ϱ übrig, die nicht in eindeutiger Weise aus nur einer Gleichung bestimmt werden können.
Zu dem (unstrittigen) Prinzip der Massenerhaltung treten daher zusätzliche **Zustands- oder Materialgleichungen**, welche den vorliegenden Massenfluß näher beschreiben. Die Form dieses Materialgleichungen ist eine Frage physikalischer *Modellierung*, die auf eine brauchbare mathematische Beschreibung des Gesamtproblems abzuzielen hat, aus mathematischer Sicht allerdings eine (zuerst einmal) willkürliche zusätzliche Festsetzung darstellt. Diese Festsetzung ist meist nur unter gewissen physikalischen Bedingungen gültig, im allgemeinen also mit Vorsicht zu genießen!
Wir führen eine dimensionslose *Konzentration* u ein und setzen

$$\varrho = \varrho_0 u \qquad\qquad \varrho_0 = \text{const.} > 0 \quad .$$

Die grundlegenden Transportprozesse sind *Diffusion* und *Konvektion*.
Diffusion: Ficksches Gesetz (1855)

$$\boxed{\varrho v = -\alpha \nabla u \qquad \alpha > 0}$$

Anschaulich: Der Massenfluß erfolgt proportional zum steilsten Abstieg der Konzentration.
Konvektion:

$$\boxed{\varrho v = \vec{\beta} u \ , \quad \vec{\beta} = \text{const.}}$$

Anschaulich: Die Ausbreitungsgeschwindigkeit $v = \vec{\beta}/\varrho_0$ ist vorgegeben.
Konvektion-Diffusion:

$$\boxed{\varrho v = -\alpha \nabla u + \vec{\beta} u}$$

Anschaulich: Beide Phänomene wirken zusammen.
Einsetzen in die Erhaltungsgleichung (1.8) ergibt

$$\boxed{\varrho_0 u_t = \operatorname{div}(\alpha \nabla u - \vec{\beta} u) + f \quad .} \tag{1.11}$$

Im allgemeinen sind α, $\vec{\beta}$ Funktionen von x und u! Im Falle $\alpha, \vec{\beta} = \text{const.}$ erhält man die **Konvektions-Diffusionsgleichung:**

$$\varrho_0 u_t = \alpha \Delta u - \vec{\beta} \nabla u + f \quad .$$

Bemerkung: Aus Erhaltung von Masse und Impuls nebst Zustandsgleichungen folgen die *Navier-Stokes-Gleichungen*

$$\begin{array}{rcl} -\varepsilon\Delta\vec{u} + (\vec{u}\cdot\nabla)\vec{u} + \nabla p &=& f \\ \nabla\vec{u} &=& 0 \end{array} \qquad \text{in } \Omega \quad .$$

Dabei bedeutet ε die Viskosität des Fluids, $\vec{u} = (u_1, u_2, u_3)^T$ die Geschwindigkeit und p den Druck. Der Laplace-Operator $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$ und $(\vec{u}\cdot\nabla) = u_1\frac{\partial}{\partial x_1} + u_2\frac{\partial}{\partial x_2} + u_3\frac{\partial}{\partial x_3}$ werden komponentenweise angewandt.
Unter den (schwerwiegenden!) Vereinfachungen:

- bekannte Flußrichtung $\vec{\beta}\cdot\nabla = \vec{u}\cdot\nabla$
- bekannter Druck p

erhält man eine Konvektions-Diffusionsgleichung für jede Komponente u_i.

Offenbar kann man nur dann eine eindeutige Lösung erwarten, wenn man Konzentration und/oder Massenfluß am Rand angemessen beschreibt. Die häufigsten Randbedingungen sind
Dirichlet-Randbedingungen:

$$\boxed{\varrho = \varrho_0 u = g \qquad \text{auf } \partial\Omega \quad .}$$

Anschaulich: Vorgabe der Dichte bzw. der Konzentration.
Neumann-Randbedingungen:

$$\boxed{\varrho v\cdot n = g \qquad \text{auf } \partial\Omega}$$

mit der nach außen gerichteten Normalen n auf $\partial\Omega$.
Anschaulich: Vorgabe des Massenabflußes.
Cauchy-Randbedingungen:

$$\boxed{\varrho v\cdot n = \alpha\varrho \qquad \text{auf } \partial\Omega \quad .}$$

Anschaulich: Der Massenabfluß ist proportional zur Konzentration.
Statt von Cauchy-Randbedingungen spricht man verschiedentlich auch von *Robin-Randbedingungen*. Außerdem heißen Dirichlet-, Neumann- und Cauchy-Randbedingungen auch *Randbedingungen 1., 2. und 3. Art*.

Bemerkung: Natürlich kann man auch verschiedene Randbedingungen auf verschiedenen Teilstücken von $\partial\Omega$ vorgeben.

Neben den Standard-Typen von Randbedingungen können auch andere Randbedingungen sinnvoll sein. Als Beispiel für eine *nichtlokale Randbedingung* sei die *Massenerhaltung*

$$\int_{\partial\Omega} \varrho v \cdot n \, d\sigma = 0$$

erwähnt.

Bemerkung: Ähnlich wie Zustandsgleichungen sollen Randbedingungen eine *spezielle physikalische Situation* beschreiben. *Mathematisch* sinnvolle Randbedingungen sichern zusammen mit den vorliegenden (Differential)-Gleichungen Existenz und Eindeutigkeit einer Lösung. Die Wahl physikalisch und mathematisch sinnvoller Randbedingungen ist in vielen praktischen Anwendungen alles andere als einfach.

Natürlich hängt die Konzentration $u(x,t)$ von der Situation zum Anfangszeitpunkt $t = 0$ ab. Damit wir eine eindeutige Lösung u erwarten können müssen wir also noch eine **Anfangsbedingung:**

$$\boxed{u(x,0) = u_0(x) \qquad x \in \Omega \quad .}$$

festlegen.

1.2.2 Energieerhaltung (Wärmeleitungsgleichung)

Außer der Masse können auch noch andere Größen wie zum Beispiel die Energie erhalten bleiben. Als Beispiel betrachten wir den Wärmefluß in einem Rechengebiet $\Omega \subset \mathbb{R}^3$, der durch die folgenden Größen beschrieben werden soll.

Energiedichte : $\mathcal{E}$ $\quad Joule/(m^3)$

Wärmefluß : q $\quad Joule/(m^2 sec)$

Wärmequelle : f $\quad Joule/(m^3 sec)$

Genau wie im Abschnitt 1.2.1 erhält man die
Energieerhaltung in Integralform:

$$\boxed{\frac{\partial}{\partial t}\int_{\Omega'} \mathcal{E} \, dx + \int_{\partial\Omega'} q \cdot n \, d\sigma = \int_{\Omega'} f \, dx \qquad \forall \Omega' \subset \Omega \quad .}$$

Unter geeigneten Regularitätsvoraussetzungen (vgl. Satz 1.6) ist damit gleichbedeutend die **Kontinuitätsgleichung:**

$$\boxed{\mathcal{E}_t + \operatorname{div} q = f \qquad \text{in } \Omega \quad .}$$

Als erste **Zustandsgleichung** postulieren wir

$$\boxed{\mathcal{E} = \mathcal{E}(\theta) = c\theta,}$$

wobei θ die Temperatur und $c > 0$ die spezifische Wärmekapazität bedeutet. Einen (einfachen, aber z.B. in der Nähe des absoluten Nullpunkts sicher falschen!) Zusammenhang zwischen Wärmefluß und Temparatur liefert die nächste Zustandsgleichung.

Fouriersches Gesetz (1822)

$$\boxed{q = -\kappa\nabla\theta \quad .}$$

Der Proportionalitätsfaktor $\kappa > 0$ heißt *Wärmeleitfähigkeit*. Das Fouriersche Gesetz beschreibt die Wärmeausbreitung durch **Diffusion**. Einsetzen in die Differentialform ergibt die

Wärmeleitungsgleichung:

$$\boxed{\tfrac{\partial}{\partial t}(c\theta) = \operatorname{div}(\alpha\nabla\theta) + f \quad .}$$

Im allgemeinen sind c, κ Funktionen von x und u. Sind diese Funktionen konstant, so vereinfacht sich die Wärmeleitungsgleichung zu

$$\theta_t = \gamma\Delta u + g \quad \text{mit} \quad \gamma = \frac{\kappa}{c},\ g = \frac{f}{c} \quad .$$

Wir betrachten noch kurz die Wärmeleitung in zeitlich bewegtem Medium, d.h.

$$x = x(t), \quad x \in \Omega \quad .$$

Die Geschwindigkeit an der Stelle $x = x(t)$ ist dann $\vec{\beta} = \dot{x}(t)$. Mit der Kettenregel erhält man in diesem Fall eine **Konvektions-Diffusionsgleichung**

$$\frac{\partial}{\partial t}\theta(x(t), t) = \vec{\beta} \cdot \nabla\theta + \theta_t = \gamma\Delta\theta + g \quad .$$

Beachte, daß die *Konvektion* $\vec{\beta} \cdot \nabla\Theta$ durch die Bewegung des Mediums hervorgerufen wird.

Wieder hat man noch *Rand-* und *Anfangsbedingungen* festzulegen. Welche physikalische Interpretation haben Dirichlet- Neumann- und Cauchy-Randbedingungen in diesem Fall?

1.2.3 Grundwasserfluß im gesättigten Boden

Den Grundwasserfluß durch ein Gebiet $\Omega \subset \mathbb{R}^3$ kann man durch folgende Größen beschreiben.

Dichte	:	ϱ	kg/m^3
Porosität	:	n	dimensionslos
Sättigung	:	S	dimensionslos
Abstandsgeschwindigkeit	:	v	m/sec
Filtergeschwindigkeit	:	v_F	m/sec
Druck	:	p	$Newton/m^2$
Quelldichte	:	f	$(kg/m^3\,sec)$

Die *Porosität* $n(x)$, $x \in \Omega$ beschreibt den Anteil des *Porenraums* an einem (infinitesimalen) Referenzvolumen. Die *Sättigung* $S(x,t)$ beschreibt den Anteil des Fluids am Porenraum. Dann lautet die

Massenerhaltung in Differentialform:

$$(Sn\varrho)_t + \operatorname{div}(Sn\varrho v) = f \quad . \tag{1.12}$$

Wieder sind eine Reihe von Zustandsgleichungen nötig. Wir wollen einen *gesättigten* Fluß betrachten, also

$$S(x,t) \equiv 1 \qquad \forall x \in \Omega, \quad t > 0 \quad .$$

Weiter sollen Temparaturunterschiede keine Rolle spielen. Dann ist

$$\varrho(x,t) \equiv \varrho_0 = \text{const.} \qquad \forall x \in \Omega, \quad t > 0,$$

denn Wasser ist *inkompressibel*. Die Porosität ist eine Funktion des Drucks, also

$$\varrho_0 \frac{\partial(n(p))}{\partial t} = \varrho_0 \frac{dn}{dp} p_t$$

Wir nehmen nun an, daß

$$\varrho_0 \frac{dn}{dp} = \frac{S_0}{g} = \text{const.}, \tag{1.13}$$

wobei g die *Erdbeschleunigung* (Einheit m/sec^2) und S_0 den *spezifischen Speicherkoeffizienten* (Einheit $1/m$) bezeichnen. Durch ausführliche Beobachtung der Brunnen von Dijon fand Darcy eine Zustandsgleichung für die Filtergeschwindigkeit.

Darcysches Gesetz (1856):

$$\boxed{v_F = -K\nabla h}$$

Die symmetrische 3×3–Matrix K mit den Spalten K_i, $i = 1, \ldots, 3$, heißt *hydraulische Leitfähigkeit* und beschreibt die Durchlässigkeit des Bodens. Die Funktion h heißt *Standrohrspiegelhöhe* und kann als Stand des Grundwassers im Bohrloch beobachtet werden. Es gilt

$$\boxed{h = \frac{p}{\varrho_0 g} + x_3 \quad .}$$

Schließlich nehmen wir an, daß die (mikroskopische) Abstandsgeschwindigkeit v mit der (makroskopischen) Filtergeschwindigkeit v_F gemäß der Beziehung

$$v_F = nv$$

korreliert. Einsetzen all dieser Beziehungen in (1.11) ergibt schließlich die sogenannte **Darcy-Gleichung:**

$$\boxed{S_0 p_t = \operatorname{div}(K \nabla p) + \varrho_0 g \operatorname{div} K_3 + gf \quad .} \tag{1.14}$$

Bemerkung: Anstelle der in Zustandsgleichung (1.12) beschriebenen Situation kann auch ein *druckstabiles Korngerüst* vorliegen, also

$$\frac{dn}{dp} = 0 \; .$$

In diesem Fall reduziert sich die Darcy–Gleichung (1.13) auf

$$-\operatorname{div}(K \nabla p) = \varrho_0 g \operatorname{div} K_3 + gf \; .$$

Bemerkung: Wir sind von $S(x,t) \equiv 1$ und $K = K(x)$ ausgegangen. Im Falle *ungesättigter Strömungen* ist die Sache komplizierter, nämlich $S(x,t) \neq 1$ und $K = K(p,x)$. Um die unbekannte Sättigung S aus der Differentialgleichung zu eliminieren, verwendet man in diesem Fall aus Messungen abgeleitete Druck–Sättigungsbeziehungen und erhält eine *nichtlineare* Differentialgleichung, die sog. *Richardsgleichung*.
Vorgegebene Druck–Sättigungsbeziehungen gelten nur unter gewissen Bedingungen. Sind diese Bedingungen nicht erfüllt, so muß man zu noch komplexeren Modellen übergehen, die aus zwei getrennten nichtlinearen Differentialgleichungen für Druck und Sättigung bestehen. Bei solchen *Mehrphasenströmungen* sind wir schon recht nah an der aktuellen Forschung.

Nun fehlen noch geeignete **Randbedingungen**. Welche physikalische Bedeutung haben in diesem Fall Dirichlet-, Neumann- und Cauchy-Randbedingungen?

1.2.4 Impulserhaltung (Wellengleichung)

Bislang stand die Kontinuitätsgleichung im Mittelpunkt. Zum Schluß das
Prinzip der Impulserhaltung:
Die Zunahme an Impuls in einem materiellen Kontrollvolumen ist gleich der daran von außen angreifenden Kraft.
Übrigens spiegelt sich die Impulserhaltung auch in den Maßeinheiten wieder (Erinnerung: $1 Newton = kg\ m/sec^2$) Für jedes Kontrollvolumen $\Omega' \subset \Omega \subset \mathbb{R}^d$ (mit genügend glattem Rand) lässt sich dieses physikalische Grundgesetz direkt als Formel schreiben,

$$\boxed{\frac{\partial}{\partial t}\int_{\Omega'} \varrho v\, dx = \int_{\Omega'} f\, dx + \int_{\partial\Omega'} \sigma \cdot n ds \quad .}$$

Dabei haben wir die folgenden Bezeichnungen verwendet.

Dichte	:	ϱ	kg/m^d
Geschwindigkeit	:	v	m/sec
Volumenkraftdichte	:	f	$Newton/m^d$
Spannungstensor	:	σ	$Newton/m^{d-1}$

Wieder kann man diese Bilanzgleichung (unter zusätzlichen Regularitätsbedingungen) in eine Differentialgleichung umformulieren (in welche?).
Wir wollen das Prinzip der Impulserhaltung nun zur Beschreibung einer schwingenden Saite (keine Gravitation, keine Biegesteifigkeit) verwenden. Die Saite soll in a und b eingespannt sein und uns interessiert zu jedem Zeitpunkt $t > 0$ ihre

$$\boxed{\textbf{Auslenkung:} \quad u : [a, b] \to \mathbb{R}.}$$

In diesem (eindimensionalen) Fall erhält die Impulsbilanz die Gestalt

$$\frac{\partial}{\partial t}\int_{x_0}^{x_1} \varrho v\, dx = \int_{x_0}^{x_1} f\, dx + \sigma(x_0) - \sigma(x_1) \qquad \forall x_0, x_1 \quad a \le x_0 < x_1 \le b$$

Wir haben nun ϱ, v, f und σ festzulegen. Gravitationskräfte sollen vernachlässigt werden, also ist

$$f \equiv 0\ .$$

Bei vertikalen Schwingungen gilt

$$\boxed{v = u_t.}$$

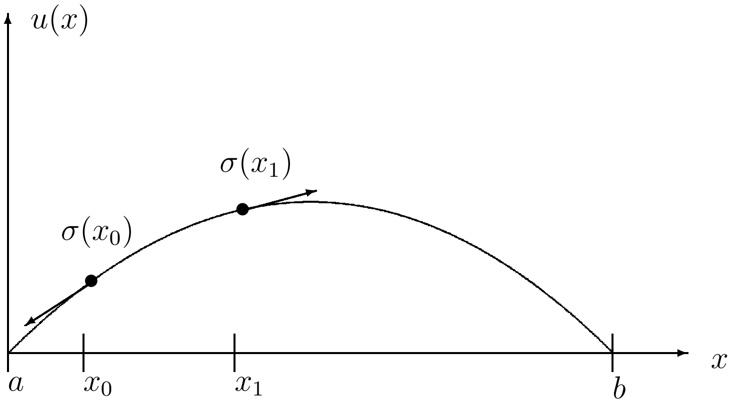


Abbildung 3: Schwingende Saite

Für die Spannungen $\sigma(x_0)$ und $\sigma(x_1)$ benötigen wir wieder eine *Zustandsgleichung*. Dazu definieren wir die Funktion $L : [a,b] \to \mathbb{R}$ durch

$$L(x) = \text{Länge der Kurve } \Gamma(x) = \{(s, u(s)) \in \mathbb{R}62 \mid s \in [a,x]\} \ .$$

Die Längenänderung $\frac{dL}{dx}$ ist bekanntlich

$$\frac{dL}{dx}(x) = \int\limits_0^x \sqrt{1 + u_x^2(s)} - 1 \ ds = \sqrt{1 + u_x^2(x)} - 1 \ .$$

Für eine *elastische Saite* gilt das **Hookesche Gesetz:**

$$\boxed{\sigma = \alpha \frac{dL}{dx} \quad .}$$

Der Proportionalitätsfaktor $\alpha > 0$ heisst Elastizität und ist materialabhängig und damit i.a. ortsabhängig. Wir wollen uns auf *kleine Verzerrungen* $u_x \approx 0$ beschränken. Dann gilt

$$\frac{dL}{dx} \doteq \tfrac{1}{2} u_x \ .$$

Aus dem Hooke'sche Gesetz folgt dann die *lineare Zustandsgleichung*

$$\boxed{\sigma = \tfrac{1}{2} \alpha u_x \quad .}$$

Einsetzen all dieser Beziehungen liefert nun

$$\frac{\partial}{\partial t}\int\limits_{x_0}^{x_1} \varrho u_t\, dx = \tfrac{1}{2}\big(\alpha(x_1)u_x(x_1) - \alpha(x_0)u_x(x_0)\big) \qquad \forall x_0, x_1 \quad a < x_0 < x_1 < b$$

Unter geeigneten *Regularitätsbedingungen* an die Lösung u und die Daten ϱ, α erhalten wir in gewohnter Weise die
Wellengleichung:

$$\boxed{(\varrho u_t)_t = \tfrac{1}{2}(\alpha u_x)_x \qquad \text{in } (a,b).}$$

Im Falle $\varrho, \alpha = \text{const.}$ vereinfacht sich die Wellengleichung mit $\gamma = \frac{\alpha}{2\varrho}$ zu

$$\boxed{u_{tt} - \gamma u_{xx} = 0 \qquad \text{in } (a,b)\ .}$$

Dazu treten zum Beispiel **Dirichlet - Randbedingungen:**

$$u(a,t) = u_b(t) \qquad\qquad u(b,t) = u_b(t) \qquad t > 0 \quad .$$

Übrigens: Welche physikalische Bedeutung haben Neumann - Randbedingungen? Sind Cauchy - Randbedingungen physikalisch sinnvoll?
Natürlich hängt die Bewegung der Saite entscheidend davon ab, wie die Auslenkung zum Zeitpunkt $t = 0$ aussieht und welche Geschwindigkeit sie zu diesem Zeitpunkt hat. Um eine eindeutige Lösung erwarten zu können, werden wir also voraussichtlich noch die
Anfangsbedingungen:

$$u(x,0) = u_0(x) \qquad\qquad u_t(x,0) = u_1(x) \qquad x \in (a,b)$$

festlegen müssen.

1.2.5 Verkehrsfluß (Autoerhaltung)

Verkehrsdichte :	ϱ	Autos/m
Geschwindigkeit	v	m/sec

Wir machen gleich am Anfang die *Kontinuumshypothese*

$$\varrho : [a,b] \to \mathbb{R} \quad .$$

Diese ist trotz steigenden Verkehrsaufkommens im vorliegenden Fall zweifelhaft. Schon deshalb sollte man das folgende Modell nicht zu ernst nehmen. Analog zur Massenerhaltung folgt nun die
Autoerhaltungsgleichung:

$$\boxed{\varrho_t + (\varrho v)_x = 0}$$

Offenbar gilt

$$0 \leq \varrho(x) \leq \varrho_{max} = \text{const.}$$

Im Falle $\varrho = \varrho_{max}$ stehen die Autos Stoßstange an Stoßstange. Weiter gehen wir von halbwegs vernünftiger Fahrweise aus, also von

$$0 \leq v(x) \leq v_{max} = \text{const.}$$

Interpoliert man die Extremfälle $v = 0$ (im Falle $\varrho = \varrho_{max}$) und $v = v_{max}$ (im Falle $\varrho = 0$) einfach linear, so erhält man die **Zustandsgleichung**

$$v(\varrho) = v_{max}(1 - \varrho/\varrho_{max}) \quad .$$

Einsetzen ergibt eine
Nichtlineare Erhaltungsgleichung:

$$\boxed{\varrho_t + v_{max}(\varrho(1 - \varrho/\varrho_{max}))_x = 0 \quad .}$$

2 Elliptisch, parabolisch, hyperbolisch

2.1 Das Cauchy-Problem

Wir betrachten die folgende Differentialgleichung *2. Ordnung* in den zwei Variablen x und y

$$\boxed{a u_{xx} + 2b u_{xy} + c u_{yy} = d.} \tag{2.1}$$

Die Differentialgleichung (2.0) heißt

$$\begin{array}{lll} \textit{quasilinear}, & \text{falls} & a,b,c,d = a,b,c,d(x,y,u,u_x,u_y) \\ \textit{semilinear}, & \text{falls} & a,b,c = a,b,c(x,y),\ d = d(x,y,u,u_x,u_y) \\ \textit{linear}, & \text{falls} & a,b,c,d = a,b,c,d(x,y). \end{array}$$

In jedem Fall sollen a, b, c nicht gleichzeitig verschwinden (2. Ordnung!). Beachte, daß Linearkombinationen von Lösungen nur im linearen Fall wieder Lösungen sind.
Wir wollen uns nun mit der Existenz von Lösungen von (2.0) beschäftigen. Wir haben schon gesehen, daß Eindeutigkeit nur unter zusätzlichen Bedingungen zu erwarten ist. Da die Differentialgleichung von zweiter Ordnung ist, liegen Bedingungen an Funktion und erste Ableitungen nahe, also

$$\begin{aligned} u|_\gamma &= u_0 \\ u_x|_\gamma &= g_1 \\ u_y|_\gamma &= g_2 \end{aligned}$$

Dabei sei $\gamma = (\gamma_1, \gamma_2)$ eine glatte Kurve in $\mathbb{R}^2$ (genauer: $\gamma_1,\ \gamma_2 \in C^\infty(I)$, $I \subset \mathbb{R}$, Intervall) mit Tangente $\gamma' = (\gamma_1', \gamma_2')$ der Länge $\sqrt{(\gamma_1')^2 + (\gamma_2')^2} > 0$.
Näheres Hinschauen zeigt, daß die Lösung durch diese Bedingungen überbestimmt ist. Es gilt nämlich die *Kompatibilitätsbedingung*

$$\frac{d}{ds} u_0(\gamma_1(s), \gamma_2(s)) = g_1 \gamma_1' + g_2 \gamma_2' \quad .$$

Es sind also nur **zwei Bedinungen frei wählbar**![1] Also sollten wir z.B. nicht Bedingungen für $u_x|_\gamma$ und $u_y|_\gamma$, sondern nur für die Normalableitung

$$\frac{\partial}{\partial n} u \;=\; \nabla u \cdot (\gamma')^\perp = \frac{-u_x \gamma_2' + u_y \gamma_1'}{\sqrt{{\gamma_1'}^2 + {\gamma_2'}^2}}$$

bzgl. γ angeben. Wir betrachten daher das *Cauchy-Problem*

[1] Herleitung durch Differenzieren laut Kettenregel und Einsetzen von $u_x|_\gamma = g_1$ und $u_y|_\gamma = g_2$.

$$\boxed{\begin{aligned} au_{xx} + 2bu_{xy} + cu_{yy} &= d \quad \text{in } \mathbb{R}^2 \\ u &= u_0 \quad \text{auf } \gamma \\ \frac{\partial}{\partial n} u &= u_1 \quad \text{auf } \gamma \end{aligned} \qquad (2.2)}$$

Dabei sollen die Bedingungen auf γ stetig angenommen werden, also

$$\lim_{(x,y)\to\gamma(s_0)} u(x,y) = u_0(\gamma(s_0)) \quad .$$

Beachte, daß sich mit Hilfe der Kompatibilitätsbedingung die partiellen Ableitungen $u_x|_\gamma$, $u_y|_\gamma$ aus Funktionswert und Normalenableitung berechnen lassen.
Es wäre eine (gute!) Idee für einen Existenzbeweis, die Lösung u in einer Umgebung von γ in eine Taylorreihe zu entwickeln, also

$$u(x,y) = u(\gamma(s_0)) + \sum_{k=1}^{\infty} \sum_{|\beta|=k} \frac{1}{\beta!} \partial^\beta u(\gamma(s_0)) h^\beta \qquad \forall x, y \in K_\varepsilon(\gamma(s_0)) \; .$$

Dabei ist $\beta = (\beta_1, \ldots, \beta_k)$, $\beta_i \in \{x, y\}$, $i = 1, \ldots, k$, ein Multiindex mit $|\beta| = k$,

$$\partial^\beta = \partial_{\beta_1} \cdots \partial_{\beta_k}, \qquad h^\beta = h_{\beta_1} \cdots h_{\beta_k}, \quad h_x = x - \gamma_1(s_0), h_y = y - \gamma_2(s_0)$$

und $\beta! = \beta_x!\beta_y!$ gesetzt. Schließlich bezeichnet β_x bzw. β_y die Anzahl der Indizes $\beta_i = x$, bzw. $\beta_i = y$.
Zur Bestimmung von u brauchen wir also die höheren Ableitungen $\partial^\beta u$ von u auf γ. Die Ableitungen erster Ordnung, also u_x, u_y, haben wir schon. Jetzt geht es um die Ableitungen zweiter Ordnung. Die Kettenregel liefert

$$\begin{aligned} \frac{d}{ds} u_x &= \gamma_1' u_{xx} + \gamma_2' u_{xy} \\ \frac{d}{ds} u_y &= \gamma_1' u_{xy} + \gamma_2' u_{yy} \quad , \end{aligned}$$

und es gilt ja

$$au_{xx} + 2bu_{xy} + cu_{yy} = d \quad .$$

Unter der Bedingung

$$\boxed{\begin{vmatrix} \gamma_1' & \gamma_2' & 0 \\ 0 & \gamma_1' & \gamma_2' \\ a & 2b & c \end{vmatrix} = a{\gamma_2'}^2 - 2b\gamma_1'\gamma_2' + c{\gamma_1'}^2 \neq 0 \qquad (2.3)}$$

hat dieses lineare Gleichungssystem für jede rechte Seite die eindeutig bestimmte Lösung u_{xx}, u_{xy}, u_{yy}. Sind die Ableitungen zweiter Ordnung bekannt, so können wir durch weiteres Differenzieren alle partiellen Ableitungen von $\partial^\beta u(\gamma(s_0))$, $|\beta| = 3, 4, \ldots$, an jeder Stelle $\gamma(s_0) \in \gamma$ berechnen (Übung).
Die Eigenschaft (2.2) von γ ist offenbar von gewisser Bedeutung und wir geben ihr vorsorglich einen Namen.

Definition 2.1

Eine Kurve $\gamma(s)$ heißt charakteristisch oder auch Charakteristik in $s \in I$, falls

$$\boxed{\det(s) := a\gamma_2'(s)^2 - 2b\gamma_1'(s)\gamma_2'(s) + c\gamma_1'(s)^2 = 0.}$$

Eine Kurve $\gamma(s)$ heißt nichtcharakteristisch in $s \in I$, falls $\det(s) \neq 0$. Ist γ überall, d.h. für alle $s \in I$, charakteristisch (nicht-charakteristisch), so heißt γ Charakteristik (Nicht-Charakteristik).

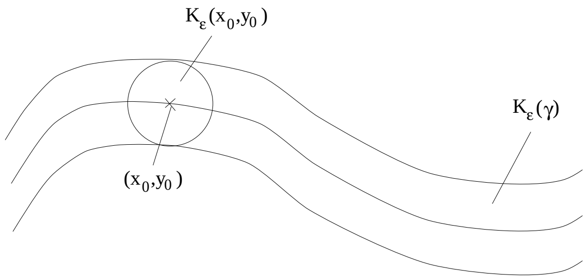
Um den oben angedeuteten Existenzbeweis zu vollenden, muß man die lokale Konvergenz der Taylor–Entwicklung sichern. Das ist der schwierigere Teil des Beweises.

Satz 2.2 (Cauchy-Kowalevskaya) *Für das Cauchy-Problem* (2.2) *gegeben auf der Kurve $\gamma = \gamma(s)$ gilt das folgende: Es sei $\gamma(s)$ nicht-charakteristisch in s_0 (und somit in einer Umgebung von s_0). In einer Umgebung von $\gamma(s_0)$,[2] seien u_0, u_1 bzw. die Funktionen a, b, c, d reell analytisch (d.h. durch multivariate Potenzreihen darstellbar). Dann existiert in einer genügend kleinen Umgebung $K_\varepsilon(\gamma(s_0))$ eine eindeutig bestimmte analytische Lösung des Cauchy-Problems* (2.2).

Beweis:

Siehe zum Beispiel John [14] Chapter 3.3 oder Renardy und Rogers [22] Chapter 2.2

∎



[2] Genauer: In einer Umgebung von $(\gamma(s_0)), u_0(\gamma(s_0)), u_x(\gamma(s_0)), u_y(\gamma(s_0)))$

Bemerkung: Dieser grundlegende Existenzsatz lässt sich zwar noch etwas verallgemeinern. Es bleiben aber wesentliche Schönheitsfehler, nämlich

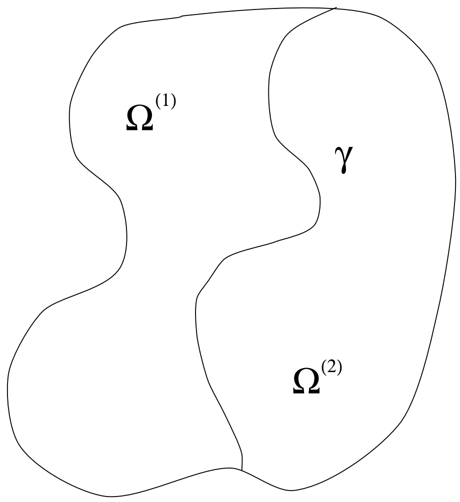
- sehr starke Regularitätsvoraussetzungen an die Daten
- nur lokale Existenz

Um die grundlegende Bedeutung von Charakteristiken weiter zu illustrieren, betrachten wir noch die *Ausbreitung von Unstetigkeiten.*

Satz 2.3 *Die Funktionen a, b, c, d seien stetig. Weiter sei $\gamma = (\gamma_1, \gamma_2)$ eine glatte Kurve und $\Omega = \Omega^{(1)} \cup \gamma \cup \Omega^{(2)}$. Schließlich habe $u \in C^1(\Omega)$ die folgenden Eigenschaften. Die Restriktionen $u^{(i)}$,*

$$u|_{\Omega^{(i)}} =: u^{(i)} \in C^2(\Omega^{(i)} \cup \gamma)\ ,\ i = 1, 2,$$

seien jeweils Lösungen von (2.0) in $\Omega^{(i)} \cup \gamma$ und in jedem Punkt auf γ sei mindestens eine der zweiten Ableitungen u_{xx},u_{xy} oder u_{yy} unstetig. Dann ist γ eine Charakteristik.



Beweis:

Den Sprung einer Funktion $v : \Omega \to \mathbb{R}$ entlang γ definieren wir als

$$[v] := v^{(1)}(\gamma_1(s), \gamma_2(s)) - v^{(2)}(\gamma_1(s), \gamma_2(s)) \quad .$$

Da $u \in C^1(\Omega)$, erhält man aus der Kettenregel

$$\begin{aligned} 0 &= \frac{d}{ds}[u_x] = u_{xx}^{(1)}\gamma_1' + u_{xy}^{(1)}\gamma_2' - u_{xx}^{(2)}\gamma_1' - u_{xy}^{(2)}\gamma_2' = [u_{xx}]\gamma_1' + [u_{xy}]\gamma_2' \\ 0 &= \frac{d}{ds}[u_y] = [u_{xy}]\gamma_1' + [u_{yy}]\gamma_2' \end{aligned}$$

Die Stetigkeit von u, u_x, u_y und a, b, c, d liefert

$$a[u_{xx}] + 2b[u_{xy}] + c[u_{yy}] = 0 \quad .$$

Insgesamt ergibt sich das lineare Gleichungssystem

$$\begin{pmatrix} \gamma_1' & \gamma_2' & 0 \\ 0 & \gamma_1' & \gamma_2' \\ a & b & c \end{pmatrix} \begin{pmatrix} [u_{xx}] \\ [u_{xy}] \\ [u_{yy}] \end{pmatrix} = 0 \quad . \tag{2.4}$$

Da nach Voraussetzung $[u_{xx}]^2 + [u_{xy}]^2 + [u_{yy}]^2 \neq 0$ ist, muß γ eine Charakteristik sein.

□

2.2 Klassifizierung

2.2.1 Quasilineare Differentialgleichungen zweiter Ordnung in zwei Variablen

Die Grundlage der Klassifizierung von quasilinearen Differentialgleichungen zweiter Ordnung (vgl. (2.0)) ist die Existenz oder Nichtexistenz von Charakteristiken.
Zur Berechnung der *charakteristischen Richtung* $(\gamma_1', \gamma_2')^\top$ steht die Gleichung

$$a{\gamma_2'}^2 - 2b\gamma_1'\gamma_2' + c{\gamma_1'}^2 = 0$$

zur Verfügung. Seien die Argumente x, y, u, u_x, u_y fest gewählt. Über die Existenz von reellen Lösungen γ_1', γ_2' und damit von Charakteristiken entscheidet dann die Diskriminante

$$b^2 - ac \begin{cases} < 0 & \text{keine reelle Lösung} \\ = 0 & \text{eine relle Lösung} \\ > 0 & \text{zwei reelle Lösungen} \end{cases} \tag{2.4}$$

Dies motiviert die

Definition 2.4 *Die Differentialgleichung* (2.0) *heißt in* x, y, u, u_x, u_y

$$\boxed{\begin{array}{rcccl} \text{elliptisch} & \Longleftrightarrow & b^2 - ac < 0 & \Longleftrightarrow & \text{keine reelle Charakteristik} \\ \text{parabolisch} & \Longleftrightarrow & b^2 - ac = 0 & \Longleftrightarrow & \text{genau eine reelle Charakteristik} \\ \text{hyperbolisch} & \Longleftrightarrow & b^2 - ac > 0 & \Longleftrightarrow & \text{zwei reelle Charakteristiken.} \end{array}}$$

Ohne das folgende Ergebnis wäre die ganze Definition sinnlos!

Satz 2.5 *Der Typ einer Differentialgleichung ist invariant gegenüber Koordinatentransformationen.*

Beweis:

Übung

□

Bemerkung: Der Typ einer Differentialgleichung kann von x, y und sogar von der (durch zusätzliche Bedingungen festgelegten) Lösung u abhängen.

Bemerkung: Nur der Hauptteil

$$au_{xx} + 2bu_{xy} + cu_{yy}$$

und nicht $d(x, y, u, u_x, u_y)$ entscheidet über den Typ einer Differentialgleichung.

Bemerkung: Im Falle $v \in C_0^\infty(\mathbb{R}^2)$ liefert die Anwendung der Fourier-Transformation

$$\hat{v}(\xi, \eta) = \mathcal{F}v(\xi, \eta) = \frac{1}{2\pi} \int\limits_{\mathbb{R}^2} e^{-i(x\xi + y\eta)} v(x, y) d(x, y)$$

auf die partielle Ableitung $v = u_x$ die Beziehung $\mathcal{F}u_x = i\xi\hat{u}$ (partielle Integration). Im Falle konstanter, lösungsunabhängiger Koeffizienten $a, b, c \in \mathbb{R}$ und $u \in C_0^\infty(\mathbb{R}^2)$ liefert die Anwendung der Fouriertransformation auf (2.0) dementsprechend

$$\boxed{(a\xi^2 + 2b\xi\eta + c\eta^2)\, \hat{u}(\xi, \eta) = -\hat{d}(\xi, \eta).}$$

Das Polynom $a\xi^2 + 2b\xi\eta + c\eta^2$ heißt *Symbol* der Differentialgleichung. Im elliptischen, parabolischen oder hyperbolischen Fall beschreibt das Symbol eine Ellipse, Parabel oder Hyperbel. Daher die Namensgebung.

Beispiel:

a) konstante Koeffizienten

Laplace-Gleichung:	$u_{xx} + u_{yy} = 0$	$(a = c = 1,\ b = d = 0)$
elliptisch	keine Charakteristiken	
Wärmeleitungsgleichung:	$u_y - u_{xx} = 0$	$(a = 1,\ b = c = 0,\ d = u_y)$
parabolisch	Charakteristik $\gamma = (s, 0)$	
Wellengleichung:	$u_{xx} - u_{yy} = 0$	$(a = -c = 1,\ b = d = 0)$
hyperbolisch	Charakteristiken $\gamma_\pm = (s, \pm s + \text{const.})$	

b) lösungsabhängige Koeffizienten

Potentialströmung

$$(c_0^2 - u_x^2)u_{xx} - 2u_x u_y u_{xy} + (c_0^2 - u_y^2)u_{yy} = 0$$

$$\begin{array}{lcll} M & = & (c_0^{-2}(u_x^2 + u_y^2))^{\frac{1}{2}} & \text{Machsche Zahl} \\ M & < & 1 & \text{ellipisch} \\ M & = & 1 & \text{parabolisch} \\ M & > & 1 & \text{hyperbolisch} \end{array}$$

2.2.2 Systeme erster Ordnung in zwei Variablen

Wir kommen nun zur Klassifizierung von Systemen 1. Ordnung. Diese sollte nicht im Widerspruch zu unseren bisherigen Abmachungen stehen. Die Betrachtung von Systemen erster Ordnung wird uns helfen, die Bedeutung der Charakteristiken insbesondere für den hyperbolischen Fall noch besser zu verstehen.
Die oben betrachtete (skalare) Gleichung zweiter Ordnung lässt sich in ein System 1. Ordnung umschreiben. Dazu führen wir die neuen Variablen v, w durch

$$\begin{array}{rcl} u_x & = & v \\ u_y & = & w \end{array} \tag{2.5}$$

ein. Dann erhält die Gleichung zweiter Ordnung

$$\boxed{au_{xx} + 2bu_{xy} + cu_{yy} = d}$$

die Gestalt

$$\boxed{\begin{array}{rcl} av_x + 2bv_y + cw_y & = & d \\ w_x - v_y & = & 0 \quad , \end{array}}$$

wobei die zweite Gleichung schlicht aus

$$v_y = u_{xy} = u_{yx} = w_x \quad .$$

resultiert. Wir nehmen zur Vereinfachung

$$\boxed{a \neq 0}$$

an. Dann läßt sich System (2.5) in Matrixschreibweise schreiben als

$$\boxed{U_x + AU_y + D = 0 \qquad \text{mit} \quad U = \begin{pmatrix} v \\ w \end{pmatrix},}$$

wobei

$$A = \begin{pmatrix} 2b/a & c/a \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}, \quad \text{und} \quad D = \begin{pmatrix} -d/a \\ 0 \end{pmatrix}.$$

Die Matrix A hat die Eigenwerte

$$\boxed{\lambda_\pm = \frac{1}{a}\left(b \pm \sqrt{b^2 - ac}\right).}$$

Wir haben also analog zu (2.4) die folgende Situation

$$b^2 - ac \begin{cases} < 0 & \text{kein reeller Eigenwert} \\ = 0 & \text{ein (doppelter) reeller Eigenwert} \\ > 0 & \text{zwei (verschiedene) reelle Eigenwerte} \end{cases}$$

Darüberhinaus ergibt sich im Fall $b^2 - ac > 0$, dass A reell diagonalisierbar ist! Wir verwenden diese Beobachtung zur Klassifikation von Systemen 1. Ordnung.

Definition 2.6 *Das System*

$$U_x + AU_y + BU + D = 0 \tag{2.6}$$

mit $A, B(x, y, v, w) \in \mathbb{R}^{n,n}$, $D(x, y, v, w) \in \mathbb{R}^n$ und unbekannter Funktion $U(x, y) \in \mathbb{R}^n$ heißt elliptisch, falls alle Eigenwerte von A komplex sind und hyperbolisch, falls A reell diagonalisierbar ist.

Beispiel:

- **Laplace-Gleichung.** Anwendung der Transformation (2.5) auf die (elliptische!) Laplace-Gleichung $u_{xx} + u_{yy} = 0$, ergibt die *Cauchy–Riemann'schen Differentialgleichungen*, nämlich ein System der Form (2.5) mit

  $$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad , \quad a = c = 1,\ b = 0.$$

  Offenbar hat A die imaginären Eigenwerte $\pm i$. Achtung! Querverbindung zur Funktionentheorie: Die komplexwertige Funktion $W(z) = w(z) + iv(z)$ mit komplexem Argument $z = x + iy$ ist genau dann holomorph (d.h. komplex differenzierbar), wenn v, w Lösungen des obigen elliptische Systems sind.

- **Wellengleichung.** Verwendet man die Transformation (2.5) auf die (hyperbolische!) Wellengleichung $u_{xx} - u_{yy} = 0$, so erhält man ein System der Form (2.5) mit

  $$A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad , \quad a = 1,\ b = 0,\ c = -1.$$

  Offenbar hat A die reellen Eigenwerte $\lambda_1 = -1$, $\lambda_2 = 1$ und die orthogonalen Eigenvektoren

  $$e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad e_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad .$$

Dementsprechend können wir das System (2.5) in diesem Fall durch die Transformation

$$T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad , \quad T^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

diagonalisieren. Wir erhalten

$$TU_x + TAT^{-1}TU_y = 0 \tag{2.7}$$

und somit ein *entkoppeltes* System $\tilde{U}_x + TAT^{-1}\tilde{U}_y = 0$ in den neuen Unbekannten

$$\tilde{U} = \begin{pmatrix} \tilde{v} \\ \tilde{w} \end{pmatrix} = TU = \begin{pmatrix} u_x + u_y \\ u_x - u_y \end{pmatrix}.$$

Bemerkung: Analog zum obigen Vorgehen können wir auch die Wärmeleitungsgleichung in ein parabolisches System 1. Ordnung überführen, welches tatsächlich weder hyperbolisch noch elliptisch ist. Mangels Bedeutung spricht man allerdings nicht von parabolischen Systemen erster Ordnung.

Charakteristiken für skalare Differentialgleichungen erster Ordnung. Um die Bedeutung dieser Einsichten näher verstehen zu können, betrachten wir für einen Moment eine skalare Differentialgleichung der Form

$$\boxed{a_1 u_x + a_2 u_y + bu + d = 0.} \tag{2.8}$$

mit x, y-abhängigen Koeffizienten aus $C^1(\mathbb{R}^2)$. Man definiert in diesem Fall:

Definition 2.7 *Ein stetig differenzierbare Kurve $\gamma = (\gamma_1, \gamma_2) : I = [0, T] \to \mathbb{R}^2$, das in I die gewöhnlichen Differentialgleichungen*

$$\boxed{\gamma_1' = a_1(\gamma), \qquad \gamma_2' = a_2(\gamma)} \tag{2.9}$$

erfüllt, heißt Charakteristik von (2.7).

Es gelten folgende Aussagen:

Satz 2.8 *Sei $u \in C^1(\mathbb{R}^2)$ eine Lösung von (2.7) und γ eine Charakteristik. Dann erfüllt $U(t) = u(\gamma(t))$ die gewöhnliche Differentialgleichung*

$$\boxed{U'(t) + b(\gamma(t))\, U(t) + d(\gamma(t)) = 0. \qquad \forall t \in I} \tag{2.10}$$

Beweis.
Übung.

Satz 2.9 *Zu jedem Punkt $(x,y) \in \mathbb{R}^2$ existiere eine Charakteristik γ mit zugehörigem Parameterintervall $I = [0,T]$, so daß $\gamma(t) = (x,y)$ für ein $t \in [0,T]$ gilt. Sei $u \in C^1(\mathbb{R}^2)$ eine Funktion mit der Eigenschaft, daß die Einschränkung $U(t) = u(\gamma(t))$ für jede beliebige Charakteristik γ die gewöhnliche Differentialgleichung (2.9) erfüllt. Dann ist u eine Lösung der partiellen Differentialgleichung (2.7).*

Beweis.
Übung.

Diese beiden Sätze zeigen also, dass die Charakteristiken von (2.7) gerade die Kurven sind, auf denen sich die Werte $u(\gamma(t))$ aus der *gewöhnlichen* Differentialgleichung (2.9) berechnen lassen. Entlang der Charakteristiken wird die partielle Differentialgleichung also zu einer gewöhnlichen Differentialgleichung! Da keine Interaktion mit Werten außerhalb der Chrakteristik auftritt, sagt man auch, daß der Wert der Lösung wird entlang der Charakteristiken *transportiert* wird.

Bemerkung: Die Bedeutung dieser Einsicht für unsere obigen Betrachtungen zeigt sich vor allem anhand der Wellengleichung $u_{xx} - u_{yy} = 0$: Da das zweidimensionale System (2.6) entkoppelt, kann also das zur Wellengleichung gehörige System erster Ordnung in zwei unabhängige skalare partieller Differentialgleichungen der Form (2.7) zerlegt werden. Deren Charakteristiken $\gamma^{(\pm)}$ ergeben sich als Lösungen zu

$$(\gamma_1')^{(\pm)} = 1, \quad (\gamma_2')^{(\pm)} = \pm 1,$$

sind also durch die zwei Geraden $\gamma^{(\pm)}(t) = \gamma^{(\pm)}(0) + (t, \pm t)$ gegeben und tatsächlich identisch mit den Charakteristiken, die wir bei der Untersuchung der Wellengleichung als Gleichung zweiter Ordnung bestimmt hatten. Wegen $b = d = 0$ lautet (2.9) hier einfach $\dot{U} = 0$, d.h. die Anfangsbedingungen werden einfach unverändert entlang der Charakteristiken transportiert. Als Lösung des sog. *Cauchy'schen Anfangswert–Problems*

$$\begin{aligned} u_{xx} - u_{yy} = 0 &\qquad (x,y) \in \mathbb{R} \times \mathbb{R}_+ \\ u(x,0) = u_0(x) &\qquad x \in \mathbb{R} \\ u_y(x,0) = u_1(x) &\qquad x \in \mathbb{R} \end{aligned}$$

erhält man auf diese Weise

$$\begin{aligned} u_x(x,y) &= \frac{1}{2}\big(u_1(x-y) - u_1(x+y) + u_0'(x-y) + u_0'(y+x)\big) \\ u_y(x,y) &= \frac{1}{2}\big(u_1(x-y) + u_1(y+x) + u_0'(x-y) - u_0'(y+x)\big) \end{aligned}$$

wobei wir $u_x(x,0) = u_0'(x)$ benutzt haben. Daraus erhält man die sogenannte *d'Alembert'sche Lösung* u des Cauchy'schen Anfangswertproblems. Wie sieht die aus (Übung)?

2.3 Korrekt oder schlecht gestellte Probleme

2.3.1 Anfangswertprobleme

Als Beispiel betrachten wir das Cauchy-Problem (2.2) für die **Laplace-Gleichung**

$$\boxed{\begin{array}{rcll} \Delta u &=& 0 & \text{in } \Omega \\ u &=& g_0 & \text{auf } \gamma \\ \dfrac{\partial}{\partial n} u &=& g_1 & \text{auf } \gamma \end{array}}$$

auf dem Rechengebiet Ω

$$\Omega = \mathbb{R} \times \mathbb{R}_+, \qquad \mathbb{R}_+ = \{y \in \mathbb{R} \mid y > 0\},$$

mit $\gamma = \partial\Omega = \{(x,y) \in \mathbb{R}^2 : \ y = 0\}$. Da die zweite Variable y meist als Zeit interpretiert werden kann, nennt man die entsprechenden Cauchy-Bedingungen

$$u(x,0) = g_0, \qquad u_y(x,0) = g_1 \qquad \forall x \in \mathbb{R}$$

für $y = 0$ in diesem Fall *Anfangsbedingungen* und das gesamte Problem *Cauchy'sches Anfangswertproblem.*
Sei $j \in \mathbb{N}$ fest gewählt. Im Falle

$$g_0(x) = \cos(j\pi x), \qquad g_1(x) = j\pi \cos(j\pi x) \tag{2.11}$$

kann man eine Lösung des resultierenden Anfangswertproblems für die Laplace-Gleichung mit Hilfe des *Produktansatzes*

$$u_j(x,y) = w(y) \cos(j\pi x)$$

leicht bestimmen. Es folgt nämlich

$$u_j(x,y) = e^{j\pi y} \cos(j\pi x) \quad .$$

Die Lösung

$$u^n(x,y) = \sum_{j=1}^{n} \frac{1}{(j\pi)^4} e^{j\pi y} \cos(j\pi x)$$

zu den Anfangsbedingungen

$$g_0^n(x) = \sum_{j=1}^{n} \frac{1}{(j\pi)^4} \cos(j\pi x), \qquad g_1^n(x) = \sum_{j=1}^{n} \frac{1}{(j\pi)^3} \cos(j\pi x)$$

erhält man durch Superposition. Offenbar konvergieren die Funktionenfolgen g_0^n, g_1^n gleichmäßig gegen die Lipschitz-stetigen Funktionen

$$\boxed{g_0(x) = \textstyle\sum_{j=1}^{\infty} \frac{1}{(j\pi)^4} \cos(j\pi x), \qquad g_1(x) = \sum_{j=1}^{\infty} \frac{1}{(j\pi)^3} \cos(j\pi x).} \tag{2.12}$$

Für beliebiges $y > 0$ gilt jedoch

$$\boxed{u^n(0,y) = \sum_{j=1}^n \frac{1}{(j\pi)^4} e^{j\pi y} \to \infty.}$$

Beliebig kleine Veränderungen der Anfangsbedingungen haben also (unendlich) große Wirkung auf die zugehörigen Lösungen. Dies motiviert die folgende Definition.

Definition 2.10 *Das Cauchy-Problem* (2.2) *heißt korrekt gestellt (in X Y bzgl. der Normen $\|\cdot\|$), falls gilt*

- *Existenz und Eindeutigkeit: $\forall g_0,\ g_1 \in X$ existiert eine eindeutig bestimmte Lösung $u \in Y$*
- *Stetige Abhängigkeit von den Daten: es ex. $c_0,\ c_1 \in \mathbb{R}$, so daß*

$$\|u - \tilde{u}\| \le c_0\|g_0 - \tilde{g}_0\| + c_1\|g_1 - \tilde{g}_1\|.$$

Andernfalls heißt das Problem schlecht gestellt (in X Y bzgl. der Normen $\|\cdot\|$).

Wir haben durch unser Gegenbeispiel gezeigt, daß das Cauchy'sche Anfangswertproblem für die Laplace-Gleichung bezüglich der Maximum-Norm $\|\cdot\|_\infty$ schlecht gestellt ist. Dennoch kommen elliptische Probleme auf unbeschränkten Gebieten in der Praxis vor (z.B. Potential einer Punktladung). Im diesem Falle muß man geeignete Abklingbedingungen stellen, um ein korrekt gestelltes Problem zu erhalten.

Wellengleichung. Analog zum obigen Vorgehen (Produktansatz, Superposition, Grenzübergang) erhält man als Lösung des Cauchy'schen Anfangswertproblems für die (hyperbolische!) Wellengleichung $u_{xx} - u_{yy} = 0$ zu den Anfangsbedingungen (2.11) die beschränkte Grenz-Lösung

$$u^n(x,y) \to u(x,y) = \sum_{j=1}^{\infty} \frac{1}{(j\pi)^4} (\sin(j\pi y) + \cos(j\pi y)) \cos(j\pi x) \quad .$$

Dieses Problem ist offenbar **korrekt gestellt** im Sinne der Maximumsnorm.

Wärmeleitungsgleichung. Im Falle der (parabolischen!) Wärmeleitungsgleichung $u_y = u_{xx}$ ist die Lösung durch die beiden Anfangsbedingungen (2.10) *überbestimmt*, denn es gilt zumindest für glatte Lösungen

$$g_1(x) = u_y(x,0) = \lim_{y\to 0} u_y(x,y) = \lim_{y\to 0} u_{xx}(x,y) = g_0''(x) \quad .$$

Diese Schwierigkeit ist nicht überraschend, denn $\gamma = \{(x,0) \mid x \in \mathbb{R}\}$ ist gerade die Charakteristik der Wärmeleitungsgleichung! Wir können also, obwohl wir es mit einer Differentialgleichung zweiter Ordnung zu tun haben, beim Cauchy'schen Anfangswertproblem für die Wärmeleitungsgleichung nur eine Anfangsbedingung, nämlich

$$u(x,0) = g_0(x) \qquad\qquad \forall x \in \mathbb{R}$$

fordern. Dieser Sachverhalt lässt sich auch physikalisch interpretieren. ("Unendliche Ausbreitungsgeschwindigkeit").

Bemerkung: In unserem Beispiel

$$g_0(x) = \sum_{j=1}^{\infty} \frac{1}{(j\pi)^4} \cos(j\pi x)$$

erhält man mittels Produktansatz, Superposition und Grenzübergang die Lösung

$$u(x,y) = \sum_{j=1}^{\infty} \frac{1}{(j\pi)^4} e^{-(j\pi)^2 y} \cos(j\pi x) \quad .$$

Welche Schwierigkeiten stellen sich ein, wenn man an Stelle von $u_y - u_{xx} = 0$ die rückwärts genommene Wärmeleitungsgleichung $u_y + u_{xx} = 0$ betrachtet? Wie lassen sich diese Schwierigkeiten physikalisch interpretieren?

2.3.2 Randwert– und Anfangsrandwertprobleme

Randwertprobleme. Für partielle Differentialgleichungen kann man bei beschränktem Rechengebiet Ω meistens jeweils nur eine Bedingung am Rand stellen. Wir haben bereits Dirichlet–, Neumann– und Cauchy–Randbedingungen in Kapitel 1 kennengelernt. Man spricht dann von einem *Randwertproblem.*
Mit elliptischen Randwertproblemen werden wir uns im nächsten Abschnitt ausführlich beschäftigen. Es lässt sich durch Gegenbeispiele zeigen, daß Randwertprobleme für hyperbolische und parabolische Differentialgleichungen im allgemeinen schlecht gestellt sind (Übung).

Anfangsrandwertprobleme. Ist das Rechengebiet in eine Koordinatenrichtung unbeschränkt (meist in Zeitrichtung) und in alle anderen Richtungen beschränkt, so sind neben Anfangsbedingungen noch Randbedingungen nötig. Zum Beispiel könnten wir die Wärmeleitungsgleichung auf

$$\Omega = (0,1) \times \mathbb{R}_+$$

unter der Anfangsbedingung

$$u(x,0) = u_0(x), \qquad \forall x \in \mathbb{R}$$

und den Dirichlet-Randbedingungen

$$u(0,y) = g_0(y), \quad u(1,y) = g_1(y) \qquad \forall y \in \mathbb{R}_+$$

betrachten (physikalische Interpretation?) Dies ist ein typisches *Anfangsrandwertproblem*. Auch Anfangsrandwertprobleme für elliptische Differentialgleichungen sind im allgemeinen schlecht gestellt (Gegenbeispiel?).

3 Klassische Lösungen elliptischer Probleme

3.1 Lösungsdarstellung und Greensche Funktion

Als Prototyp elliptischer Differentialgleichungen betrachten wir die *Poisson-Gleichung*

$$-\Delta u = f \qquad \text{in } \Omega \quad . \tag{3.1}$$

Dabei ist $f \in C(\Omega)$ vorausgesetzt und

$$\Delta u = \sum_{j=1}^{n} u_{x_j x_j}$$

der Laplace-Operator. Im Fall $f \equiv 0$ heißt (3.0) *Laplace-Gleichung*. Jede Lösung u der Laplace-Gleichung heißt *harmonisch*. Wir betrachten (3.0) auf einem beschränkten Rechengebiet $\Omega \subset \mathbb{R}^n$ (offen und zusammenhängend) mit glattem Rand $\partial\Omega$ (Greenscher Bereich).
Auf $\partial\Omega$ legen wir Dirichlet-Randbedingungen fest

$$u(x) = g(x) \qquad \text{für} \quad x \in \partial\Omega \tag{3.2}$$

und setzen dabei $g \in C(\partial\Omega)$ voraus.

Definition 3.1 *Genügt $u \in C^2(\Omega) \cap C(\overline{\Omega})$ in Ω der Poisson-Gleichung* (3.0) *und auf dem Rand $\partial\Omega$ den Dirichlet-Randbedingungen* (3.1)*, so heißt u klassische Lösung des Randwertproblems* (3.0), (3.1).

Der Lösungsraum $C^2(\Omega) \cap C(\overline{\Omega})$ gewährleistet die Existenz der nötigen Ableitungen von u im Inneren von Ω und die stetige Annahme der Randbedingungen.
Folgendes Beispiel zeigt, daß es nicht sinnvoll ist den Lösungsraum weiter einzuschränken.
Beispiel:

Sei $\Omega = \{(r,\varphi)\,|\, r \in (0,1)\ ,\ \varphi \in (0, \frac{3}{2}\pi)\}$ und

$$\begin{array}{rcll} f(r,\varphi) & \equiv & 0 & (x,y) \in \Omega \\ g(r,\varphi) & = & \sin(\frac{2}{3}\varphi) & (x,y) \in \partial\Omega \quad . \end{array}$$

Beachte, daß $g \in C(\partial\Omega)$! Wir transformieren den Laplace-Operator auf Polarkoordinaten

$$\begin{array}{rcl} x & = & r\cos\varphi \\ y & = & r\sin\varphi \quad . \end{array}$$

und erhalten mit

$$u(r,\varphi) = r^{\frac{2}{3}} \sin(\tfrac{2}{3}\varphi)$$

die Lösung des entsprechenden Randwertproblems. Wir zeigen nun, daß $u \notin C^1(\overline{\Omega})$ gilt.
Mit u_x und u_y müßte auch u_r beschränkt sein, denn

$$u_r = u_x x_r + u_y y_r = u_x \cos\varphi + u_y \sin\varphi \quad ,$$

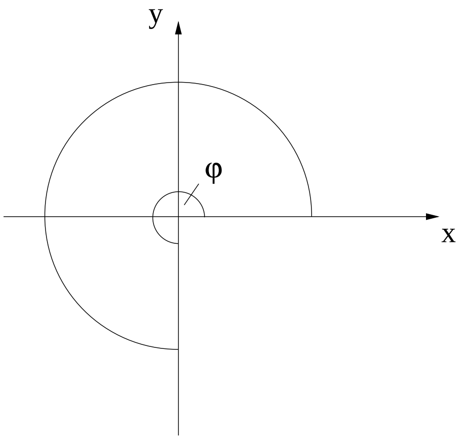


Abbildung 4: Einspringende Ecke

aber

$$u_r = \tfrac{2}{3} r^{-\frac{1}{3}} \sin(\tfrac{2}{3}\varphi) \to \infty \quad \text{für} \quad r \to 0 \quad .$$

Der Grund für die Schwierigkeiten ist die *einspringende Ecke* im Rechengebiet Ω.

Wir wollen nun untersuchen, ob eine klassische Lösung u durch g und f eindeutig bestimmt ist und stetig von diesen Daten abhängt. Dazu werden wir zunächst ein *Integraldarstellung* von (genügend glatten) Lösungen herleiten. Grundlegend sind dabei (wieder einmal) geeignete Produktregeln.

Lemma 3.2 *Es seien $u, v \in C^2(\overline{\Omega})$. Dann gilt die 1. Greensche Formel*

$$\boxed{\int_\Omega u \Delta v \, dx = -\int_\Omega \nabla u \cdot \nabla v \, dx + \int_{\partial\Omega} u \, \tfrac{\partial v}{\partial n} \, d\sigma} \tag{3.3}$$

und die 2. Greensche Formel

$$\boxed{\int_\Omega (u\Delta v - v\Delta u) \, dx = \int_{\partial\Omega} (u \, \tfrac{\partial v}{\partial n} - v \, \tfrac{\partial u}{\partial n}) \, d\sigma.} \tag{3.4}$$

Beweis:

(3.2) folgt durch Wahl von $V = \nabla v$ direkt aus dem Divergenzsatz. (3.3) folgt durch Vertauschen von u und v in (3.2) und Subtraktion.

□

Folgerung

Das Dirichlet-Problem für die Poissongleichung hat *höchstens* eine Lösung $u \in C^2(\overline{\Omega})$ (Übung). Insbesondere ist u durch Vorgabe der Randwerte von u *und* $\frac{\partial u}{\partial n}$ überbestimmt.

Folgerung

Im Falle von Neumann-Randbedingungen

$$\frac{\partial}{\partial n} u = h \qquad \text{auf } \partial\Omega$$

ist die Kompatibilitätsbedingung

$$\int\limits_{\Omega} f\, dx = -\int\limits_{\partial\Omega} h\, d\sigma$$

notwendig für die Existenz einer Lösung.[3]

Man bestätigt durch Nachrechnen, daß der Laplace-Operator invariant gegenüber Drehungen der Koordinatenachsen ist (physikalische Interpretation?). Dies motiviert die Betrachtung rotationssymmetrischer Lösungen, d.h. von Lösungen der Gestalt

$$\boxed{u(x) = s(r) \qquad r = |x-a| \quad , \quad a \in \mathbb{R}^n.}$$

Lemma 3.3
Rotationssymmetrische, in $\mathbb{R}^n \setminus \{a\}$ harmonische Funktionen haben die Gestalt

$$\boxed{s(r) = \begin{cases} c \log r + C & n = 2 \\ c\, r^{2-n} + C & n > 2 \end{cases}} \tag{3.5}$$

mit frei wählbaren Konstanten $c, C \in \mathbb{R}$.

Beweis:

Durch Tranformation des Laplace-Operators in Polar-Koordinaten. Zur Übung.

□

Wir wählen nun spezielle Werte für die Konstanten c, C.

Definition 3.4 *Sei $a \in \mathbb{R}^n$. Die in $\mathbb{R}^n \setminus \{a\}$ harmonische Funktion*

$$\boxed{s(x,a) = \begin{cases} -\frac{1}{2\pi} \log|x-a| & n = 2 \\ -\frac{1}{(2-n)\omega_n} |x-a|^{2-n} & n > 2 \end{cases}}$$

heißt Singularitätenfunktion. Dabei ist $\omega_n = \int_{|x|=1} d\sigma$ die Oberfläche der Einheitskugel im $\mathbb{R}^n$.

[3]Direkt durch Einsetzen von $\Delta v = -f$ und $\frac{\partial v}{\partial n} = g$ mit $u = 1$ in (3.2).

Mit Hilfe der Singularitätenfunktion definiert man die Fundamentallösungen:

Definition 3.5 *Sei* $\Phi(\cdot, a) \in C^2(\overline{\Omega})$ *harmonisch für jedes* $a \in \Omega$. *Dann heißt*

$$\boxed{\gamma(x,a) := s(x,a) + \Phi(x,a)}$$

Grundlösung oder Fundamentallösung der Laplace-Gleichung in Ω.

Nun können wir unseren Darstellungssatz formulieren.

Satz 3.6 (Darstellungssatz) *Es sei* $u \in C^2(\overline{\Omega})$ *und* γ *eine Fundamentallösung. Dann gilt für jeden Punkt* $a \in \Omega$

$$\boxed{u(a) = \int_{\partial\Omega} \left(\gamma(x,a)\,\tfrac{\partial}{\partial n} u(x) - u(x)\,\tfrac{\partial}{\partial n}\gamma(x,a)\right)\, d\sigma - \int_\Omega \gamma(x,a)\Delta u(x)\, dx.} \tag{3.6}$$

Beweis:

Den Beweis findet man z.B. bei Jost [15, p. 11]. Dabei wird folgende Eigenschaft der Singularitätenfunktion ausgenutzt: Sei $K_\varrho(a) := \{x \in \mathbb{R}^n \mid |x-a| < \varrho\} \subset \Omega$ und $v \in C(\Omega)$ beliebig. Dann gilt:

$$\lim_{\varrho\to 0} \int\limits_{K_\varrho(a)} v(x)\, s(x,a)\, dx = 0 \quad \text{ und } \quad \lim_{\varrho\to 0} \int\limits_{\partial K_\varrho(a)} v(x)\, s(x,a)\, d\sigma = 0,$$

sowie

$$-\lim_{\varrho\to 0} \int\limits_{\partial K_\varrho(a)} v(x)\, \frac{\partial}{\partial n_K} s(x,a)\, d\sigma = v(a),$$

wobei n_K die äußere Normale auf $K_\varrho(a)$ bezeichnet. Diese Gleichungen implizieren dann die behauptete Aussage.

□

Beachte, daß die Formel (3.5) den Wert der Lösung $u(a)$ durch die rechte Seite $f = -\Delta u$ und die Cauchy-Daten u und $\frac{\partial u}{\partial n}$ auf $\partial\Omega$ darstellt. Wir haben schon gesehen, daß wir in jedem $x \in \partial\Omega$ entweder u oder $\frac{\partial u}{\partial n}$ vorgeben können. Um eine Darstellung von $u(a)$ durch bekannte Daten zu erhalten, treffen wir nun eine geeignete Wahl der bislang unbestimmten Funktion Φ. Sind wie im vorliegenden Fall Dirichletsche Randbedingungen gegeben, so möchten wir Φ **so bestimmen**, daß für alle $a \in \Omega$ die Bedingung

$$\boxed{\gamma(x,a) = 0 \qquad \forall x \in \partial\Omega} \tag{3.7}$$

oder äquivalent

$$\Phi(x,a) = -s(x,a) \qquad \forall x \in \partial\Omega$$

erfüllt ist.

Definition 3.7 *Eine Grundlösung $G(x,a) = \gamma(x,a)$ mit der Eigenschaft $G(x,a) = 0$ für alle $x \in \partial\Omega$ heißt Greensche Funktion 1. Art.*

Die Darstellungsformel (3.5) einer Lösung u des Dirichlet-Problems geht dann über in

$$\boxed{u(a) = \int_\Omega G(x,a)\, f(x)\, dx - \int_{\partial\Omega} \frac{\partial}{\partial n} G(x,a)\, g(x)\, d\sigma.} \tag{3.8}$$

Beispiel:

Sei $g = 0$ und $f_n \in C(\Omega)$, $n \in \mathbb{N}$ eine Folge von Funktionen mit den Eigenschaften

$$\int_\Omega f_n(x)\, dx = 1, \qquad \operatorname{supp} f_n \subset K_{\frac{1}{n}}(x_0) \quad .$$

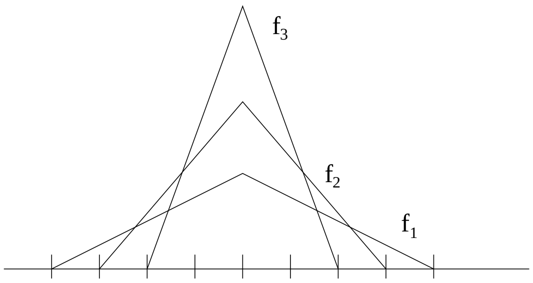


Abbildung 5: Funktionenfolge f_n

Dabei bedeutet

$$\operatorname{supp} f_n = \overline{\{x \in \Omega \mid f_n(x) \neq 0\}}$$

den Träger (engl. support) von f_n und $x_0 \in \Omega$ ist beliebig aber fest gewählt. Wir nehmen an daß die zugehörigen Lösungen u_n existieren und in $C^2(\overline{\Omega})$ liegen. Dann gilt für jedes $a \in \Omega$, $a \neq x_0$

$$u_n(a) \to u(a) = G(x_0, a) \quad \text{für} \quad n \to \infty \quad .$$

Damit kann man sich $u(x) = G(x_0, x)$ als Lösung von

$$\boxed{-\Delta G(x_0, x) = \delta_{x_0}(x)}$$

vorstellen, wobei die rechte Seite

$$\delta_{x_0} = \lim_{n\to\infty} f_n \tag{3.9}$$

die Eigenschaften

$$\int_\Omega \delta(x_0)\, dx, \qquad \delta_{x_0}(x) = 0 \quad \text{für} \quad x \neq x_0$$

haben muß. So eine Funktion gibt es offenbar nicht (obwohl der Physiker Dirac sie erfolgreich verwendete und sie heute in der Physik ein unentbehrliches Hilfsmittel ist). Es handelt sich um eine *verallgemeinerte*

Funktion, nämlich um die Diracsche δ-Distribution. Auf Distributionen werden wir später noch kommen. Dann wird auch klar, wie der Grenzübergang (3.8) zu interpretieren ist.
Vorläufig können wir uns δ_{x_0} als Punktquelle vorstellen. Das Abklingen der zugehörigen Lösung der Greenschen Funktion $G(x, x_0)$ beschreibt insbesondere den Einfluß von Störungen der rechten Seite f in x_0. Das Abklingverhalten von $G(x, x_0)$ wird durch die Singularitätenfunktion dominiert und erfolgt daher wie

$$\begin{array}{ll} \mathcal{O}\left(-\log|x-x_0|\right) & \text{für } n=2 \\ \mathcal{O}\left(|x-x_0|^{2-n}\right) & \text{für } n>2 \end{array}$$

Es gilt $G(x_0, x) > 0$ (Übung). Damit beeinflußen lokale Störungen der rechten Seite f (an einer Stelle $x_0 \in \Omega$) die zugehörige Lösung u zwar global (in allen $x \in \Omega$), der Einfluß klingt aber mit wachsendem Abstand zu x_0 aber sehr schnell ab. Lokale Störungen haben also (fast) lokale Wirkungen.

Die Existenz einer Greenschen Funktion ist im allgemeinen nicht gesichert. Die Schwierigkeiten illustriert das folgende Kriterium.

Satz 3.8 *Für alle $a \in \Omega$ sei das Dirichlet-Problem*

$$\begin{aligned} \Delta\Phi(\cdot, a) &= 0 && \text{in } \Omega \\ \Phi(\cdot, a) &= -s(\cdot, a) && \text{auf } \partial\Omega \end{aligned}$$

lösbar, und es gelte $\Phi(\cdot, a) \in C^2(\overline{\Omega})$. Dann existiert die Greensche Funktion erster Art zum entsprechenden Dirichlet-Problem.

Wir haben schon gesehen, daß insbesondere bei Gebieten mit einspringenden Ecken die Regularitätsbedingung $\Phi(\cdot, a) \in C^2(\overline{\Omega})$ Probleme bereiten kann. Für spezielle Gebiete kann man die Greensche Funktion aber explizit angeben. Das einfachste ist erwartungsgemäß die Kugel, siehe Jost [15, p. 14].

3.2 Existenz

Bislang haben wir die Existenz einer Lösung $u \in C^2(\overline{\Omega})$ vorausgesetzt und schließlich deren Darstellung durch die Daten mittels einer Greenschen Funktion erhalten. Nun kann man hoffen, daß *umgekehrt* durch die Formel (3.7) eine Lösung des Dirichlet-Problems für die Poisson-Gleichung *definiert* wird. Im allgemeinen ist das jedoch *nicht* der Fall. Insbesondere gibt es rechte Seiten $f \in C(\overline{\Omega})$, so daß (3.7) keine Lösung des Dirichlet-Problems ergibt. Man braucht im allgemeinen mehr Regularität.

Satz 3.9 *Es sei $g = 0$ und f Hölder-stetig zum Exponenten $\lambda \in (0, 1)$, d.h. es gelte*

$$|f(x) - f(y)| \leq c\,|x - y|^\lambda \qquad \forall x, y \in \Omega$$

mit einer von x, y unabhängigen Konstanten $c \in \mathbb{R}$. Weiter existiere die zugehörige

Greensche Funktion erster Art $G(\xi, x)$. *Dann ist*

$$\boxed{u(x) = \int_\Omega G(\xi, x) f(\xi)\, d\xi}$$

eine Lösung des Dirichletproblems.

Beweis:

Siehe z.B. Hackbusch [12] S. 38.

□

Bemerkung: Man beachte die Regularitätsbedingungen an f und vor allem an Ω (Existenz einer Greenschen Funktion $G(\xi, x)$!).

Um die Existenz einer Lösung des Dirichlet-Problems (3.0), (3.1) zu gewährleisten, brauchen wir nun nur noch ein Existenzresultat für den homogenen Fall $f = 0$. Wir beschränken uns dabei auf eine Kugel.

Satz 3.10 (Poissonsche Integralformel) *Sei* $\Omega = \{x \in \mathbb{R}^n \mid |x| < \varrho\}$, $f = 0$ *und* $g \in C(\partial\Omega)$. *Dann ist*

$$u(x) = \frac{\varrho^2 - |x|^2}{\varrho \omega_n} \int\limits_{\partial\Omega} \frac{g(\xi)}{|\xi - x|^n}\, d\sigma \quad \text{für} \quad x \in \Omega \tag{3.10}$$

eine klassische Lösung des Dirichlet-Problems (3.0), (3.1). *Weiter gilt* $u \in C^\infty(\Omega) \cap C(\overline{\Omega})$.

Beweis:

Siehe John [14] oder Hackbusch [12]. Der Beweis verwendet die spezifische Form der Greenschen Funktion für die Kugel und Einsetzen in (3.7).

□

Weitergehende Existenzresultate finden sich z.B. bei John [14]. Dabei spielen insbesondere Regularitätsbedingungen an den Rand $\partial\Omega$ eine Rolle.

3.3 Eindeutigkeit und stetige Abhängigkeit von den Daten

Eine wichtige Folgerung aus der Darstellungsformel (3.7) ist die *Mittelwerteigenschaft*.

Satz 3.11 *Sei* u *harmonisch in* Ω. *Ist* $x \in \Omega$ *und* $\varrho > 0$ *so klein, daß* $\overline{K_\varrho(x)} \subset \Omega$, *so gilt*

$$u(x) = \frac{1}{\omega_n \varrho^{n-1}} \int\limits_{\partial K_\varrho(x)} u(\xi)\, d\sigma \quad . \tag{3.11}$$

Beweis:

Für $x = 0$ hat die Darstellungsformel (3.7) die Gestalt (3.9). Einsetzen von $u = g$ auf $\partial\Omega$ und $|\xi| = \varrho$ liefert (3.10). Den Fall $x \neq 0$ führt man durch Translation $\tilde{\Omega} := \Omega - x$ auf den Fall $x = 0$ zurück.

□

Satz 3.12 (Starkes Maximumprinzip) *Ist u harmonisch in $\Omega \subset \mathbb{R}^n$ und nimmt u in $x \in \Omega$ ein Minimum oder Maximum an, so ist u konstant auf Ω.*

Beweis:

Sei $M := \sup\{u(x)\,|\,x \in \Omega\} < \infty$. Wir zerlegen Ω in

$$\Omega_1 = \{x \in \Omega\,|\,u(x) = M\} \cup \Omega_2 = \{x \in \Omega\,|\,u(x) < M\} \quad .$$

Da u stetig ist, ist Ω_2 offen. Wir zeigen, daß auch Ω_1 offen ist. Sei $x \in \Omega_1$, also $u(x) = M$. Aus der Mittelwerteigenschaft folgt für genügend kleine $\varrho > 0$

$$\begin{aligned} 0 &= \frac{1}{\omega_n \varrho^{n-1}} \int\limits_{\partial K_\varrho(x)} u(\xi)\, d\sigma - u(x) \\ &= \frac{1}{\omega_n \varrho^{n-1}} \int\limits_{\partial K_\varrho(x)} (u(\xi) - M)\; d\sigma \end{aligned}$$

Da $u(\xi) - M \leq 0$ und stetig ist, muß $u(\xi) = M \quad \forall \xi \in \partial K_\varrho(x)$ sein. Durch Widerspruchsbeweis folgt, daß $K_\varrho(x) \subset \Omega_1$. Da Ω zusammenhängend ist, muß $\Omega_1 = \emptyset$ oder $\Omega_2 = \emptyset$ gelten. Das liefert die Behauptung.

Die Minimum-Eigenschaft folgt durch Anwendung der Maximum-Eigenschaft auf die harmonische Funktion $-u$.

□

Bemerkung: Aus dem starken Maximumprinzip folgt

$$G(\xi, x) > 0 \qquad \forall \xi, x \in \Omega, \xi \neq x \quad .$$

Folgerung (schwaches Maximumprinzip)

Es gilt für jede harmonische Funktion $u \in C(\overline{\Omega}) \cap C^2(\Omega)$

$$\boxed{\min_{\xi\in\partial\Omega} u(\xi) \leq u(x) \leq \max_{\xi\in\partial\Omega} u(\xi) \qquad \forall x \in \Omega.}$$

Satz 3.13 (Eindeutigkeit) *Das Dirichlet-Problem* (3.0), (3.1) *besitzt höchstens eine Lösung $u \in C^2(\Omega) \cap C(\overline{\Omega})$.*

Beweis:

Seien $u_1, u_2 \in C(\overline{\Omega}) \cap C^2(\Omega)$ Lösungen. Dann löst $u = u_1 - u_2$ das homogene Problem ($f = 0, g = 0$). Wäre $u(x) \neq 0$ für ein $x \in \Omega$, so müßte in Ω ein (positives) Maximium oder ein (negatives) Minimum liegen. Das kann wegen Satz (3.12) nicht sein.

□

Eine letzte wichtige Folgerung aus dem schwachen Maximumprinzip ist

Satz 3.14 (Stetige Abhängigkeit von den Randdaten) *Seien u_1, u_2 klassische Lösungen der Randwertaufgaben*

$$-\Delta u_i = f \text{ in } \Omega \, , \quad u_i = g_i \text{ auf } \partial\Omega \, , \, i = 1, 2 \quad .$$

Dann gilt

$$\max_{x \in \Omega} |u_1(x) - u_2(x)| \leq \max_{y \in \partial\Omega} |g_1(y) - g_2(y)| \quad .$$

Beweis:

Ist $g_1 = g_2$, so folgt $u_1 = u_2$ aus dem Eindeutigkeitssatz (3.13). Sei also $g_1 \neq g_2$.
Offenbar ist $u = u_1 - u_2$ harmonisch und genügt den Randbedingungen $g = g_1 - g_2$. Wäre

$$u(x) > \max_{y \in \partial\Omega} |g(y)| \geq \max_{y \in \partial\Omega} g(y) \quad ,$$

so hätte u ein Maximum in Ω ohne konstant zu sein. Ähnlich schließt man den Fall

$$-u(x) > \max_{y \in \partial\Omega} |g(y)| \geq - \min_{y \in \partial\Omega} g(y)$$

aus.

□

4 Differenzenverfahren

Wir betrachten wie im vorigen Kapitel das Dirichletproblem

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega \\ u &= g && \text{auf } \partial\Omega \end{aligned} \tag{4.1}$$

mit $f \in C(\Omega)$, $g \in C(\partial\Omega)$ und einem beschränkten, glatt berandeten Gebiet Ω (z.B. Parametrisierung von $\partial\Omega$ stetig differenzierbar). Die Existenz einer klassischen Lösung $u \in C^2(\Omega) \cap C(\overline{\Omega})$ wird vorausgesetzt.

4.1 Gitter und Differenzenquotienten

Wir wollen ein Differenzenverfahren zur näherungsweisen Lösung von (4.1) entwickeln. Dafür überziehen wir zunächst das Rechengebiet Ω mit einem *äquidistanten Gitter* zur *Schrittweite* $h > 0$,

$$\Omega_h = \{x \in \Omega \,|\, x = (ih, jh)\ ,\ i, j \in \mathbb{Z}\} \quad .$$

Als *Randdiskretisierung* verwenden wir

$$\partial\Omega_h = \{x \in \partial\Omega \,|\, x = (x_1, x_2)\ ,\ x_1 = ih \text{ oder } x_2 = ih,\ \ i \in \mathbb{Z}\} \quad .$$

Auch andere Diskretisierungen des Randes sind denkbar. Die obige Variante führt später auf das sogenannte *Shortley–Weller–Verfahren*.

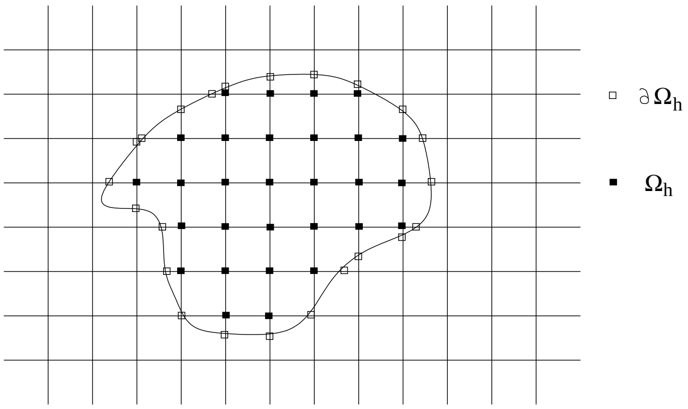


Abbildung 6: Diskretisierung von Ω

Man beachte die Störung des äquidistanten Gitters am Rand. Eine wichtige Teilmenge des äquidistanten Gitters Ω_h ist die Menge der *inneren Punkte*, bezeichnet mit Ω_h°. Diese ist wie folgt definiert:

$$\Omega_h^\circ = \{x \in \Omega_h \mid Nb(x) \subset \Omega_h\}$$

Dabei haben wir die *Menge der Nachbarn* $Nb(x)$, verwendet:

$$Nb(x) = \{y \in \Omega_h \cup \partial\Omega_h \ : \ |x_1 - y_1| + |x_2 - y_2| \leq h\}.$$

Die Nachbarpunkte werden entsprechend Abbildung 7 bezeichnet. In Analogie zum kontinuierlichen Fall setzen wir noch

$$\overline{\Omega}_h = \Omega_h \cup \partial\Omega_h \quad .$$

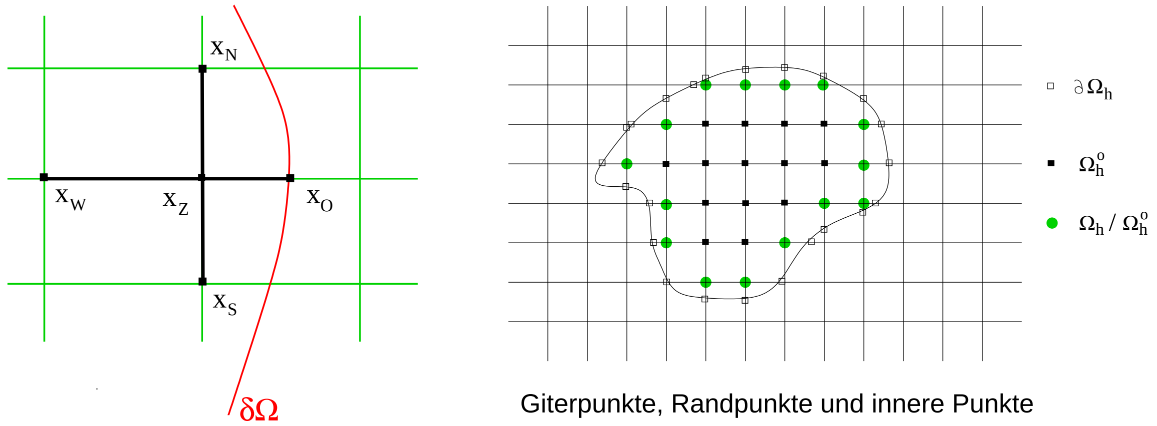


Abbildung 7: Links: Gitterpunkte zur Approximation des Laplace-Operators. Gezeigt ist die Bezeichnung der Nachbarn $Nb(x_Z) = \{x_W, x_N, x_O, x_S\}$ des zentralen Punktes $x_Z \in \Omega_h$. In dem gezeigten Fall ist offensichtlich $x_Z \notin \Omega_h^\circ$. Rechts: Illustration innere Punkte.

Definition 4.1 $\overline{\Omega}_h$ *heißt* *diskret zusammenhängend, wenn man je zwei Gitterpunkte* $x, y \in \overline{\Omega}_h$ *durch eine Folge benachbarter Punkte verbinden kann.*

Lemma 4.2 *Sei* Ω *zusammenhängend. Dann gibt es ein* $h_0 > 0$*, so daß für alle* $h < h_0$ $\overline{\Omega}_h$ *diskret zusammenhängend ist.*

Beweis:

Übung.

□

Von nun an sei $\overline{\Omega}_h$ zusammenhängend, also h genügend klein. Auf $\overline{\Omega}_h$ suchen wir nun eine *Gitterfunktion*

$$U : \overline{\Omega}_h \to \mathbb{R} \quad ,$$

welche die kontinuierliche Lösung $u \in C^2(\Omega) \cap C(\overline{\Omega})$ von (4.1) möglichst gut approximiert. Um eine Berechnungsvorschrift für U zu erhalten, haben wir einerseits die Differentialgleichung und andererseits die Randbedigung zu approximieren.
Zur Approximation der im Laplace-Operator auftretenden Ableitungen wollen wir Differenzenquotienten verwenden. Naheliegende Differenzenapproximationen erster Ordnung sind (mit den Bezeichnungsweisen aus Abb. 7):

$$\boxed{\begin{array}{rcll} D_1^+U(x_Z) & = & \dfrac{U(x_O)-U(x_Z)}{|x_O-x_Z|} & \text{(vorwärts)} \\[2ex] D_1^-U(x_Z) & = & \dfrac{U(x_Z)-U(x_W)}{|x_Z-x_W|} & \text{(rückwärts)} \\[2ex] D_1U(x_Z) & = & \frac{1}{2}\left(D_1^+U(x_Z)+D_1^-U(x_Z)\right) & \text{(zentral)} \quad . \end{array}}$$

Lemma 4.3 *Es sei $u \in C^2(\overline{\Omega})$ und $x \in \Omega_h$. Dann gilt für $D = D_1^+, D_1^-, D_1$*

$$u_{x_1}(x) - Du(x) = \mathcal{O}\left(h\right) \quad . \tag{4.2}$$

Gilt sogar $u \in C^3(\overline{\Omega})$ und ist $x \in \Omega_h^\circ$, so folgt für D_1

$$\boxed{u_{x_1}(x) - D_1u(x) = \mathcal{O}\left(h^2\right).} \tag{4.3}$$

Beweis:

Wir zeigen nur (4.2).
Taylorentwicklung liefert für $x = x_{ij} = (ih, jh) \in \Omega_h^\circ$

$$u(x_{i\pm1,j}) = u(x_{ij}) \pm hu_{x_1}(x_{ij}) + \frac{h^2}{2}u_{x_1x_1}(x_{ij}) \pm \frac{h^3}{6}u_{x_1x_1x_1}(x_{ij} \pm \omega_\pm h)$$

mit $\omega_\pm \in (0,1)$. Einsetzen führt auf

$$\begin{aligned} D_1u(x_{ij}) &= \frac{1}{2h}\left(u(x_{i+1,j}) - u(x_{i-1,j})\right) \\ &= u_{x_1}(x_{ij}) + \mathcal{O}\left(h^2\right) \quad . \end{aligned}$$

□

Zentrale Differenzenquotienten haben also (unter entsprechenden Glattheitsvoraussetzungen!) eine höhere Approximationsgüte als einseitige.
Zur Approximation der 2. Ableitungen verwenden wir daher die *zentralen dividierten Differenzen 2. Ordnung*

$$\boxed{\begin{aligned} D_{11}U(x_Z) &= \frac{2}{|x_O - x_W|}\left(D_1^+U(x_Z) - D_1^-U(x_Z)\right) \\ D_{22}U(x_Z) &= \frac{2}{|x_N - x_S|}\left(D_2^+U(x_Z) - D_2^-U(x_Z)\right) \quad . \end{aligned}}$$

Als Diskretisierung Δ_h des Laplace-Operators Δ wählen wir schließlich

$$\boxed{\Delta_h = D_{11} + D_{22}.}$$

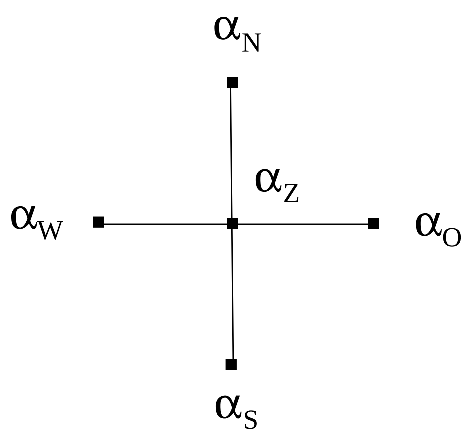
Ausschreiben liefert

$$\boxed{\Delta_h U(x_Z) = \alpha_Z\, U(x_Z) + \alpha_W\, U(x_W) + \alpha_O\, U(x_O) + \alpha_N\, U(x_N) + \alpha_S\, U(x_S).}$$

mit entsprechenden Gewichten α_Z, α_W,α_O, α_S und α_N. Diese lassen sich auch als Werte einer Gitterfunktion

$$\alpha(x, x_D) = \alpha_D, \qquad D = Z, W, O, N, S$$

in $x = x_Z \in \Omega_h$ interpretieren. Anschaulicher ist folgender *Differenzenstern*.



Im Falle $x_Z \in \Omega_h^\circ$ ergibt sich der *Standard-5-Punkte-Stern*

$$\boxed{\alpha_Z = -\tfrac{4}{h^2}\,, \quad \alpha_W = \alpha_O = \alpha_N = \alpha_S = \tfrac{1}{h^2}.}$$

Ansonsten hängen die Gewichte im allgemeinen von $x_Z \in \Omega_h \setminus \Omega_h^\circ$ ab. Für alle $x_Z \in \Omega_h$ gelten jedoch die Beziehungen

$$\alpha_Z < 0\,, \quad \alpha_W, \alpha_O, \alpha_N, \alpha_S > 0\,, \quad \alpha_Z + \alpha_W + \alpha_O + \alpha_N + \alpha_S = 0 \quad .$$

Differenzenverfahren mit diesen Eigenschaften heißen *von positivem Typ*.
Wir wollen kurz die Genauigkeit unserer Differenzenapproximation untersuchen.

Lemma 4.4 *Es sei* $u \in C^3(\overline{\Omega})$ *und* $x \in \Omega_h$. *Dann gilt*

$$\Delta u(x) - \Delta_h u(x) = \mathcal{O}\left(h\right).$$

Ist sogar $u \in C^4(\overline{\Omega})$ *und* $x \in \Omega_h^\circ$, *so folgt*

$$\boxed{\Delta u(x) - \Delta_h u(x) = \mathcal{O}\left(h^2\right).}$$

Beweis:

Taylorentwicklung.

□

4.2 Das Shortley–Weller–Verfahren

Als Differenzenapproximation von (4.1) erhalten wir schließlich das *Shortley–Weller–Verfahren*

$$\boxed{\begin{array}{rcll} -\Delta_h U(x) &=& f_h(x) & x \in \Omega_h \qquad (4.3) \\ U(x) &=& g_h(x) & x \in \partial\Omega_h. \qquad (4.3) \end{array}}$$

Dabei sind f_h und g_h geeignete Approximationen der Daten f und g. Wir wählen einfach

$$f_h(x) = f(x) \text{ für } x \in \Omega_h \quad , \quad g_h(x) = g(x) \text{ für } x \in \partial\Omega_h \quad . \tag{4.4}$$

Mit Hilfe der Randbedingung in (4.3) können wir die Unbekannten $U(x)$, $x \in \partial\Omega_h$ eliminieren.
Zur Bestimmung der übrigen Werte haben wir ein lineares Gleichungssystem mit $N = \#\Omega_h$ Gleichungen und Unbekannten zu lösen. Nach Einführung einer geeigneten Numerierung der Gitterpunkte

$$\Omega_h = \{x_i \,|\, i = 1, \ldots, N\}$$

lautet es

$$\boxed{AU = F \qquad (4.5)}$$

mit Koeffizientenmatrix A,

$$A = (a_{ij})_{i,j=1}^N \qquad a_{ij} = -\alpha(x_i, x_j) \quad ,$$

Unbekanntenvektor U,

$$U = (U_i)_{i=1}^N \quad ,$$

und rechter Seite F,

$$F = (F_i)_{i=1}^N \quad , \quad F_i = f_h(x_i) + \sum_{y \in Nb(x_i) \cap \partial\Omega_h} \alpha(x_i, y) g(y) \quad .$$

Unabhängig von einer bestimmten Numerierung sind in jeder Zeile von A nur höchstens 5 Koeffizienten von Null verschieden. A ist also *schwach besetzt.*
Gilt $\Omega_h = \Omega_h^\circ$, so ist A *symmetrisch*, d.h. $a_{ij} = a_{ji}$. Im allgemeinen gilt das leider nicht!

Definition 4.5 *Der* Graph *einer Matrix $A = (a_{ij})_{i,j=1}^N$ hat N Knoten und es gibt eine gerichtete Kante zwischen den Knoten i und j, falls $a_{ij} \neq 0$ ist. Die Matrix A heißt irreduzibel, falls der Graph von A zusammenhängend ist, d.h. falls je zwei Knoten i, j durch eine Folge von Knoten verbunden werden können.*

Bemerkung: Ist Ω_h diskret zusammenhängend, so ist A irreduzibel.

Auch andere Eigenschaften unserer Diskretisierung (4.3) lassen sich in entsprechende Eigenschaften des algebraische Gleichungssystems (4.4) übersetzen und umgekehrt. Wir haben zwei verschiedene Standpunkte zur Verfügung. Um Analogien zum kontinuierlichen Fall zu betonen, reden wir im folgenden lieber von Gitterfunktionen als von Vektoren und kommen nur hin und wieder auf die algebraische Blickweise zurück.

Definition 4.6 *Eine Gitterfunktion $U : \overline{\Omega}_h \to \mathbb{R}$ mit der Eigenschaft*

$$-\Delta_h U = 0 \qquad \text{auf } \Omega_h$$

heißt diskret harmonisch. Gilt nur

$$-\Delta_h U \leq 0 \qquad \text{auf } \Omega_h$$

so heißt U diskret subharmonisch.

Lemma 4.7 (Diskrete Mittelwerteigenschaft) *Ist U diskret subharmonisch, so hat U die diskrete Mittelwerteigenschaft*

$$U(x_Z) \;\leq\; \frac{1}{\alpha_W + \alpha_O + \alpha_S + \alpha_N} \left(\alpha_W U(x_W) + \alpha_O U(x_O) + \alpha_S U(x_S) + \alpha_N U(x_N) \right) \tag{4.6}$$

für alle $x_Z \in \Omega_h$.

Beweis:
Die Behauptung folgt aus

$$-\alpha_Z = \alpha_W + \alpha_O + \alpha_S + \alpha_N > 0 \quad .$$

□

Da wir vorausgesetzt haben, daß $\overline{\Omega}_h$ diskret zusammenhängend ist, folgt aus der diskreten Mittelwerteigenschaft eine weitere Analogie zum kontinuierlichen Fall.

Satz 4.8 (Diskretes Maximumprinzip) *Es sei U diskret subharmonisch. Dann nimmt U seinen maximalen Wert auf $\partial\Omega_h$ an oder ist konstant auf $\overline{\Omega}_h$.*

Beweis:

Sei $x_Z \in \Omega_h$ und

$$U(x_Z) = M = \max_{x\in\overline{\Omega}_h} U(x) \quad .$$

Dann folgt aus der diskreten Mittelwerteigenschaft

$$\begin{aligned} 0 \;\leq\; & \alpha_W\,(U(x_W) - U(x_Z)) \;+\; \alpha_O\,(U(x_O) - U(x_Z)) \;+ \\ & \alpha_S\,(U(x_S) - U(x_Z)) \;+\; \alpha_N\,(U(x_N) - U(x_Z)) \quad . \end{aligned}$$

Alle Summanden sind nicht-positiv da $U(x_Z) = M$. Es folgt also

$$U(x) = M, \quad \forall x \in Nb(x_Z) \quad .$$

Da je zwei Gitterpunkte aus $\overline{\Omega}_h$ durch eine Folge endlich vieler benachbarter Punkte verbunden werden können, folgt die Behauptung.

□

Weiter geht es wie im kontinuierlichen Fall.

Satz 4.9 *Das diskrete Randwertproblem* (4.3) *ist eindeutig lösbar.*

Beweis:

Sind $U^{(1)}, U^{(2)}$ zwei Lösungen von (4.3), so genügt $U = U^{(1)} - U^{(2)}$ den Gleichungen

$$\begin{aligned} -\Delta_h U(x) &= 0 \qquad x \in \Omega_h \\ U(x) &= 0 \qquad x \in \partial\Omega_h \quad . \end{aligned}$$

Insbesondere ist U diskret subharmonisch. Ist $\max_{x\in\Omega_h} U(x) = M > 0$, so folgt $U(x) = M$, $\forall x \in \overline{\Omega}_h$ aus dem diskreten Maximumprinzip. Das ist ein Widerspruch zu $U(x) = 0$, $x \in \partial\Omega_h$. Also muß $U(x) \leq 0$ für alle $x \in \Omega_h$ gelten. Die Annahme $\min_{x\in\Omega_h} U(x) = M < 0$ führt in gleicher Weise auf einen Widerspruch, denn auch $-U$ ist diskret subharmonisch.

Wir haben es mit einer linearen Abbildung auf einem *endlichdimensionalen Raum* zu tun. Aus der Injektivität folgt daher die Surjektivität.

□

Als letzte Folgerung aus dem diskreten Maximumprinzip erschließen wir die *inverse Monotonie* von $-\Delta_h$.

Lemma 4.10 *Der diskrete Laplace-Operator $-\Delta_h$ ist invers monoton, d.h. aus den Bedingungen*

$$\begin{aligned} -\Delta_h U(x) &\leq -\Delta_h V(x) \qquad x \in \Omega_h \\ U(x) &\leq V(x) \qquad x \in \partial\Omega_h \end{aligned}$$

folgt

$$U(x) \leq V(x) \quad , \quad x \in \overline{\Omega}_h \quad .$$

Beweis:

Offenbar gilt für $W = U - V$

$$\begin{aligned} -\Delta_h W(x) &\leq 0 \qquad & x &\in \Omega_h \\ W(x) &\leq 0 \qquad & x &\in \partial\Omega_h \quad . \end{aligned}$$

Wäre dann im Widerspruch zur Behauptung

$$\max_{x\in\Omega_h} W(x) = M > 0 \quad ,$$

so würde nach dem diskreten Maximumprinzip $W(x) = M > 0$, $\forall x \in \overline{\Omega}_h$ gelten. Dies wäre ein Widerspruch zu $W(x) \leq 0$, $x \in \partial\Omega_h$.

□

Lemma 4.10 werden wir später bei der Stabilität noch brauchen. Vorerst ergibt sich eine weitere interessante Eigenschaft der Koeffizientenmatrix A.

Da (4.3) eindeutig lösbar ist, muß A regulär sein. Setzt man in Lemma 4.10 $U(x) = 0$, $\forall x \in \partial\Omega_h$, so erhält man insbesondere

$$AU = F \leq 0 \quad \Longrightarrow \quad U \leq 0 \quad ,$$

wobei "$\leq$" komponentenweise gemeint ist. Wählt man $F = -e_i$ (i-ter Einheitsvektor), so erhält man die Lösung

$$U = -(A^{-1})_i \leq 0 \quad \text{i-te Spalte von } A^{-1} \quad .$$

Wir haben damit gezeigt, daß komponentenweise

$$A^{-1} \geq 0$$

gilt. Eine sorgfältigere Analyse zeigt sogar $A^{-1} > 0$ (siehe Übung). Das kommt nicht von ungefähr, denn A^{-1} spielt die Rolle einer diskreten Greenschen Funktion! Die Interpretation der Spalte $(A^{-1})_i$ als Gitterfunktion entspricht dabei $G(x_i, \cdot)$, denn

$$A(A^{-1})_i = e_i$$

ist ein diskretes Analogon zu

$$-\Delta_x G(x_i, x) = \delta_{x_i}(x) \qquad x \in \Omega \quad .$$

Die Greensche Funktion $G(x_i, \cdot)$ ist positiv. Diskretisierungen, welche diese Eigenschaft erhalten, sind von besonderer Bedeutung. Das motiviert die

Definition 4.11 (M-Matrix) *Eine Matrix $A = (a_{ij})_{i,j=1}^N$ heißt M-Matrix, falls gilt*

$$a_{ii} > 0 \quad , \quad a_{ij} \leq 0 \quad , \quad i, j = 1, \ldots, N$$
$$A \text{ ist regulär und } A^{-1} \geq 0 \quad .$$

Wir haben oben gezeigt, daß die unsere Diskretisierung (4.4) auf eine M-Matrix führt.

4.3 Konsistenz, Stabilität und Konvergenz

Zur Vorbereitung des abschließenden Konvergenzsatzes zeigen wir zunächst ein diskretes Analogon zur *a-priori-Stabilitätsabschätzung*

$$\max_{x\in\overline{\Omega}} |u(x)| \;\leq\; \max_{x\in\partial\Omega} |u(x)| \;+\; c \sup_{x\in\Omega} |\Delta u(x)| \,.$$

Erinnerung: Aus dieser Abschätzung folgt, daß das kontinuierliche Randwertproblem sachgemäß gestellt ist.

Satz 4.12 (Stabilität) *Sei U eine Gitterfunktion auf $\overline{\Omega}_h$ und sei R der Radius eines Kreises $K_R(0)$ mit $\overline{\Omega} \subset K_R(0)$. Dann gilt*

$$\max_{x\in\overline{\Omega}_h} |U(x)| \;\leq\; \max_{x\in\partial\Omega_h} |U(x)| \;+\; \frac{R^2}{2} \max_{x\in\Omega_h} |\Delta_h U(x)| \quad .$$

Beweis:

Der Beweis ist etwas technisch (trickreich). Wir betrachten die Wirkung von Δ_h auf

$$W(x) = \frac{1}{2}x_1^2 \quad , \quad V(x) = 1 \quad , \quad x = (x_1, x_2) \in \overline{\Omega}_h \quad .$$

Einsetzen liefert

$$\Delta_h W(x) = 1 \quad , \quad \Delta_h V(x) = 0 \quad , \quad x \in \Omega_h \quad .$$

Wir setzen

$$M = \max_{x\in\Omega_h} |\Delta_h U(x)| \quad , \quad N = \max_{x\in\partial\Omega_h} |U(x)| \quad .$$

Damit folgt für jedes der beiden Vorzeichen

$$\begin{aligned}
-\Delta_h \left(\pm U + MW\right)(x) \quad &\leq \quad 0 = -\Delta_h \left((N + \frac{R^2}{2}M)V\right)(x) \quad , \quad x \in \Omega_h \\
\pm U(x) + MW(x) \quad &\leq \quad (N + \frac{R^2}{2}M)V(x) \quad , \quad x \in \partial\Omega_h \quad .
\end{aligned}$$

Die inverse Monotonie aus Lemma 4.10 liefert

$$\pm U(x) + MW(x) \leq N + \frac{R^2}{2}MV(x) \quad , \quad x \in \overline{\Omega}_h \quad .$$

Wegen $MW(x) \geq 0$ ist also

$$\pm U(x) \leq N + \frac{R^2}{2}M \quad , \quad x \in \overline{\Omega}_h \quad .$$

Das ist die Behauptung.

□

Der obige Satz sichert die Stabilität von U gegenüber Störungen von $\Delta_h U$ und Störungen der Randdaten. Das nutzen wir jetzt aus.

Definition 4.13 *Setzt man die exakte Lösung in die Differenzenapproximation ein, so erhält man den* <u>*lokalen Abschneidefehler*</u>

$$\boxed{\begin{aligned}
\tau_h(x) &= f_h(x) + \Delta_h u(x) \qquad && x \in \Omega_h \\
\tau_h(x) &= u(x) - g_h(x) && x \in \partial\Omega_h.
\end{aligned}}$$

Das Differenzenverfahren heißt <u>*konsistent*</u>*, falls*

$$\max_{x\in\overline{\Omega}_h} |\tau_h(x)| \;\to\; 0 \quad , \quad h \;\to\; 0$$

und <u>*konsistent mit der Ordnung p*</u>*, falls*

$$\max_{x\in\overline{\Omega}_h} |\tau_h(x)| \;=\; \mathcal{O}\left(h^p\right)$$

gilt.

Satz 4.14 (Konsistenz) *Es sei $u \in C^3(\overline{\Omega})$. Dann gilt*

$$\boxed{\max_{x\in\overline{\Omega}_h} |\tau_h(x)| = \mathcal{O}\left(h\right).}$$

Beweis:

Offenbar ist $\tau_h(x) = 0 \quad$ für $x \in \partial\Omega_h$, und wegen $f_h(x) = f(x)$ haben wir

$$\tau_h(x) = f_h(x) + \Delta_h u(x) = -\Delta u(x) + \Delta_h u(x) = \mathcal{O}\left(h\right) \qquad \text{für } x \in \Omega_h$$

nach Lemma (4.4)

□

Bemerkung: Es gilt sogar $\tau_h(x) = \mathcal{O}\left(h^2\right)$, falls $x \in \Omega_h^\circ$ (und u genügend glatt) ist.

Aus Konsistenz und Stabilität folgt jetzt die Konvergenz.

Satz 4.15 (Konvergenz) *Es sei $u \in C^3(\overline{\Omega})$. Dann gilt*

$$\max_{x\in\overline{\Omega}_h} |U(x) - u(x)| = \mathcal{O}\left(h\right) \quad .$$

Beweis:

Wir setzen $V(x) = U(x) - u(x)$ für $x \in \overline{\Omega}_h$. Dann gilt nach Definition des Abschneidefehlers

$$\begin{aligned} -\Delta_h V(x) &= f_h(x) + \Delta_h u(x) = \tau_h(x) && x \in \Omega_h \\ V(x) &= \tau_h(x) && x \in \partial\Omega_h \quad . \end{aligned}$$

Aus der Stabilität folgt

$$\begin{aligned} \max_{x\in\overline{\Omega}_h} |U(x) - u(x)| &= \max_{x\in\overline{\Omega}_h} |V(x)| \\ &\leq \max_{x\in\partial\Omega_h} |\tau_h(x)| + \frac{R^2}{2} \max_{x\in\Omega_h} |\tau_h(x)| \;=\; \mathcal{O}\left(h\right). \end{aligned}$$

□

Unser Differenzenverfahren ist also konvergent. Konvergenz– und Konsistenzordnung stimmen überein. Ein Blick auf den Konvergenzbeweis zeigt, daß nur die schlechte Konsistenz in Randnähe

$$\tau_h(x) = \mathcal{O}\left(h\right) \quad , \quad x \in \Omega_h \setminus \Omega_h^\circ$$

die Konvergenzordnung 2 verhindert (falls $u \in C^4(\overline{\Omega})$).

Für verschwindende Schrittweite h ist die Anzahl der irregulären Punkte $\#\Omega_h \setminus \Omega_h^\circ$ um eine Größenordnung kleiner als die Anzahl der inneren Punkte Ω_h°. Dies läßt auf eine mögliche Verschärfung unseres Konvergenzresultats hoffen. Eine genauere Analyse zeigt tatsächlich

Satz 4.16 *Es sei $u \in C^4(\overline{\Omega})$. Dann gilt*

$$\max_{x \in \overline{\Omega}_h} |U(x) - u(x)| = \mathcal{O}\left(h^2\right) \quad .$$

Beweis:

Siehe z.B. Hackbusch [12], S. 82 .

□

5 Schwache Lösungen

Aus Abschnitt 1.1.1 wissen wir, daß die Auslenkung u einer fest eingespannten Membran Ω durch eine Kraftdichte f Lösung der elliptischen Randwertaufgabe

$$\begin{aligned} -\operatorname{div}(\alpha \nabla u) &= f \qquad \text{in } \Omega \\ u &= 0 \qquad \text{auf } \partial\Omega \end{aligned} \tag{5.1}$$

ist. Dabei ist $\alpha > 0$ eine Materialkonstante. Wir betrachten nun eine Membran aus zwei verschiedenen Materialien. Dann ist

$$\alpha(x) = \begin{cases} \alpha_1 & x \in \Omega_1 \\ \alpha_2 & x \in \Omega_2 \end{cases} ,$$

wobei α_1, $\alpha_2 \in \mathbb{R}$ und die Teilgebiete Ω_1, Ω_2 mit $\Omega_1 \cup \Omega_2 \cup \Gamma = \Omega$, $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ den verschiedenen Materialien entsprechen.

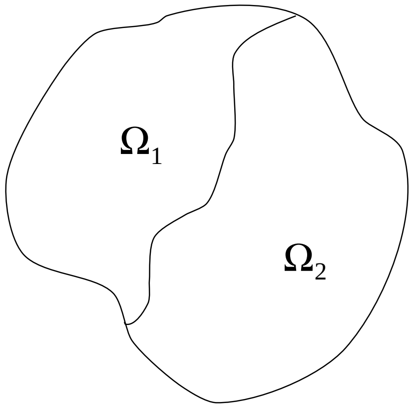


Abbildung 8: Übergangsproblem

Aus $\alpha_1 \neq \alpha_2$ ergibt sich nun das Dilemma , daß *entweder* ∇u *oder* $\alpha \nabla u$ in Ω stetig sein kann. In beiden Fällen kann (5.1) (im klassischen Sinn) nicht erfüllt sein. Ein möglicher Ausweg besteht darin, (5.1) durch ein sogenanntes *Übergangsproblem* zu ersetzen. Dabei fordern wir die Gültigkeit von (5.1) nur im Innern von Ω_1 und Ω_2. Am inneren Rand $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ stellt man die Übergangsbedingungen

$$\begin{aligned} u_1 &= u_2 \qquad \text{auf } \Gamma \\ \frac{\partial}{\partial n} u_1 &= \frac{\partial}{\partial n} u_2 \qquad \text{auf } \Gamma \quad , \end{aligned} \tag{5.2}$$

wobei $u_i = u_{|\Omega_i}$ für $i = 1, 2$ gesetzt ist. Nun wären Existenz, Eindeutigkeit und Stabilität der Lösung von Übergangsproblemen zu klären.

Bei Anwendung von Differenzenverfahren müssen die Übergangsbedingungen (5.2) gesondert diskretisiert werden. Bei krummlinigem Übergang Γ ist insbesondere die Diskretisierung der Normalenableitung unbequem. Die Schwierigkeiten werden noch größer, wenn $\Gamma = \Gamma(u)$ von der unbekannten Funktion u abhängt (Wärmeleitung, poröse Medien). Im Laufe der iterativen Lösung des resultierenden nichtlinearen Gleichungssystems muß dann die gesamte Diskretisierung der Übergangsbedinungen an den jeweils aktuellen Übergang $\Gamma(U_h^\nu)$ angepasst werden.

Solche Schwierigkeiten sind typisch für Differenzenverfahren. Deren grundsätzliche Nachteile sind

- mangelnde Flexibilität,
- hohe Regularitätsanforderungen.

Um robustere Lösungsverfahren zu entwickeln, erinnern wir uns, daß die mathematische Modellierung zunächst auf das Minimierungsproblem

$$u \in H : \; J(u) \leq J(v) \qquad \forall v \in H \tag{5.3}$$

für das Energiefunktional

$$J(v) = \frac{1}{2} \int_{\Omega} \alpha \, |\nabla v|^2 \, dx - \int_{\Omega} f v \, dx \tag{5.4}$$

führte. Erst anschließend wurde unter zusätzlichen Regularitätsannahmen die Formulierung als Differentialgleichung abgeleitet. Diese Annahmen treffen bei nichtglatten Daten, z.B. bei unstetigem α nicht zu! Vor diesem Hintergrund konzentrieren wir uns von nun an direkt auf die Lösung des Minimierungsproblems (5.2). Dabei haben wir wieder Existenz, Eindeutigkeit und stetige Abhängigkeit von den Daten zu klären. Das läuft insbesondere auf die richtige Wahl des Lösungsraums H hinaus.

5.1 Hilberträume

Wir erinnern kurz an einige Grundbegriffe aus der Funktionalanalysis. Eine anwendungsorientierte Einführung in dieses Gebiet gibt zum Beispiel Alt [2].

Definition 5.1

(a) *Sei V ein linearer Raum. Dann heißt eine Abbildung*

$$\|\cdot\| : \; V \to \mathbb{R}$$

Norm auf V, wenn für alle $v, w \in V$ und $\mu \in \mathbb{R}$ gilt:

$$\begin{aligned} \|v\| &\geq 0 \quad \text{und} \quad \|v\| = 0 \iff v = 0 \\ \|\mu v\| &= |\mu| \, \|v\| \\ \|v + w\| &\leq \|v\| + \|w\| \end{aligned}$$

Ist auf V eine Norm $\|\cdot\|_V$ definiert, so ist V ein normierter, linearer Raum.

(b) *Ein normierter, linearer Raum V heißt vollständig, wenn jede Cauchy–Folge in V konvergiert.*

(c) *Ein vollständiger, normierter, linearer Raum B heißt Banachraum.*

Beispiel:

Sei $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet. Dann ist $C(\overline{\Omega})$, versehen mit der Norm

$$\|v\|_\infty = \max_{x \in \overline{\Omega}} |v(x)| \quad ,$$

ein Banachraum.

Beispiel:

Der lineare Raum $C^1(\overline{\Omega})$, versehen mit der Norm

$$\|v\|_{1,\infty} = \|v\|_\infty + \sum_{i=1}^{n} \|v_{x_i}\|_\infty \quad ,$$

ist ein Banachraum.

Definition 5.2

(a) *Seien V und W normierte, lineare Räume mit den Normen $\|\cdot\|_V$ und $\|\cdot\|_W$. Dann heißt eine lineare Abbildung $L : V \to W$ beschränkt, wenn es eine Zahl $c > 0$ gibt mit*

$$\|Lv\|_W \leq c \, \|v\|_V \qquad \forall v \in V \quad .$$

(b) *Die Menge aller beschränkten, linearen Abbildungen $L : V \to W$ bezeichnen wir mit $\mathcal{L}(V, W)$.*

(c) *$V' = \mathcal{L}(V, \mathbb{R})$ ist der Dualraum von V. Die Elemente von V' heißen Funktionale.*

Satz 5.3 *Ein lineare Abbildung $L : V \to W$ ist genau dann stetig, wenn sie beschränkt ist.*

Satz 5.4 *Durch*

$$\|L\| = \sup_{v \in V} \frac{\|Lv\|_W}{\|v\|_V}, \quad L \in \mathcal{L}(V, W),$$

wird die kanonische Norm auf $\mathcal{L}(V, W)$ definiert.

Definition 5.5

(a) *Eine Abbildung $a(\cdot, \cdot) : V \times V \to \mathbb{R}$ heißt Bilinearform (auf V), wenn sie linear in beiden Argumenten ist.*

(b) *$a(\cdot, \cdot)$ heißt symmetrisch, falls*

$$a(v, w) = a(w, v) \quad \forall v, w \in V \quad .$$

(c) *$a(\cdot,\cdot)$ heißt positiv definit, falls*

$$a(v,v) \geq 0 \quad und \quad a(v,v) = 0 \iff v = 0 \quad \forall v \in V \quad .$$

(d) *$a(\cdot,\cdot)$ heißt beschränkt, falls es eine Zahl Γ gibt, so daß*

$$|a(v,w)| \leq \Gamma \|v\|_V \|w\|_V \quad \forall v,w \in V \quad .$$

Satz 5.6 *Eine Bilinearform $a(\cdot,\cdot)$ ist genau dann stetig, wenn sie beschränkt ist.*

Satz 5.7 *Eine symmetrische, positiv definite Bilinearform $a(\cdot,\cdot)$ genügt der Cauchy–Schwarzschen Ungleichung*

$$|a(v,w)| \leq a(v,v)^{\frac{1}{2}} \cdot a(w,w)^{\frac{1}{2}} \qquad \forall v,w \in V$$

und der Dreiecksungleichung

$$a(v+w,v+w)^{1/2} \leq a(v,v)^{\frac{1}{2}} + a(w,w)^{\frac{1}{2}} \qquad \forall v,w \in V$$

Insbesondere ist durch

$$\|v\| = a(v,v)^{\frac{1}{2}}, \qquad v \in V,$$

eine Norm, die sogenannte Energienorm, definiert.

Definition 5.8

(a) *Eine symmetrische, positiv definite Bilinearform heißt Skalarprodukt.*

(b) *Ein linearer Raum V versehen mit einem Skalarprodukt $(\cdot,\cdot)$ und der zugehörigen Norm $\|\cdot\|_V$ heißt Prä–Hilbertraum.*

(c) *Ein vollständiger Prä–Hilbertraum heißt Hilbertraum.*

Beispiel:

Q^n versehen mit dem Euklidischen Skalarprodukt $(v,w) = \sum_i v_i w_i$ ist ein Prä–Hilbertraum. Die Vervollständigung $\mathbb{R}^n$ ist ein Hilbertraum.

Beispiel:

Der lineare Raum

$$X = \{v \in C(\Omega) \mid \int_\Omega v^2(x)\, dx < \infty\}$$

mit dem Skalarprodukt

$$(v,w)_{L^2(\Omega)} = \int_\Omega v(x)w(x)\, dx$$

ist ein Prä-Hilbertraum.

Beispiel:

Der lineare Raum

$$X = \{v \in C^1(\Omega) \mid (v,v)_{H^1(\Omega)} < \infty\}$$

mit dem Skalarprodukt

$$(v,w)_{H^1(\Omega)} = (v,w)_{L^2(\Omega)} + \sum_{i=1}^{n} (v_{x_i}, w_{x_i})_{L^2(\Omega)}$$

ist ein Prä-Hilbertraum.

Beispiel:

Sei $\alpha : \Omega \to \mathbb{R}$ stückweise stetig (stetig bis auf Riemann–Nullmengen) und es gelte mit $\alpha_0, \alpha_1 \in \mathbb{R}$

$$0 < \alpha_0 \leq \alpha(x) \leq \alpha_1 < \infty \qquad \forall x \in \Omega$$

mit $\alpha_0, \alpha_1 \in \mathbb{R}$. Dann ist

$$X = \{v \in C^1(\Omega) \cap C(\overline{\Omega}) \mid v|_{\partial\Omega} = 0, (v,v)_{H^1(\Omega)} < \infty\} \subset H_C$$

ein Prä–Hilbertraum mit dem Energie–Skalarprodukt

$$a(v,w) = \int\limits_{\Omega} \alpha \nabla v \cdot \nabla w \, dx \quad .$$

Satz 5.9 *Jeder Prä–Hilbertraum V läßt sich bis auf Isomorphie in eindeutig bestimmter Weise zu einem Hilbertraum H erweitern. V liegt dann dicht in H, und man nennt H Vervollständigung von V.*

Satz 5.10 *Sei V ein linearer Raum mit der Norm $\|\cdot\|_V$. Auf V existiert genau dann ein Skalarprodukt $(\cdot,\cdot)$ welches $\|\cdot\|_V$ erzeugt, also*

$$\|v\|_V = (v,v)^{\frac{1}{2}} \qquad \forall v \in V$$

erfüllt, wenn die Parallelogramm–Gleichung

$$\|v+w\|_V^2 + \|v-w\|_V^2 = 2\|v\|_V^2 + 2\|w\|_V^2 \qquad \forall v,w \in V$$

erfüllt ist.

Beispiel:

Die Maximumsnorm $\|\cdot\|_{1,\infty}$ genügt der Parallelogrammgleichung nicht, also ist $C^1(\overline{\Omega})$ *kein* Hilbertraum!

Bevor wir uns speziellen Funktionenräumen zuwenden, wollen wir den abstrakten Rahmen nutzen, um einige grundlegende Aussagen bereitzustellen. Von nun an sei H ein Hilbertraum mit dem Skalarprodukt $(\cdot,\cdot)$ und der zugehörigen Norm $\|\cdot\|_H = (\cdot,\cdot)^{1/2}$.

Theorem 5.11 (Rieszscher Darstellungssatz) *Für jedes $l \in H'$ besitzt die Variationsaufgabe*

$$u \in H : \quad (u, v) = l(v) \qquad \forall v \in H \tag{5.5}$$

eine eindeutig bestimmte Lösung und es gilt

$$\|u\|_H = \|l\|_{H'} \quad .$$

Beweis:

Der Existenzbeweis fällt unter die *direkten Methoden der Variationsrechnung*. Wir setzen $J(v) = \frac{1}{2}(v, v) - l(v)$. Wortwörtlich wie im Beweis zu Satz 1.2 (Seite 3) folgt die Äquivalenz von (5.4) mit dem Minimierungsproblem

$$u \in H : \quad J(u) \leq J(v) \qquad \forall v \in H \quad . \tag{5.6}$$

Es reicht also die eindeutige Lösbarkeit dieses Minimierungsproblems zu zeigen.

1. Wir weisen zunächst nach, daß $J(v)$ nach unten beschränkt ist, daß also ein $M \in \mathbb{R}$ existiert, so daß

   $$-\infty < M \leq J(v) \qquad \forall v \in H \quad .$$

   Wegen der Beschränktheit von l ist

   $$|l(v)| \leq \|l\|_{H'} \|v\|_H$$

   und es folgt

   $$\begin{aligned} J(v) &= \frac{1}{2}(v, v) - l(v) = \frac{1}{2}\|v\|_H^2 - l(v) \\ &\geq \frac{1}{2}\|v\|_H^2 - |l(v)| \\ &\geq \frac{1}{2}\|v\|_H^2 - \|l\|_{H'}\|v\|_H \quad . \end{aligned}$$

   Die Parabel $\frac{1}{2}\|v\|_H^2 - \|l\|_{H'}\|v\|_H$ ist für alle $\|v\|_H$ nach unten beschränkt, also auch $J(v)$.

2. Da $J(v)$ nach unten beschränkt ist, existiert das Infimum

   $$\beta = \inf_{v \in H} J(v) \quad .$$

   Nach Definition des Infimums gibt es eine Minimalfolge $(u_n)_{n \in \mathbb{N}} \subset H$, also eine Folge mit der Eigenschaft

   $$J(u_n) \to \beta \quad \text{für} \quad n \to \infty \quad .$$

   Wir zeigen, daß $(u_n)_{n \in \mathbb{N}}$ eine Cauchy–Folge ist. Zunächst bemerkt man, daß wegen der *Konvexität von J*, also wegen

   $$\beta \leq J\left(\tfrac{u_n + u_m}{2}\right) \leq \frac{1}{2}\left(J(u_n) + J(u_m)\right) \quad \to \beta$$

auch $J(\frac{u_n+u_m}{2})$ gegen β konvergiert. Aus der Parallelogrammgleichung folgt

$$\begin{aligned} \|u_n - u_m\|_H^2 &= 2\|u_n\|_H^2 + 2\|u_m\|_H^2 - \|u_n + u_m\|_H^2 \\ &= 4J(u_n) + 4J(u_m) + 4l(u_n + u_m) - 4\left\|\frac{u_n + u_m}{2}\right\|_H^2 \\ &= 4J(u_n) + 4J(u_m) - 8\left(\frac{1}{2}\left\|\frac{u_n + u_m}{2}\right\|_H^2 - l(\frac{u_n + u_m}{2})\right) \\ &= 4J(u_n) + 4J(u_m) - 8J(\frac{u_n + u_m}{2}) \to 0 \quad . \end{aligned}$$

Damit ist $(u_n)_{n\in\mathbb{N}}$ eine Cauchy–Folge und wegen der Vollständigkeit von H gibt es ein $u^* \in H$ mit

$$\lim_{n\to\infty} u_n = u^* \qquad \text{in } H \quad .$$

3. Die Stetigkeit von J liefert

$$\beta = \lim_{n\to\infty} J(u_n) = J(u^*) \quad .$$

u^* ist also eine Lösung des Minimierungsproblems (5.5).

4. Die Lösung u^* ist eindeutig. Gäbe es zwei Lösungen u_1 und u_2 von (5.4) wäre nämlich

$$(u_1, v) = l(v) = (u_2, v) \qquad \forall v \in H$$

also

$$(u_1 - u_2, v) = 0 \qquad \forall v \in H \quad .$$

Einsetzen von $v = u_1 - u_2$ liefert die Behauptung.

5. $\|u\|_H = \|l\|_{H'}$ folgt aus den Ungleichungen

$$\|l\|_{H'} = \sup_{v\neq 0} \frac{|l(v)|}{\|v\|_H} = \sup_{v\neq 0} \frac{|(u,v)|}{\|v\|_H} \leq \|u\|_H \qquad \text{Cauchy–Schwarz}$$

und

$$\|u\|_H^2 = (u,u) = l(u) \leq \|l\|_{H'} \|u\|_H \quad .$$

□

Bemerkung: Der Rieszsche Darstellungssatz besagt, daß jedes $l \in H'$ über das Skalarprodukt durch ein $u \in H$ mit gleicher Norm dargestellt werden kann. *Der Hibertraum H und sein Dualraum H' sind damit isometrisch isomorph*, d. h. sie unterscheiden sich topologisch und algebraisch nicht.

Korollar 5.12 *Sei $l \in H'$ und $a(\cdot,\cdot)$ eine symmetrische Bilinearform auf H mit der Eigenschaft*

$$\gamma\|v\|_H^2 \leq a(v,v) \qquad \forall v \in H \tag{5.7}$$

wobei γ nicht von v abhängt. Dann hat das Minimierungsproblem

$$u \in H: \qquad J(u) \leq J(v) \qquad \forall v \in H$$

für das Energiefunktional

$$J(v) = \frac{1}{2}a(v,v) - l(v), \qquad v \in H \quad ,$$

bzw. dessen Variationsformulierung

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

eine eindeutig bestimmte Lösung. Außerdem gilt

$$\|u\|_H \leq \frac{1}{\gamma}\,\|l\|_{H'} \quad .$$

Beweis:

Wegen (5.6) ist $a(\cdot,\cdot)$ ein Skalarprodukt auf H und $l \in H'$ ist auch bezüglich der Energienorm beschränkt. Wir können also Theorem 5.11 anwenden und erhalten Existenz und Eindeutigkeit. Der Rest folgt aus

$$\gamma\,\|u\|_H^2 \leq a(u,u) = l(u) \leq \|l\|_{H'}\,\|u\|_H \quad .$$

□

Korollar 5.12 motiviert die folgende

Definition 5.13 *Eine beschränkte Bilinearform $a(\cdot,\cdot)$ heißt ($H-$)<u>elliptisch</u> oder <u>koerziv</u>, wenn die Abschätzungen*

$$\boxed{\gamma\|v\|_H^2 \leq a(v,v), \qquad |a(v,w)| \leq \Gamma\|v\|_H\|w\|_H \qquad \forall v \in H} \tag{5.8}$$

gelten.

Korollar 5.14 (Bestapproximation) *Sei S ein abgeschlossener Unterraum von H und $w_0 \notin S$. Dann existiert ein $w \in S$ mit*

$$\|w - w_0\|_H = \min_{v \in S} \|w_0 - v\|_H \quad ,$$

und es gilt die Orthogonalität

$$(w_0 - w, v) = 0 \qquad \forall v \in S \quad .$$

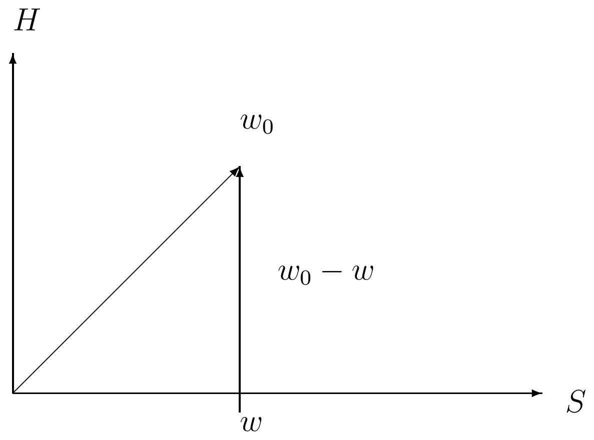


Abbildung 9: Bestapproximation im Hilbertraum

Beweis:
Setzt man $J(v) = \frac{1}{2}(v, v) - \ell_0(v)$ mit $\ell_0(v) = (v, w_0)$, so ist offenbar

$$\|v - w_0\|_H^2 = 2\,(J(v) + (w_0, w_0)) \quad .$$

S ist als abgeschlossener Unterraum wieder ein Hilbertraum. Damit existiert nach dem Rieszsche Darstellungssatz (Theorem 5.11) genau ein $w \in S$ mit der Eigenschaft

$$J(w) \leq J(v) \qquad \forall v \in S \quad .$$

Die Orthogonalität ist gerade die äquivalente Variationsformulierung

$$w \in H : \qquad (w, v) = (w_0, v) \qquad \forall v \in H \quad .$$

□

Wir wollen nun die Aussage von Korollar 5.12 verallgemeinern. Kernpunkt ist, daß die Bilinearform nicht mehr symmetrisch sein muß. *Beachte, daß der Variationsaufgabe dann kein Minimierungsproblem mehr zugrunde liegt*!

Theorem 5.15 (Lax–Milgram–Lemma) *Sei $a(\cdot,\cdot)$ H–elliptisch (nicht notwendig symmetrisch!). Dann hat die Variationsgleichung*

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

für jedes $l \in H'$ eine eindeutig bestimmte Lösung, und es gilt

$$\|u\|_H \leq \frac{1}{\gamma} \|l\|_{H'} \quad .$$

Beweis:

1.) Wir besorgen uns eine Operatordarstellung der Bilinearform.
Sei $w \in H$ beliebig. Dann ist

$$f_w := a(w,\cdot) \in H' \quad ,$$

denn f_w ist linear und auch stetig, weil

$$|f_w(v)| = |a(w,v)| \leq \Gamma \|w\|_H \|v\|_H \quad .$$

Damit gilt auch

$$\|f_w\|_{H'} \leq \Gamma \|w\|_H \quad .$$

Nach dem Rieszschen Darstellungssatz gibt es daher zu jedem $w \in H$ ein $Aw \in H$ mit der Eigenschaft

$$(Aw,v) = f_w(v) = a(w,v) \qquad \forall v \in H \quad .$$

In diesem Sinne stellt die so definierte Abbildung

$$A : H \to H$$

die Bilinearform $a(\cdot,\cdot)$ dar.

2.) Eigenschaften von A

- A ist linear, denn für alle $v \in H$ ist

$$\begin{aligned} (A(\mu_1 w_1 + \mu_2 w_2), v) &= a(\mu_1 w_1 + \mu_2 w_2, v) \\ &= \mu_1 a(w_1,v) + \mu_2 a(w_2,v) \\ &= \mu_1 (Aw_1,v) + \mu_2 (Aw_2,v) \quad . \end{aligned}$$

- A ist beschränkt, d. h. $\|Av\|_H \leq \Gamma \|v\|_H \quad \forall v \in H$, denn

$$\|Av\|_H^2 = (Av,Av) = a(v,Av) \leq \Gamma \|v\|_H \|Av\|_H \quad .$$

- A ist invers beschränkt, d. h. $\gamma \|v\|_H \leq \|Av\|_H \quad \forall v \in H$, denn

$$\gamma \|v\|_H^2 \leq |a(v,v)| = |(Av,v)| \leq \|Av\|_H \|v\|_H$$

- A ist injektiv, denn $Av = 0 \;\Rightarrow\; v = 0$.

3.) Eigenschaften von $R(A)$ und $R(A)^\perp$

Wir definieren für A den Wertebereich $R(A)$ und das orthogonale Komplement $R(A)^\perp$.

$$\begin{aligned} R(A) &:= \{w \in H \,|\, \exists v \in H \text{ mit } w = Av\} \\ R(A)^\perp &:= \{v \in H \,|\, (w,v) = 0 \quad \forall w \in R(A)\} \end{aligned}$$

- $R(A)$ ist abgeschlossen, denn aus

$$(w_n)_{n\in\mathbb{N}} \in R(A) \quad \text{und} \quad w_n \to w_0$$

folgt mit $Av_n = w_n$ wegen

$$\|w_n - w_m\|_H = \|A(v_n - v_m)\|_H \geq \gamma \|v_n - v_m\|_H \quad ,$$

daß $(v_n)_{n\in\mathbb{N}}$ eine Cauchy–Folge ist. Also konvergiert $v_n \to v_0$, und aus der Stetigkeit von A folgt $Av_0 = w_0$, also $w_0 \in R(A)$.

- Es gilt $R(A)^\perp = \{0\}$, denn ist $w_0 \in R(A)^\perp$, also

$$(w_0, w) = 0 \qquad \forall w \in R(A) \quad ,$$

so folgt

$$0 = (w_0, w) = (w_0, Av) \qquad \forall v \in H \quad ,$$

und insbesondere

$$0 = (w_0, Aw_0) = a(w_0, w_0) \geq \gamma \|w_0\|_H^2 \quad .$$

4.) Wir zeigen nun $R(A) = H$. Dazu verwenden wir Korollar 5.14 mit $S = R(A)$ (R(A) ist abgeschlossen!). Ist $w_0 \in H \setminus R(A)$, so existiert dann nämlich ein $w \in R(A)$ mit

$$(w_0 - w, v) = 0 \qquad \forall v \in R(A) \quad ,$$

also ist $w_0 - w \in R(A)^\perp$ und damit aber $w_0 - w = 0$ im Widerspruch zu $w_0 \notin R(A)$.

5.) Existenz und Eindeutigkeit

Da $R(A) = H$ ist, gibt es zu jedem $w \in H$ ein $u \in H$ mit $Au = w$. Wählt man nach dem Rieszschen Darstellungssatz w so, daß

$$(w, v) = l(v) \qquad \forall v \in H \quad ,$$

so folgt

$$a(u, v) = (Au, v) = (w, v) = l(v) \qquad \forall v \in H \quad .$$

Die Eindeutigkeit folgt aus der Injektivität von A.

Die Stabilitätsabschätzung folgt wie im Beweis von Korollar 5.12.

□

Approximation der Lösung durch Ritz-Galerkin-Verfahren. Soweit zur Existenz und Eindeutigkeit. Nun zur Approximation von u durch das Ritz–Galerkin–Verfahren.

Theorem 5.16 (Céa–Lemma) *Sei $a(\cdot,\cdot)$ H–elliptisch, $l \in H'$ und $S \subset H$ ein abgeschlossener Unterraum von H. Dann hat das Variationsproblem*

$$u_s \in S : \qquad a(u_s, v) = l(v) \qquad \forall v \in S$$

genau eine Lösung u_s, und es gilt die a priori Fehlerabschätzung

$$\|u - u_s\|_H \leq \frac{\Gamma}{\gamma} \inf_{v \in S} \|u - v\|_H \quad .$$

Beweis:

1.) Existenz und Eindeutigkeit

Die eindeutige Lösbarkeit folgt unmittelbar aus dem Lax–Milgram–Lemma, da S als abgeschlossener Unterraum eines Hilbertraums wieder ein Hilbertraum ist.

2.) Galerkin–Orthogonalität

Offenbar gilt

$$a(u_s, v) = l(v) = a(u, v) \qquad \forall v \in S \quad ,$$

also

$$a(u - u_s, v) = 0 \qquad \forall v \in S \quad .$$

Damit ist der Fehler a–orthogonal auf S.

3.) Fehlerabschätzung

Aus

$$\begin{aligned}
\gamma \|u - u_s\|^2 &\leq a(u - u_s, u) - \underbrace{a(u - u_s, u_s)}_{=0} = a(u - u_s, u) \\
&= a(u - u_s, u) - \underbrace{a(u - u_s, v)}_{=0} = a(u - u_s, u - v) \\
&\leq \Gamma \|u - u_s\| \, \|u - v\|
\end{aligned}$$

für alle $v \in S$ folgt sofort

$$\|u - u_s\| \leq \frac{\Gamma}{\gamma} \|u - v\| \quad .$$

□

5.2 Sobolev–Räume

5.2.1 Vervollständigung

Damit wir die abstrakten Resultate aus dem vorigen Abschnitt auf unser Minimierungsproblem (5.2) anwenden können, müssen wir als Lösungsraum einen geeigneten Hilbertraum H wählen. Dabei stehen wir vor folgendem Dilemma.

a) Der Raum aller Funktionen $v \in C^1(\overline{\Omega})$ mit der Eigenschaft $v|_{\partial\Omega} = 0$ ist, versehen mit der Norm $\|v\|_{1,\infty}$, zwar vollständig, aber kein Hilbertraum. Wegen

$$\|v+w\|_{1,\infty}^2 + \|v-w\|_{1,\infty}^2 \neq 2(\|v\|_{1,\infty}^2 + \|w\|_{1,\infty}^2)$$

(Übung!) gibt es nämlich kein Skalarprodukt, das $\|\cdot\|_{1,\infty}$ erzeugt.

b) Als Beispiel für einen Prä–Hilbertraum haben wir auf Seite 61 den Raum

$$X = \{v \in C^1(\Omega) \cap C(\overline{\Omega}) \mid v|_{\partial\Omega} = 0, (v,v)_{H^1(\Omega)} < \infty\} \tag{5.9}$$

mit dem Skalarprodukt

$$(v,w)_{H^1(\Omega)} = (v,w)_{L^2(\Omega)} + \sum_{i=1}^{n} (v_{x_i}, w_{x_i})_{L^2(\Omega)}$$

eingeführt. Dieser Raum ist zwar größer als unser erstbester Lösungsraum H_C für die ausgelenkte Membran (vgl. Seite 3). Trotzdem ist X nicht vollständig.

Das ergibt sich aus folgendem Gegenbeispiel. Sei $\Omega = (-1,1)$. Die Folge $(v_k)_{k\in\mathbb{N}} \subset X$ definiert durch

$$v_k(x) = \begin{cases} 1+x & -1 \leq x \leq -\frac{1}{n} \\ 1 - \frac{1}{2n} - \frac{n}{2}x^2 & -\frac{1}{n} \leq x \leq +\frac{1}{n} \\ 1-x & +\frac{1}{n} \leq x \leq 1 \end{cases}$$

ist zwar Cauchy–Folge bezüglich $\|\cdot\|_{H^1(\Omega)}$ und konvergiert gleichmäßig gegen

$$v^*(x) = \begin{cases} 1+x & -1 \leq x \leq 0 \\ 1-x & 0 \leq x \leq 1 \end{cases}$$

aber $v^* \notin X$, denn $v^* \notin C^1(\Omega)$ (Übung).

Um den Prä–Hilbertraum X aus (5.8) zu einem Hilbertraum zu machen, müssen wir X vervollständigen. Wir setzen

$$H_0^1(\Omega) = \text{Vervollständigung von } X \text{ bezüglich } \|\cdot\|_{H^1(\Omega)} = (\cdot,\cdot)_{H^1(\Omega)}.$$

und können dann den Hilbertraum $H = H_0^1(\Omega)$ als Lösungsraum wählen. Leider wissen wir davon noch nicht, wie wir uns eine Lösung, oder allgemeiner, eine Funktion $v \in H$ vorzustellen haben.

5.2.2 Grundlegende Eigenschaften von Sobolev–Räumen

$H_0^1(\Omega)$ ist ein spezieller *Sobolevraum*. Wir werden im folgenden weitere Sobolevräume durch Vervollständigung definieren und einige grundlegende Eigenschaften angeben. Wer etwas mehr erfahren will, kann sich zum Beispiel bei Alt [2] informieren. Noch mehr steht bei Adams [1].

1. Äquivalenzklassen. Die Elemente von Sobolev–Räumen sind Äquivalenzklassen $[v]$ von Funktionen. Die zugehörige Äquivalenzrelation ist gegeben durch

$$v_1 \sim v_2 \iff \text{meas}\{x \in \Omega \,|\, v_1(x) \neq v_2(x)\} = 0 \quad .$$

Dabei bedeutet meas das n–dimensionale Lebesgue–Maß.

Definition 5.17 *$\Omega' \subset \mathbb{R}^n$ ist eine Lebesgue–Nullmenge genau dann, wenn es zu jedem $\varepsilon > 0$ eine Menge von Quadern $\{I_k \,|\, k \in \mathbb{N}\}$ gibt, so daß*

$$\Omega' \subset \bigcup_{k \in \mathbb{N}} I_k \quad \text{und} \quad 0 \leq \sum_{k \in \mathbb{N}} \text{meas}\{I_k\} < \varepsilon \quad .$$

Gilt eine Aussage für alle $x \in \Omega$ bis auf eine Lebesgue–Nullmenge, so sagen wir, die Aussage gilt fast überall (f.ü.) in Ω.

Bemerkung: Im Unterschied zum Riemann–Maß sind *abzählbar unendlich* viele Quader erlaubt!

Beispiel:

Jede abzählbare Menge $\Omega' = \{z_k \in \mathbb{R}^n \,|\, k \in \mathbb{N}\}$ ist Lebesgue–Nullmenge.

Als erstes charakterisieren wir den Sobolev–Raum, welcher durch Vervollständigung der stetigen, quadratintegrablen Funktionen entsteht.

Satz 5.18 *Durch Vervollständigung von*

$$X = \{v \in C(\Omega) \mid \int_\Omega v^2 \, dx < \infty\}$$

mit dem Skalarprodukt

$$(v, w)_{L^2(\Omega)} = \int_\Omega vw \, dx$$

erhält man den Raum aller Äquivalenzklassen $[v]$ von (Lebesgue–meßbaren) Funktionen v mit der Eigenschaft

$$\|v\|_{L^2(\Omega)} = \left(\int_\Omega |v(x)|^2 \, dx \right)^{\frac{1}{2}} < \infty \quad .$$

Beispiel:

Die *Dirichlet–Funktion*

$$v = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

liegt in $L^2(\mathbb{R})$ und es gilt $v \in [0]$.

Alle äquivalenten Funktionen haben dieselbe Norm! In Zukunft wird der Einfachheit halber nicht mehr zwischen der Äquivalenzklasse $[v]$ und einem Repräsentanten $v \in [v]$ unterschieden. Aber Vorsicht!

Noch ein Approximationsresultat zum Abschluß.

Satz 5.19 *Der Teilraum*

$$C_0^\infty(\Omega) = \{v \in C^\infty(\Omega) \mid \text{supp } v \subset \Omega\}$$

der unendlich oft differenzierbaren Funktionen mit kompaktem Träger in Ω

$$\text{supp } \varphi = \overline{\{x \in \Omega \,|\, \varphi(x) \neq 0\}}$$

liegt dicht in $L^2(\Omega)$*, d.h. zu jedem* $v \in L^2(\Omega)$ *und jedem* $\varepsilon > 0$ *existiert ein* φ_ε *mit der Eigenschaft*

$$\|v - \varphi_\varepsilon\|_{L^2(\Omega)} \leq \varepsilon \quad .$$

Beweis:

Siehe Alt [2], Seite 73.

□

2. Schwache Ableitungen. Funktionen $v \in L^2(\Omega)$ sind nur bis auf Nullmengen bestimmt und daher sicher nicht mehr im klassischen Sinne differenzierbar. Wir erweitern also unseren Ableitungsbegriff.

Definition 5.20 (Schwache Ableitung) *Sei* $u \in L^2(\Omega)$*. Existiert dann ein* $g \in L^2(\Omega)$ *mit der Eigenschaft*

$$\int_\Omega u\varphi_{x_k}\, dx = -\int_\Omega g\varphi\, dx \qquad \forall \varphi \in C_0^\infty(\Omega) \quad ,$$

so heißt $u_{x_k} := g$ *<u>schwache Ableitung</u> von* u *in Richtung* x_k.

Bemerkung: Die schwache Ableitung genügt *per definitionem* der Produktregel (Greensche Formel!).

Beispiel:

Wir betrachten die stückweise linearen Spline–Funktionen

$$S = \{v \in C[0,1] \mid v_{[x_{i-1},x_i]} \in \Pi_1,\ i = 1,\ldots,N\}$$

(Π_1 steht für die Polynome 1. Grades) bezüglich des Gitters

$$0 = x_0 < x_1 < x_2 < \cdots < x_N = 1 \quad .$$

Offenbar ist $v \in S$ i.a. nicht im klassischen, wohl aber im schwachen Sinne differenzierbar, denn mit $\varphi \in C_0^\infty(0,1)$ gilt

$$\begin{aligned}
\int_0^1 v\varphi'\,dx &= \sum_{i=1}^N \int_{x_{i-1}}^{x_i} v\varphi'\,dx \\
&= \sum_{i=1}^N \left(-\int_{x_{i-1}}^{x_i} v'\varphi\,dx \right) + \sum_{i=1}^{N-1} \underbrace{\left(v(x_i)^{\text{links}} - v(x_i)^{\text{rechts}} \right)}_{=0} \varphi(x_i) \\
&= -\int_0^1 g\varphi\,dx \quad ,
\end{aligned}$$

mit $g(x) := v'(x)\ \forall x \neq x_i$.

Beispiel:

Sei $v : [-1,1] \to \mathbb{R}$ gegeben durch

$$H(x) = \begin{cases} 0 & -1 \le x < 0 \\ 1 & 0 \le x \le 1 \end{cases} \qquad \text{(Heavyside–Funktion)} \quad .$$

Um zu prüfen, ob H schwach differenzierbar ist, wählen wir $\varphi \in C_0^\infty(-1,1)$ und berechnen

$$\int_{-1}^1 H\varphi'\,dx = 0 + \int_0^1 1 \cdot \varphi'\,dx = \varphi(1) - \varphi(0) = -\varphi(0) =: -\delta_0\varphi \quad .$$

Offenbar ist durch $\delta_0\varphi = \varphi(0)$ eine lineare Abbildung von $C_0^\infty(-1,1)$ nach $\mathbb{R}$ definiert. Diese Abbildung heißt *Dirac'sche Delta–Distribution.* δ_0 ist als Abbildung von $L^2(-1,1)$ nach $\mathbb{R}$ nicht stetig und daher existiert auch kein $g \in L^2(-1,1)$ mit der Eigenschaft $\int_{-1}^1 g\varphi\,dx = \delta_0\varphi$. Die Heavyside–Funktion H ist daher *nicht* schwach differenzierbar.

Wir erinnern an die Multiindex–Schreibweise $\partial^\beta = \frac{\partial}{\partial x_{\beta_1} \ldots \partial x_{\beta_k}}$, $\beta_i = 1, \ldots, n$, $|\beta| = k$.

Definition 5.21 (Sobolev–Raum) *Die Vervollständigung von*

$$X = \{v \in C^\infty(\Omega) \mid (v,v)_{H^m(\Omega)} < \infty\}$$

bezüglich der vom Skalarprodukt

$$(v,w)_{H^m(\Omega)} = \sum_{|\beta| \le m} (\partial^\beta v, \partial^\beta w)$$

erzeugten Norm ist der *Sobolev–Raum* $H^m(\Omega)$.

Satz 5.22 *Der Sobolev–Raum $H^m(\Omega)$ besteht aus allen (Äquivalenzklassen von) Funktionen $v \in L^2(\Omega)$ mit schwachen Ableitungen*

$$\partial^\beta v \in L^2(\Omega), \quad |\beta| \le m \quad .$$

$H^m(\Omega)$ ist ein Hilbertraum mit dem Skalarprodukt

$$(v,w)_m = \sum_{|\beta| \le m} (\partial^\beta v, \partial^\beta w) \quad .$$

Näheres erfährt man von Alt [2] auf den Seiten 31 ff. und 78.

Bemerkung: Nach Konstruktion ist $C^\infty(\Omega) \cap H^m(\Omega)$ dicht in $H^m(\Omega)$.

3. Randbedingungen im schwachen Sinn. Wir haben gesehen, daß Funktionen $u \in L^2(\Omega)$ nur bis auf (Lebesgue–)Nullmengen bestimmt sind. Der Rand $\partial\Omega$ ist eine Nullmenge. Welchen Sinn soll dann aber z.B. die Dirichlet–Randbedingung $v|_{\partial\Omega} = 0$ noch für Funktionen $v \in H_0^1(\Omega)$ haben, die nicht in $C\overline{\Omega}$ liegen? Eine erste Antwort, nämlich im Falle

$$\Omega = \mathbb{R}^n_+ = \{(x,y) \,|\, x \in \mathbb{R}^{n-1}, y > 0\}, \qquad n \in \mathbb{N} \ge 2$$

mit dem Rand $\partial\Omega = \mathbb{R}^{n-1}$, gibt der folgende *Spursatz*.

Satz 5.23 (Spursatz im Halbraum) *Es sei $\Omega = \mathbb{R}^n_+$. Dann gibt es eine beschränkte, lineare Abbildung*

$$tr:\ H^1(\Omega) \to L^2(\partial\Omega)$$

mit der Eigenschaft

$$tr\, v = v|_{\partial\Omega} \qquad \forall v \in H^1(\Omega) \cap C(\overline{\Omega}) \quad .$$

Beweis:
Wir werden später sehen, daß für genügend glatt berandete Gebiete der lineare Raum

$$X = \{u|_\Omega \mid u \in C_0^\infty(\mathbb{R}^n)\}$$

dicht in $H^1(\Omega)$ ist. Der Rand $\partial\Omega = \mathbb{R}^{n-1}$ von $\Omega = \mathbb{R}^n_+$ ist äußerst glatt. Wir setzen

$$tr\, v = v|_{\partial\Omega} \qquad \forall v \in X$$

und wollen nun zeigen, daß für ein geeignetes $c > 0$ die Abschätzung

$$\|tr\, v\|_{L^2(\partial\Omega)} \leq c\|v\|_{H^1(\Omega)} \qquad \forall v \in X$$

besteht. Dazu sei $v \in X$ beliebig aber fest gewählt.

1. Lokalisierung
   Wir konstruieren zunächst eine *Abschneidefunktion* $\varphi \in X$ mit der Eigenschaft

$$\varphi(x,y) = \begin{cases} 1 & 0 \leq y \leq \frac{1}{2} \\ 0 & 1 \leq y \end{cases} \qquad \forall (x,y) \in \overline{\mathbb{R}^n_+} \quad .$$

   Dann gilt für $w := \varphi v \in C^\infty(\overline{\mathbb{R}^n_+})$

$$w(x,0) = w(x,y) - \int_0^y w_y(x,s)\, ds \quad .$$

   Integration über y von 0 bis 1 liefert

$$w(x,0) = \int_0^1 w(x,y)\, dy - \int_0^1 \int_0^y w_y(x,s)\, ds\, dy \quad .$$

   Die Produktregel ergibt

$$w(x,0) = \int_0^1 w(x,y)\, dy - \int_0^1 (1-y) w_y(x,y)\, dy \quad .$$

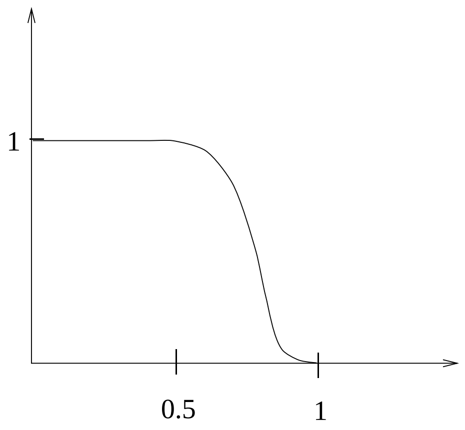


Abbildung 10: Abschneidefunktion φ

Youngsche Ungleichung ($(a+b)^2 \leq 2a^2 + 2b^2$) und Cauchy–Schwarzsche Ungleichung liefern

$$\begin{aligned} w^2(x,0) &\leq 2\int_0^1 1^2\,dy \cdot \int_0^1 w^2(x,y)\,dy + 2\int_0^1 (1-y)^2\,dy \int_0^1 w_y^2(x,y)\,dy \\ &= 2\int_{\mathbb{R}_+} w^2(x,y)\,dy + \frac{2}{3}\int_{\mathbb{R}_+} w_y^2(x,y)\,dy \quad . \end{aligned}$$

Integration über x liefert

$$\int_{\mathbb{R}^{n-1}} w^2(x,0)\,dx \leq 2\int_{\mathbb{R}^n_+} w^2(x,y)\,d(x,y) + 2\int_{\mathbb{R}^n_+} w_y^2(x,y)\,d(x,y) \quad ,$$

also wegen $tr\,w = w|_{\partial\mathbb{R}^n_+}$

$$\|tr\,w\|^2_{L^2(\partial\mathbb{R}^n_+)} \leq 2(\|w\|^2_{L^2(\mathbb{R}^n_+)} + \|w_y\|^2_{L^2(\mathbb{R}^n_+)}) \quad .$$

2. Stabilitätsabschätzung für $v \in X$

   Einsetzen von $w = v\varphi$ liefert wegen $\varphi_{|\partial\mathbb{R}^n_+} \equiv 1$

$$tr\,w = tr\,v \quad .$$

   Weiter gilt

$$\int_{\mathbb{R}^n_+} w^2\,d(x,y) = \int_{\mathbb{R}^n_+} \varphi^2 v^2\,d(x,y) \leq \max_{(x,y)\in\mathbb{R}^n_+}\{\varphi(x,y)\} \int_{\mathbb{R}^n_+} v^2\,d(x,y) \leq \|v\|^2_{L^2(\mathbb{R}^n_+)}$$

und

$$\begin{aligned}
\int\limits_{\mathbb{R}^n_+} w_y^2 \, d(x,y) &= \int\limits_{\mathbb{R}^n_+} (\varphi_y v + v_y \varphi)^2 \, d(x,y) \\
&\leq 2 \int\limits_{\mathbb{R}^n_+} \varphi_y^2 v^2 + v_y^2 \varphi^2 \, d(x,y) \\
&\leq 2(\max_{(x,y)\in\mathbb{R}^n_+} \{\varphi_y^2(x,y)\} \, \|v\|^2_{L^2(\mathbb{R}^n_+)} + \max_{x\in\mathbb{R}^n_+} \{\varphi^2(x)\} \, \|v_y\|^2_{L^2(\mathbb{R}^n_+)}) \\
&\leq c \, \|v\|^2_{H^1(\mathbb{R}^d_+)} \quad .
\end{aligned}$$

Insgesamt folgt

$$\|tr\, v\|^2_{L^2(\partial\Omega)} \leq c \, \|v\|^2_{H^1(\Omega)} \qquad \forall v \in X \quad . \tag{5.10}$$

3. Dichteschluß

   Es sei $v \in H^1(\Omega)$. Dann existiert eine Folge $(v_k)_{k\in\mathbb{N}} \in X$ mit

   $$\|v - v_k\|_{H^1(\Omega)} \to 0 \quad .$$

   Also ist $(v_k)_{k\in\mathbb{N}}$ eine Cauchy–Folge in $H^1(\Omega)$. Wegen

   $$\|tr\, v_k - tr\, v_j\|_{L^2(\partial\Omega)} \leq c \, \|tr\, v_k - tr\, v_j\|_{H^1(\Omega)}$$

   ist $tr\, v_k = v_{k|_{\partial\Omega}}$ eine Cauchy–Folge in $L^2(\partial\Omega)$. Sei $v^* \in L^2(\partial\Omega)$ ihr Grenzwert. Dann setzen wir $tr\, v := v^*$. Diese Definition ist unabhängig von der Wahl der Folge $(v_k)_{k\in\mathbb{N}}$ (nachprüfen). Alle $v_k \in X$ genügen der Abschätzung (5.9). Grenzübergang liefert mit der Stetigkeit der Normen

   $$\|tr\, v\|_{L^2(\partial\Omega)} \leq c \, \|v\|_{H^1(\Omega)} \qquad \forall v \in H^1(\Omega) \quad ,$$

   also die Stetigkeit von $tr : H^1(\Omega) \to L^2(\Omega)$.

   Um schließlich zu zeigen, daß

   $$tr\, v = v|_{\partial\Omega} \qquad \forall v \in H^1(\Omega) \cap C(\overline{\Omega})$$

   gilt, müssen wir noch zu jedem $v \in H^1(\Omega) \cap C(\overline{\Omega})$ eine Folge $(v_k)_{k\in\mathbb{N}} \in X$ konstruieren, die gegen v konvergiert und zusätzlich die Eigenschaft

   $$v_n|_{\partial\Omega} \to v|_{\partial\Omega} \qquad \text{in } L^2(\Omega)$$

   hat. Das technische Handwerkszeug dazu findet sich bei Alt [2] in den Abschnitten 2.9 und 2.12.

□

Der Spursatz im Halbraum lässt sich auf Gebiete mit *Lipschitz–Rand* erweitern.

Definition 5.24 (Lipschitz–Rand) *Ein beschränktes Gebiet $\Omega \subset \mathbb{R}^n$ hat einen Lipschitz–Rand falls es endlich viele offene Mengen O_i und eine Zahl $\varepsilon > 0$ gibt, so daß für alle $x \in \partial\Omega$ der Ball $\overline{K_\varepsilon(x)}$ in einem gewissen O_i liegt und $O_i \cap \Omega = O_i \cap \Omega_i$ gilt. Dabei ist*

$$\Omega_i = \{(x,y) \in \mathbb{R}^n \,|\, x \in \mathbb{R}^{n-1}, y \in \mathbb{R}, y < \varphi_i(x)\}$$

mit Lipschitz–stetigem φ_i, *also*

$$|\varphi_i(x_1) - \varphi_i(x_2)| \leq L|x_1 - x_2|$$

für eine feste (Lipschitz–)Konstante $L \in \mathbb{R}$.

Beispiel: Der Rand $\partial\Omega$ des geschlitzten Kreisgebiets in Abbildung 11 ist nicht Lipschitz. Beachte: In diesem Fall sind auch für $v \in C(\overline{\Omega})$ die Randwerte $v_{|\partial\Omega}$ nicht ohne weiteres definiert (doppelte Werte für $y = 0, x \in (0,1)$)!

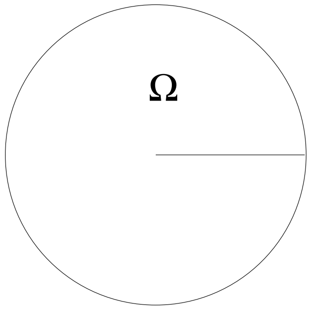


Abbildung 11: Geschlitztes Kreisgebiet

Satz 5.25 (Spursatz im Lipschitz–Gebiet) *Sei* $\Omega \subset \mathbb{R}^n$ *ein Gebiet mit Lipschitz–Rand. Dann existiert zu jedem* $v \in H^1(\Omega)$ *eine Folge* $(v_k)_{k\in\mathbb{N}}$,

$$(v_k)_{k\in\mathbb{N}} \subset \{w|_\Omega \mid w \in C_0^\infty(\mathbb{R}^n)\}$$

mit der Eigenschaft

$$v = \lim_{k\to\infty} v_k \qquad \text{in } H^1(\Omega) \quad .$$

Der Spuroperator $tr : H^1(\Omega) \to L^2(\partial\Omega)$, *ist durch*

$$tr\, v = \lim_{k\to\infty} v_k|_{\partial\Omega} \qquad \text{in } L^2(\partial\Omega) \quad ,$$

wohldefiniert und beschränkt. Weiter gilt

$$tr\, v = v|_{\partial\Omega} \qquad \forall v \in H^1(\Omega) \cap C(\overline{\Omega}) \quad .$$

Beweis:

Siehe Alt [2], Seite 190.

□

Bemerkung: Der Satz läßt sich verschärfen zu

$$\|tr\, v\|_{H^{\frac{1}{2}}(\partial\Omega)} \leq c\, \|v\|_{H^1(\Omega)} \quad .$$

Der Sobolev–Raum (rationaler Ordnung!) $H^{\frac{1}{2}}(\partial\Omega)$ ist genau der richtige Spurraum, denn der Spuroperator $tr : H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)$ ist surjektiv, d.h. zu jedem $g \in H^{\frac{1}{2}}(\partial\Omega)$ gibt es mindestens ein $v \in H^1(\Omega)$ mit der Eigenschaft $tr\, v = g$.

Wir kommen nun kurz auf $H_0^1(\Omega)$ zurück.

Satz 5.26 *Der Teilraum*

$$C_0^\infty(\Omega) \subset H_0^1(\Omega)$$

liegt dicht in $H_0^1(\Omega)$.

Wir hätten also $H_0^1(\Omega)$ auch direkt als Vervollständigung von $C_0^\infty(\Omega)$ bezüglich $\|\cdot\|_{H^1(\Omega)}$ definieren können. Üblicherweise wird das auch so gemacht.

Bemerkung: Offenbar gilt nach Satz 5.25

$$tr\ v = 0 \qquad \forall v \in H_0^1(\Omega)$$

im Sinne des Spuroperators. Gleichbedeutend ist

$$H_0^1(\Omega) = \{v \in H^1(\Omega) \mid tr\, v = 0\ \} \quad .$$

So sind also die Nullrandbedingungen zu verstehen.

Wir haben gesehen, daß Funktionen $v \in H^k(\Omega)$ zumindest für genügend große k ein ähnliches Verhalten zeigen wie stetige Funktionen (Randwerte). Das führt auf die Frage, wie groß $k > m$ sein muß, damit

$$v \in H^k(\Omega) \implies v \in C^m(\Omega)$$

gilt. Antwort gibt der folgende *Einbettungssatz*

Satz 5.27 (Sobolevscher Einbettungssatz) *Sei* $\Omega \subset \mathbb{R}^n$ *ein beschränktes Gebiet mit Lipschitz–Rand. Ist dann* $v \in H^k(\Omega)$ *mit*

$$k > m + \frac{n}{2} \quad , \tag{5.11}$$

so folgt

$$v \in C^m(\Omega) \qquad (\textit{genauer, es existiert}\ v \in [v] \cap C^m(\Omega))$$

und

$$\sup_{x\in\Omega} |\partial^\beta v\,| \leq c\, \|v\|_{H^k(\Omega)} \quad \forall |\beta| \leq m \quad .$$

Beweis:

Siehe Alt [2], Seite 244.

□

Bemerkung: Es gilt somit $H^1(\Omega) \subset C(\overline{\Omega})$ falls $n = 1$. Da wir für $n = 2$ leider nur Gleichheit in (5.10) haben, ist das für $n \geq 2$ nicht mehr der Fall. Es gibt Gegenbeispiele (Übung).

Satz 5.28 (Auswahlsatz von Rellich) *Sei $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet mit Lipschitz–Rand. Ist dann $(v_k)_{k\in\mathbb{N}} \subset H^{m+1}(\Omega)$ beschränkt, also*

$$\|v_k\|_{H^{m+1}} \leq c \qquad \forall k \in \mathbb{N} \quad ,$$

so existiert eine Teilfolge $\left(v_{k_j}\right)_{k_j\in\mathbb{N}}$ und ein $v \in H^{m+1}(\Omega)$, so daß

$$v_{k_j} \to v\,, \quad \text{für} \quad j \to \infty \quad \text{in} \quad H^m(\Omega) \quad .$$

Beweis:

Für Funktionalanalytiker: $H^{m+1}(\Omega)$ ist reflexiv, also konvergiert eine Teilfolge $\left(v_{k_j}\right)_{k_j\in\mathbb{N}}$ schwach in $H^{m+1}(\Omega)$ gegen ein $v \in H^{m+1}(\Omega)$. Der Rest steht bei Alt [2], Seite 184.

□

5.3 Schwache Lösungen elliptischer Randwertprobleme

Als Lösungsraum für unser Minimierungsproblem (5.2), bzw. für dessen Variationsformulierung

$$u \in H \; : \quad a(u,v) = l(v) \qquad \forall v \in H$$

haben wir nun $H = H_0^1(\Omega)$ identifiziert. Um den Rieszschen Darstellungssatz anwenden zu können, müssen wir sicherstellen, daß $a(\cdot,\cdot)$ und l die entsprechenden Voraussetzungen erfüllen. Vorher vereinbaren wir die Abkürzungen

$$\|v\|_{m,\Omega} = \|v\|_{H^m(\Omega)} \qquad \|v\|_{0,\Omega} = \|v\|_{L^2(\Omega)} \quad .$$

Wenn klar ist, daß sich die Normen auf Ω beziehen, lassen wir das Rechengebiet im Index manchmal weg. Außerdem schreiben wir

$$(v,w) = (v,w)_{L^2(\Omega)} \qquad (\nabla v, \nabla w) = \sum_{k=1}^{n} (v_{x_k}, w_{x_k}) \quad .$$

Lemma 5.29 *Es sei $f \in L^2(\Omega)$. Dann ist*

$$l(v) = (f,v) \tag{5.12}$$

ein beschränktes Funktional auf $H = H_0^1(\Omega)$ und es gilt $\|l\|_{H'} \leq \|f\|_0$.

Beweis:
Die Schwarzsche Ungleichung liefert

$$|l(v)| \le \|f\|_0 \, \|v\|_0 \le \|f\|_0 \, \|v\|_1$$

und damit die Behauptung.

□

Lemma 5.30 (Poincaré-Friedrichs'sche Ungleichung) *Es sei Ω beschränkt und $0 < \alpha_0 \le \alpha(x) \le \alpha_1 < \infty$ fast überall in Ω. Dann gibt es ein $C > 0$, so daß die Bilinearform*

$$a(u,v) = \int\limits_\Omega \alpha \, \nabla u \cdot \nabla v \, dx \tag{5.13}$$

die Poincaré-Friedrichs'sche Ungleichung

$$C \, \|v\|_0^2 \le a(v,v) \qquad \forall v \in H_0^1(\Omega) \quad .$$

erfüllt.

Beweis:
Wegen $\alpha(x) \ge \alpha_0 > 0$, fast überall, gilt

$$a(v,v) \ge \alpha_0 \int\limits_\Omega |\nabla v|^2 \, dx \quad .$$

Es reicht also, ein $c > 0$ zu finden, so daß

$$\|v\|_0^2 \le c \| \mid \nabla v \mid \|_0^2 \tag{5.14}$$

gilt. Wir wählen nun ein offenes Quadrat $Q = (a,b) \times (a,b)$ mit $\overline{\Omega} \subset Q$ und setzen ein beliebiges aber festes $v \in C_0^\infty(\Omega)$ mit Null zu $v \in C_0^\infty(Q)$ fort. Dann ist

$$\begin{aligned}
v(x_1,x_2) &= v(a,x_2) + \int\limits_a^{x_1} v_{x_1}(\xi,x_2) \, d\xi = \int\limits_a^{x_1} v_{x_1}(\xi,x_2) \, d\xi \\
&\overset{\text{Cauchy–Schwarz}}{\le} \left(\int\limits_a^{x_1} 1^2 \, dx \right)^{\frac{1}{2}} \cdot \left(\int\limits_a^{x_1} v_{x_1}^2(\xi,x_2) \, d\xi \right)^{\frac{1}{2}} \\
&\le (b-a)^{\frac{1}{2}} \left(\int\limits_a^{b} v_{x_1}^2(\xi,x_2) \, d\xi \right)^{\frac{1}{2}} \quad .
\end{aligned}$$

Integration liefert

$$\begin{aligned}
\|v\|_0^2 &= \int\limits_a^b \int\limits_a^b v^2(x_1,x_2) \, dx_1 \, dx_2 \\
&\le (b-a)^2 \int\limits_a^b \int\limits_a^b v_{x_1}^2(\xi,x_2) \, d\xi \, dx_2 \\
&\le (b-a)^2 \, \| \mid \nabla v \mid \|_0^2 \quad .
\end{aligned}$$

□

Als Korollar des Lemmas von Lax–Milgram erhalten wir nun das

Theorem 5.31 *Es sei* $H = H_0^1(\Omega)$ *und* $a(u,v) = \int_\Omega \alpha \nabla u \cdot \nabla v \, dx$ *bzw.* $l = (f,v)$ *wie oben gegeben durch* (5.12) *bzw.* (5.11). *Dabei gelte* $0 < \alpha_0 \leq \alpha(x) \leq \alpha_1$, *fast überall, und* $f \in L^2(\Omega)$. *Dann ist die Variationsaufgabe*

$$u \in H : \quad a(u,v) = l(v) \qquad \forall v \in H \tag{5.15}$$

eindeutig lösbar und es gilt (mit $\gamma = (C^{-1} + \alpha_0^{-1})^{-1}$ *und* C *aus Lemma 5.30)*

$$\|u\|_1 \leq \frac{1}{\gamma} \|f\|_0 \quad .$$

Beweis:

Nach Lemma 5.29 ist $l \in H'$. Weiter ist $a(\cdot,\cdot)$ trivialerweise symmetrisch und bilinear. Nach Lemma 5.30 und wegen

$$\begin{aligned}
|a(v,w)| &\leq \alpha_1 \int_\Omega |\nabla v \cdot \nabla w| \, dx \\
&\leq \alpha_1 \int_\Omega |v_{x_1}|\,|w_{x_1}| + |v_{x_2}|\,|w_{x_2}| \, dx \\
&\overset{\text{Cauchy-Schwarz in } \mathbb{R}^2}{\leq} \alpha_1 \int_\Omega (|v_{x_1}|^2 + |v_{x_1}|^2)^{\frac{1}{2}} \cdot (|w_{x_2}|^2 + |w_{x_2}|^2)^{\frac{1}{2}} \, dx \\
&\overset{\text{Cauchy-Schwarz in } L^2(\Omega)}{\leq} \alpha_1 \left(\int_\Omega (|v_{x_1}|^2 + |v_{x_1}|^2) \, dx \right)^{\frac{1}{2}} \cdot \left(\int_\Omega (|w_{x_2}|^2 + |w_{x_2}|^2) \, dx \right)^{\frac{1}{2}} \\
&\leq \alpha_1 \, \|v\|_1 \cdot \|w\|_1
\end{aligned}$$

ist $a(\cdot,\cdot)$ H-elliptisch.

□

Bemerkung: Nach Theorem 5.31 ist das Variationsproblem (5.14) korrekt gestellt: Der Lösungsoperator $f \to u$ als Abbildung von $L^2(\Omega)$ nach $H_0^1(\Omega)$ ist stetig.

Bemerkung: Ist die Lösung von (5.14) eine glatte Funktion, nämlich $u \in C^2(\overline{\Omega}) \cap H_0^1(\Omega)$, so folgt mit der Greenschen Formel, daß u auch klassische Lösung des Randwertproblems

$$\begin{aligned}
-\operatorname{div}(\alpha \nabla u) &= f \qquad \text{in } \Omega \\
u &= 0 \qquad \text{auf } \partial\Omega \quad .
\end{aligned} \tag{5.16}$$

ist.

Definition 5.32 *Die Lösung* u *des Variationsproblems* (5.14) *heißt* <u>*schwache Lösung*</u> *des Randwertproblems* (5.16).

Wir wollen nun die schwache Formulierung einiger anderer Randwertprobleme herleiten. Die Bilinearform $a(\cdot,\cdot)$ ist dabei jeweils so zu wählen, daß für genügend glatte Lösungen die Greensche Formel wieder auf das ursprüngliche Randwertproblem führt.
Ist $a(\cdot,\cdot)$ symmetrisch und positiv definit auf einem (bzgl. der Energienorm!) abgeschlossenen Lösungsraum $H \subset H^1(\Omega)$, so existiert eine eindeutig bestimmte schwache Lösung (Riesz). Ist $a(\cdot,\cdot)$ H-elliptisch im Sinne von (5.7), so brauchen wir die Symmetrie nicht und es reicht aus, die Abgeschlossenheit von H bezüglich $\|\cdot\|_1$ zu kennen, und es folgt zusätzlich die $H^1(\Omega)$–Stabilität (Lax–Milgram).

Beispiel:

Wir betrachten die schwache Form des Randwertproblems (5.16) mit

$$\alpha(x) = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}(x) \quad .$$

Für jedes $x \in \Omega$ sei $\alpha(x)$ eine symmetrische Matrix. Für die Eigenwerte $\lambda_1(x)$, $\lambda_2(x)$ gelte

$$0 < \alpha_0 \le \lambda_1(x), \lambda_2(x) \le \alpha_1 \text{ fast überall} \quad .$$

Dann ist die Bilinearform

$$a(v,w) = \int_\Omega (\alpha \nabla v) \cdot \nabla w \, dx$$

$H_0^1(\Omega)$–elliptisch.

Beispiel:

Wir betrachten das Randwertproblem

$$\begin{aligned} -\Delta u + \beta \cdot \nabla u &= f && \text{in } \Omega \\ u &= 0 && \text{auf } \partial\Omega \quad . \end{aligned} \tag{5.17}$$

Die zugehörige schwache Formulierung führt auf die Bilinearform

$$a(v,w) = \underbrace{(\nabla v, \nabla w)}_{\text{symmetrisch}} + \underbrace{(\beta \cdot \nabla v, w)}_{\text{nicht symmetrisch}} \quad .$$

Diese Bilinearform ist nicht symmetrisch, aber $H_0^1(\Omega)$-elliptisch (Übung).

Beispiel:

Wir betrachten das Randwertproblem

$$\begin{array}{rcll} -\Delta u + \lambda u & = & f & \text{in } \Omega \\ \frac{\partial}{\partial n} u & = & 0 & \text{auf } \Omega \end{array} \tag{5.18}$$

mit $\lambda > 0$. Die zugehörige schwache Formulierung (5.14) führt auf die Bilinearform

$$a(v, w) = (\nabla v, \nabla w) + \lambda(v, w) \quad .$$

und den Lösungsraum $H = H^1(\Omega)$. Offenbar ist $a(v, w)$ $H^1(\Omega)$-elliptisch für $\lambda > 0$.
Im Gegensatz zu Dirichlet–Randbedingungen sind Neumann–Randbedingungen nicht durch die Wahl eines Teilraums von $H^1(\Omega)$ erzwungen. Das kann auch nicht sein, denn $\nabla u \in (L^2(\Omega))^2$ hat keine Randwerte! Genügend glatte Lösungen des Variationsproblems erfüllen aufgrund der Greenschen Formel automatisch homogene Neumann–Bedingungen (Übung)!
Wir betrachten noch kurz den Fall $\lambda = 0$. $a(v, w) = (\nabla v, \nabla w)$ ist nicht positiv definit auf $H^1(\Omega)$, denn

$$(\nabla v, \nabla w) = 0 \iff v = \text{const.} \quad .$$

Wir identifizieren daher Funktionen, die sich nur um eine Konstante unterscheiden,

$$[v] = \{w \mid w(x) - v(x) = \text{const., fast überall in } \Omega\} \quad .$$

Auf dem Quotientenraum H der so gebildeten Äquivalenzklassen $[v]$ ist durch

$$a([v], [w]) = (\nabla v, \nabla w), \qquad v \in [v],\ w \in [w]$$

ein Skalarprodukt definiert. Durch $f \in L^2(\Omega)$ ist genau dann ein lineares Funktional

$$l([v]) = \int_\Omega f v \, dx, \qquad v \in [v]$$

erklärt, wenn die (bereits bekannte!) Kompatibiliätsbedingung

$$\int_\Omega f \, dx = 0$$

erfüllt ist. In diesem Fall hat das Variationsproblem

$$[u] \in H : \qquad a([u], [v]) = l([v]) \qquad \forall v \in H$$

eine eindeutig bestimmte Lösung $[u]$. Alle Funktionen $u \in [u]$ sind Lösungen unseres Ausgangsproblems mit $\lambda = 0$.

Beispiel:

Wir wollen nun inhomogene Randbedingungen betrachten. Dazu sei $\partial_D \cup \partial_N = \partial\Omega$ eine nichtüberlappende Zerlegung des Randes in zwei Teilstücke nichtverschwindender Länge. Vorgelegt sei das Randwertproblem

$$\begin{array}{rcll} -\Delta u & = & f & \text{in } \Omega \\ u & = & g_D & \text{auf } \partial_D \\ \frac{\partial u}{\partial n} & = & g_N & \text{auf } \partial_N \quad . \end{array} \tag{5.19}$$

Eine schwache Formulierung (5.14) erhält man mit

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

und
$$l(v) = \int_\Omega fv \, dx + \int_{\partial_N} g_N v \, d\sigma \qquad .$$

Als Lösungsmenge bietet sich an
$$X = \{v \in H^1(\Omega) \,|\, tr_{\partial_D} \, v = g_D\} \quad .$$

Problem: X ist kein linearer Raum! Wir nehmen nun an, daß zumindest $X \neq \emptyset$, daß also ein $w_0 \in H^1(\Omega)$ mit der Eigenschaft
$$tr_{\partial_D} \, w_0 = g_D$$
existiert (Dies ist genau dann der Fall, wenn $g_D \in H^{\frac{1}{2}}(\partial_D)$). Dann berechnen wir u_0 aus
$$u_0 \in H : \qquad a(u_0, v) = l(v) - a(w_0, v) \qquad \forall v \in H$$
mit dem Hilbertraum
$$H = \{v \in H^1(\Omega) \mid tr_{\partial_D} \, v = 0\}$$
und erhalten die (eindeutig bestimmte!) Lösung $u = u_0 + w_0$.
Wir haben gesehen, daß g_D gewisse Bedingungen erfüllen muß . Welche Bedingungen an g_N sind für die Lösbarkeit erforderlich?

6 Finite Elemente

6.1 Konstruktion von FE-Räumen

Wir betrachten die Variationsgleichung

$$v \in H: \; a(v,w) = l(v) \qquad \forall v \in H \tag{6.1}$$

mit einem abgeschlossenen Teilraum $H \subset H^1(\Omega)$, einer H–elliptischen Bilinearform $a(\cdot,\cdot)$, also

$$\gamma \|v\|_1^2 \leq a(v,v), \qquad |a(v,w)| \leq \Gamma \|v\|_1 \|w\|_1, \qquad v,w \in H$$

und $l \in H'$.
Aus dem Céa–Lemma folgt sofort

Theorem 6.1 *Sei $S \subset H$ ein abgeschlossener Unterraum. Dann ist u_S durch*

$$u_S \in S: \quad a(u_S, v) = l(v) \qquad \forall v \in S$$

eindeutig bestimmt, und es gilt

$$\|u - u_S\|_1 \leq \frac{\Gamma}{\gamma} \inf_{v \in S} \|u - v\|_1 \quad .$$

Bemerkung: Diese Klasse von Näherungsverfahren heißt *Galerkin–Verfahren.* Ist $a(\cdot,\cdot)$ symmetrisch, spricht man von *Ritz–Galerkin–Verfahren.*

Als Ansatzraum S wollen wir finite Elemente wählen.

Definition 6.2 *Es sei $\Omega \subset \mathbb{R}^2$ polygonal berandet. Eine Menge $\mathcal{T}$ von Dreiecken t heißt dann Triangulierung von Ω, falls gilt*

(i) $\overline{\Omega} = \bigcup_{t \in \mathcal{T}} t$

(ii) Der Schnitt zweier Dreiecke aus $\mathcal{T}$ ist entweder eine gemeinsame Kante, ein gemeinsamer Eckpunkt oder leer.

Die Menge aller inneren Kanten heißt $\mathcal{E}_\mathcal{T}$. Die Menge aller inneren Eckpunkte heißt $\mathcal{N}_\mathcal{T}$.

Definition 6.3 *Es bezeichne Π_m die Menge aller Polynome m–ter Ordnung. Dann ist*

$$\boxed{S^{(m)} := \{v \in C(\overline{\Omega}) \mid v|_t \in \Pi_m \quad \forall t \in \mathcal{T}\}.}$$

der Raum der (polynomialen) finiten Elemente m–ter Ordnung (bezüglich der Triangulierung $\mathcal{T}$).

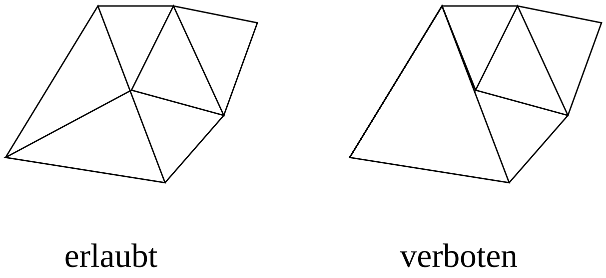


Abbildung 12: Triangulierung und verbotene Dreieckszerlegung

Bemerkung: Abgesehen von trivialen Fällen ist $S^{(m)} \not\subset C^1(\Omega)$. Im Falle $m = 1, 2, 3, \ldots$, spricht man von (stückweise) linearen, quadratischen, kubischen, ... finiten Elementen.

Um eine Basis von $S^{(m)}$ zu konstruieren, wollen wir zunächst klären, wie sich ein Polynom $P \in \Pi_m$ festlegen lässt.

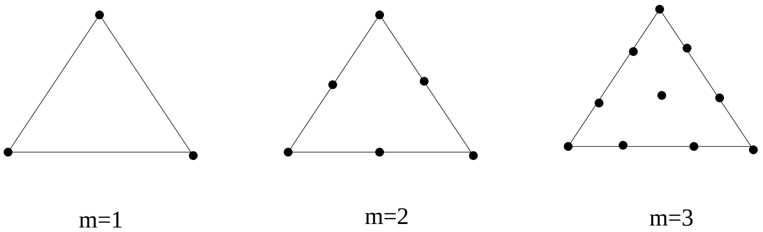


Abbildung 13: Knoten

Lemma 6.4 *Sei $m \geq 0$. In einem Dreieck t seien auf $m+1$ parallelen Linien insgesamt*

$$s = 1 + 2 + \cdots + (m+1) = \frac{(m+1)(m+2)}{2}$$

Knoten $\mathcal{N}^{(m)}(t) = \{z_1, z_2, \ldots, z_s\}$ angeordnet, wie in Abbildung 13 dargestellt. Dann hat die Interpolationsaufgabe

$$P \in \Pi_m : \qquad P(z_i) = P_i \qquad \forall i = 1, \ldots, s$$

für alle Knotenwerte P_i eine eindeutig bestimmte Lösung.

Beweis:

Vollständige Induktion:
Der Fall $m = 0$ ist klar. Sei die Behauptung für $m-1 \geq 0$ erfüllt. Ohne Beschränkung der Allgemeinheit liege die Kante mit den Knoten $z_1, z_2, \ldots, z_{m+1}$ auf der x–Achse. Die darüberliegenden Linien sind dann $\{(x, y) \mid y = y_i\}$, mit positiven $y_k \in \mathbb{R}$, $k = 1, \ldots, m$. Die univariate Interpolationsaufgabe

$$P_0 \in \Pi_m : \qquad P_0(z_i) = P_i \qquad \forall i = 1, \ldots, m+1$$

hat bekanntlich für alle P_i eine eindeutig bestimmte Lösung, die wir, z.B. konstant in y-Richtung, auf ganz $\mathbb{R}^2$ fortsetzen können. Nach Induktionsvoraussetzung gibt es ferner genau ein $Q_0 \in \Pi_{m-1}$ mit der Eigenschaft

$$Q_0(z_i) = \frac{1}{y_{k_i}}(P_i - P_0(z_i)) \qquad \forall i = m+2, \ldots, s \quad .$$

Dabei bezeichnet k_i jeweils die Linie, auf der der Knoten z_i liegt. Dann ist $P(x,y) = P_0(x,y) + yQ_0(x,y)$ eindeutig bestimmte Lösung.

□

Wir betrachten nun wieder die gesamte Triangulierung $\mathcal{T}$.

Korollar 6.5 *Es sei*

$$\mathcal{N}^{(m)} = \bigcup_{t \in \mathcal{T}_h} \mathcal{N}^{(m)}(t).$$

Dann hat die Interpolationsaufgabe

$$v \in S^{(m)} : \qquad v(p) = v_p \qquad \forall p \in \mathcal{N}^{(m)} \tag{6.2}$$

für alle Knotenwerten v_p eine eindeutig bestimmte Lösung.

Beweis:

Für jedes Dreieck $t \in \mathcal{T}$ ist $P_t \in \Pi_m$ durch

$$P_t \in \Pi_m : \qquad P_t(p) = v_p \qquad \forall p \in t \cap \mathcal{N}^{(m)}$$

eindeutig bestimmt. Da alle Eckpunkte Knoten sind, haben die zu angrenzenden Dreiecken gehörigen Polynome dort den gleichen Wert. Falls zwei Dreiecke t_1, t_2 die gemeinsame Kante $e = t_1 \cap t_2$ haben, gilt $P_{t_1}(x) = P_{t_2}(x)$ $\forall x \in e = t_1 \cap t_2$, da auf jeder Kante $m+1$ Knoten liegen. Damit ist durch

$$v(x) = P_t(x) \qquad \text{falls } x \in t$$

eine Funktion $v \in S^{(m)}$ gegeben, die offenbar die Interpolationsbedingungen erfüllt, also Lösung von (6.1) ist. Die Eindeutigkeit folgt durch Widerspruchsbeweis.

□

Im Fall $m = 1$ ist

$$\mathcal{N}^{(1)} = \{p \mid p \in \Omega \text{ ist Eckpunkt eines Dreiecks } t \in \mathcal{T}\} \quad .$$

Definition 6.6 *Durch Lösung der Interpolationsaufgaben*

$$\boxed{\lambda_p \in S^{(m)} : \qquad \lambda_p(q) = \delta_{pq} \qquad \forall q \in \mathcal{N}^{(m)} \qquad \textit{(Kronecker–}\delta\textit{)}}$$

für jedes $p \in \mathcal{N}^{(m)}$ erhält man die Knotenbasis

$$\Lambda^{(m)} = \{\lambda_p \mid p \in \mathcal{N}^{(m)}\}.$$

von $S^{(m)}$. Ist p Eckpunkt von t so heißt $\lambda_p|_t \in \Pi_m$ Formfunktion.

Bemerkung: Um diese Bezeichnung zu rechtfertigen, müssen wir nachprüfen, ob $\Lambda^{(m)}$ überhaupt eine Basis ist. Offenbar hat jedes $v \in S^{(m)}$ die Darstellung

$$v = \sum_{p \in \mathcal{N}^{(m)}} v(p)\lambda_p \quad .$$

Die Eindeutigkeit folgt aus Korollar 6.5. Im linearen Fall (m=1) sind Knotenbasisfunktionen in Abbildung 14 illustriert.

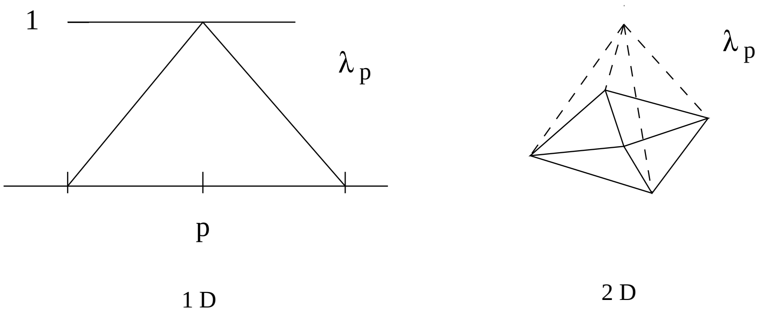


Abbildung 14: Knotenbasisfunktionen für $m = 1$

Wir betrachten nun unsere Variationsgleichung (6.0) im Falle $H = H_0^1(\Omega)$, einer $H_0^1(\Omega)$–elliptischen Bilinearform $a(\cdot,\cdot)$ und $l \in H'$. Zum Beispiel erfüllen bekanntlich

$$a(v,w) = (\alpha\nabla v, \nabla w) = \int_\Omega \alpha\nabla v \cdot \nabla w \; dx$$

mit $0 < \alpha_0 \leq \alpha(x) \leq \alpha_1$ f.ü. auf Ω, und

$$l(v) = (f,v) = \int_\Omega f \; v \; dx$$

mit $f \in L^2(\Omega)$ diese Voraussetzungen.
Zur Approximation der Lösung u unseres Modellproblems (6.0) wählen wir nun

$$\boxed{S_h := \{v \in S^{(m)} \mid v|_{\partial\Omega} = 0\}} \tag{6.3}$$

Der Diskretisierungsparameter h,

$$\boxed{h = \max\nolimits_{t \in \mathcal{T}_h} \operatorname{diam} t,}$$

beschreibt die Feinheit des Gitters. Oft nennt man h auch *Schrittweite*.

Satz 6.7 *S_h ist ein abgeschlossener Teilraum von $H_0^1(\Omega)$.*

Beweis:

Übung.

□

Anwendung des Ritz–Galerkin–Verfahrens entsprechend Theorem 6.1 liefert die Variationsgleichung

$$u_h \in S_h : \qquad a(u_h, v) = l(v) \qquad \forall v \in S_h \quad . \tag{6.4}$$

Bezeichnet

$$\mathcal{N}_h = \mathcal{N}^{(m)} \cap \Omega$$

die Menge der *inneren* Knoten von $\mathcal{T}$, so ist offenbar

$$\Lambda_h = \Lambda_h^{(m)} \cap H_0^1(\Omega) = \{\lambda_p \mid p \in \mathcal{N}_h\}$$

die Knotenbasis von S_h, denn die Werte in den Randknoten sind durch die Randbedingungen festgelegt. Einsetzen der Knotenbasisdarstellung

$$\boxed{u_h = \textstyle\sum_{p \in \mathcal{N}_h} u_p \lambda_p}$$

und von $v = \lambda_q$, $q \in \mathcal{N}$, liefert das lineare Gleichungssystem

$$\sum_{p \in \mathcal{N}_h} a(\lambda_p, \lambda_q) u_p = l(\lambda_q) \qquad \forall q \in \mathcal{N}_h$$

für die unbekannten Koeffizienten $u_p = u(p)$. Mit den Definitionen

$$\boxed{\begin{array}{rclcrcl} A &=& (a_{p,q})_{p,q \in \mathcal{N}_h}, & \qquad & a_{pq} &=& a(\lambda_q, \lambda_p) \\ b &=& (b_p)_{p \in \mathcal{N}_h} & & b_p &=& l(\lambda_p) \\ U &=& (u_p)_{p \in \mathcal{N}_h} & & & & \end{array}}$$

lässt sich das Variationsproblem (6.3) als äquivalentes Gleichungssystem

$$\boxed{AU = b} \tag{6.5}$$

schreiben. Wir notieren hier nur ein paar grundlegende Eigenschaften von A.

Satz 6.8 *Ist die Bilinearform $a(\cdot,\cdot)$ elliptisch, so ist A positiv definit. Ist die Bilinearform $a(\cdot,\cdot)$ symmetrisch, so ist auch A symmetrisch.*

Beweis:

Übung

□

Wir kommen später auf die Lösung des (großen!) Gleichungssystems (6.4) zurück und kümmern uns zunächst um die Konvergenz des Verfahrens.

6.2 Fehlerabschätzungen

6.2.1 Fehlerabschätzung in der H^1-Norm

Wir wollen wissen, wie sich der

$$\boxed{\textit{Diskretisierungsfehler} \quad \|u - u_h\|_1}$$

verhält, wenn wir $H_0^1(\Omega)$ durch $S_h \subset S^m$ (siehe (6.2)) approximieren. Nach Theorem (6.1) reicht es dazu , den

$$\boxed{\textit{Approximationsfehler} \quad \inf_{v \in S_h} \|u - v\|_1}$$

abzuschätzen. Wir wollen dies zunächst im einfachst möglichen Fall, nämlich für

$$m = 1, \qquad \Omega = (a, b) \subset \mathbb{R} \quad ,$$

tun. Aus $\mathcal{T}_h$ wird dann das *Gitter*

$$a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b$$

und wir setzen

$$t_i = [x_i, x_{i-1}], \qquad h_i = x_{i-1} - x_i, \qquad h = \max_{i=1,\ldots,N} h_i \quad .$$

S_h bezeichnet dann den Raum der stetigen, stückweise linearen Funktionen. Die folgenden Schlußweisen sind so angelegt, daß sie auf den zwei– und dreidimensionalen Fall übertragbar sind.

Satz 6.9 *Für jedes $u \in H_0^1(\Omega)$ gilt*

$$\inf_{v \in \mathcal{S}_h} \|u - v\|_{H^1(\Omega)} \to 0$$

für $h \to 0$.

Beweis:

Übung.

□

Bemerkung: Aus Satz 6.9 folgt sofort mit dem Céa–Lemma 5.16, die Konvergenz

$$\|u - u_h\|_{H^1(\Omega)} \to 0$$

der FE-Approximationen u_h für $h \to 0$. Beachte, daß dazu außer den Voraussetzungen aus dem Existenz– und Eindeutigkeitssatz 5.31 *keine zusätzlichen Regularitätsannahmen nötig sind!*
Wir wollen nun wissen, *wie schnell* die Nährungslösungen konvergieren. Um entsprechende schärferen Aussagen machen zu können, benötigen wir die

$$\boxed{\textit{Regularitätsvoraussetzung::}\ \ u \in H^2(a,b).}$$

Nach dem Sobolev'schen Einbettungssatz 5.27 ist

$$H^2(a,b) \subset C[a,b]$$

mit stetiger Einbettung. (Dies ist auch in zwei– und drei Raumdimensionen richtig, nicht aber in vier!) Damit ist wegen Korollar 6.5 der *Interpolationsoperator* $I_h : H^2(\Omega) \to S_h$ durch

$$\boxed{I_h v \in S_h : \qquad I_h v(x_i) = v(x_i) \qquad \forall i = 0, \dots, N}$$

wohldefiniert. Offenbar gilt

$$\inf_{v \in S_h} \|u - v\|_1 \leq \|u - I_h u\|_1$$

Es reicht also, den

$$\boxed{\textit{Interpolationsfehler} \quad \|u - I_h u\|_1}$$

abzuschätzen. Das geschieht in vier Schritten.

- **1. Schritt: Lokalisierung.** Aus der Additivität von Integralen folgt

$$\|u - I_h u\|_1^2 = \sum_{i=1}^{N} \|u - I_h u\|_{H^1(t_i)}^2$$

- **2. Schritt: Transformation auf das Einheitsintervall.** Für jedes feste i transformieren wir das Einheitsintervall $T = [0,1]$ affin auf t_i gemäß

$$t_i = F_i(T), \qquad x = F_i(\xi) = x_{i-1} + h_i\xi, \quad \xi = F^{-1}(x) = h_i^{-1}x - x_{i-1})$$

und setzen

$$\hat{v}(\xi) = v(F_i(\xi)) = v(x) \qquad \forall v \in H^2(t_i) \quad .$$

Nach der Transformationsregel für Integrale gilt

$$\|v\|^2_{L^2(t_i)} = \int\limits_{x_{i-1}}^{x_i} v(x)^2 \; dx = h_i \int\limits_0^1 \hat{v}(\xi)^2 \; d\xi = h_i \, \|\hat{v}\|^2_{L^2(T)} \quad .$$

Die Kettenregel liefert

$$\hat{v}'(\xi) = \frac{d}{d\xi}\hat{v}(\xi) = \frac{d}{dx}v'(x)\frac{dx}{d\xi} = h_i v'(x)$$

und daher

$$\|v'\|^2_{L^2(t_i)} = h_i^{-1} \, \|\hat{v}'\|^2_{L^2(T)}$$

Für $h_i \leq 1$ gilt also

$$\|v\|^2_{H^1(t_i)} \leq h_i^{-1} \, \|\hat{v}\|^2_{H^1(T)} \qquad \forall v \in H^2(t_i) \quad . \tag{6.6}$$

Strengenommen haben wir diese Abschätzung nur für im klassichen Sinne differenzierbare Funktionen gezeigt (punktweise Anwendung der Kettenregel). Die Gültigkeit der obigen Transformationsregel für alle $v \in H^2(t_i)$ folgt dann mit einem *Dichteschluß* (ausführlich im nächsten Schritt).

- **3. Schritt: Lokaler Interpolationsfehler.** Anwendung der Transformationsregel (6.5) auf $v = u - I_h u$ liefert

$$\|u - I_h u\|^2_{H^1(t_i)} \leq h_i^{-1} \|\hat{u}_i - \hat{I}\hat{u}_i\|^2_{H^1(T)}$$

wobei

$$\hat{I}v(x) = v(0) + (v(1) - v(0))x \qquad v \in H^2(T)$$

den Interpolationsoperator auf dem Referenzintervall T bezeichnet. Wir wollen nun die Abschätzung

$$\|v - \hat{I}v\|^2_{H^1(T)} \leq c\,\|v''\|^2_{L^2(T)} \qquad \forall v \in H^2(T) \tag{6.7}$$

mit $c = \frac{1}{3}(8 + 2\sqrt{3})$ beweisen. Beweistechnik: Wir weisen die Abschätzung erst für Funktionen aus einem dichten Teilraum glatter Funktionen nach und erweitern deren Gültigkeit dann mittels Dichteschluß.

Bekanntlich ist $X = \{\varphi|_T \mid \varphi \in C_0^\infty(\mathbb{R})\} \subset H^2(T)$ dicht. Sei also $\varphi \in X$ beliebig, aber fest gewählt. Taylor–Entwicklung liefert

$$\varphi(x) = \varphi(0) + \varphi'(0)x + \int\limits_0^x (x - z)\varphi''(z) \; dz \quad .$$

Einsetzen dieser Darstellung für $\varphi(x)$ und $\varphi(1)$ ergibt

$$\varphi(x) - \hat{I}\varphi(x) = \int\limits_0^x (x - z)\varphi''(z) \; dz - x\int\limits_0^1 (1 - z)\varphi''(z) \; dz \quad .$$

Mit der Cauchy–Schwarzschen Ungleichung erhält man

$$\begin{aligned}
\|\varphi - \hat{I}\varphi\|^2_{L^2(T)} &= \int_0^1 \left(\int_0^x (x-z)\varphi''(z)\,dz - x\int_0^1 (1-z)\varphi''(z)\,dz \right)^2 dx \\
&\leq \int_0^1 \Bigg(\left(\int_0^x (x-z)^2\,dz \right)^{\frac{1}{2}} \left(\int_0^x \varphi''(z)^2\,dz \right)^{\frac{1}{2}} \\
&\qquad + x \left(\int_0^1 (1-z)^2\,dz \right)^{\frac{1}{2}} \left(\int_0^1 \varphi''(z)^2\,dz \right)^{\frac{1}{2}} \Bigg)^2 dx \\
&\leq 4\int_0^1 (1-z)^2\,dz \int_0^1 \varphi''(z)^2\,dz \\
&= \tfrac{4}{3}\,\|\varphi''\|^2_{L^2(T)} \quad .
\end{aligned}$$

Nach bekannten Differentiationsregeln gilt

$$\big(\varphi - \hat{I}\varphi\big)'(x) = (x-x)\varphi''(x) + \int_0^x \varphi''(z)\,dz - \int_0^1 (1-z)\varphi''(z)\,dz.$$

Daraus erhält man ähnlich wie oben

$$\|\big(\varphi - \hat{I}\varphi\big)'\|^2_{L^2(T)} \leq (1 + \tfrac{1}{3}\sqrt{3})^2 \|\varphi''\|^2_{L^2(T)}.$$

Insgesamt folgt

$$\|\varphi - \hat{I}\varphi\|^2_{H^1(T)} \leq \tfrac{1}{3}(8 + 2\sqrt{3})\|\varphi''\|^2_{L^2(T)}$$

Es folgt der Dichteschluß. Dazu sei $v \in H^2(T)$ beliebig, aber fest gewählt.
Da $X \subset H^2(T)$, dicht, gibt es eine Folge $(\varphi_k)_{k\in\mathbb{N}} \subset X$, die in $H^2(T)$ gegen v konvergiert, also

$$\|v - \varphi_k\|_{H^2(T)} \to 0, \qquad k \to \infty \quad .$$

Aus dem Einbettungsresultat Satz 5.27 wissen wir außerdem, daß

$$\|\hat{I}v\|_{H^1(T)} \leq \sqrt{5} \max_{x\in[0,1]} |v(x)| \leq C\,\|v\|_{H^2(T)} \qquad \forall v \in H^2(T)$$

gilt. Insgesamt folgt nun mit der Dreiecksungleichung

$$\begin{aligned}
\|v - \hat{I}v\|_{H^1(T)} &\leq \|\varphi_k - \hat{I}\varphi_k\|_{H^1(T)} + \|v - \varphi_k\|_{H^1(T)} + \|\hat{I}(v - \varphi_k)\|_{H^1(T)} \\
&\leq c\,\|v''\|_{L^2(T)} + c\,\|v'' - \varphi_k''\|_{L^2(T)} + (1 + C)\,\|v - \varphi_k\|_{H^2(T)} \\
&\leq c\,\|v''\|_{L^2(T)} + (1 + c + C)\,\|v - \varphi_k\|_{H^2(T)} \quad .
\end{aligned}$$

Grenzübergang $k \to 0$ liefert die Behauptung.

- **4. Schritt: Rücktransformation.** Wie oben zeigt man mittels Kettenregel, daß

$$\|\hat{u}_i''\|^2_{L^2(T)} = h_i^3\,\|u''\|^2_{L^2(t_i)}$$

Jetzt können wir alles zusammenbauen, um die gewünschte Abschätzung des Interpolationsfehlers zu erhalten

$$\begin{aligned}
\|u - I_h u\|^2_{H^1(\Omega)} &= \sum_{i=1}^{N} \|u - I_h u\|^2_{H^1(t_i)} \\
&\le \sum_{i=1}^{N} h_i^{-1} \|\hat{u}_i - \hat{I}\hat{u}_i\|^2_{H^1(T)} \\
&\le c \sum_{i=1}^{N} h_i^{-1} \, \|\hat{u}_i''\|^2_{L^2(T)} \\
&\le c \sum_{i=1}^{N} h_i^{-1} h_i^3 \, \|u''\|^2_{L^2(t_i)} \\
&\le ch^2 \, \|u''\|^2_{L^2(\Omega)} \quad .
\end{aligned}$$

Damit haben wir unter der *Regularitätsvoraussetzung* $u \in H^2(\Omega)$ die *a priori Fehlerabschätzung*

$$\boxed{\|u - u_h\|_1 \le ch\|u''\|_{L^2(\Omega)} \le ch\|u\|_{H^2(\Omega)}}$$

bewiesen.

Dieses Resultat lässt sich auf den Fall $m \ge 1$ und $\Omega \subset \mathbb{R}^2$ erweitern. Analog zur Transformation auf das Einheitsintervall spielt beim Beweis die affine Transformation F_t auf des Einheitsdreiecks mit Eckpunkten $(0,0)$, $(1,0)$ und $(0,1)$ auf $t \in \mathcal{T}_h$ eine zentrale Rolle. Diese Transformation ist singulär, wenn t zu einem Intervall entartet. Der Grad der Entartung eines Dreiecks $t \in \mathcal{T}_h$ wird durch das Verhältnis von Umkeisradius r_t zum Innkreisradius ϱ_t beschrieben.

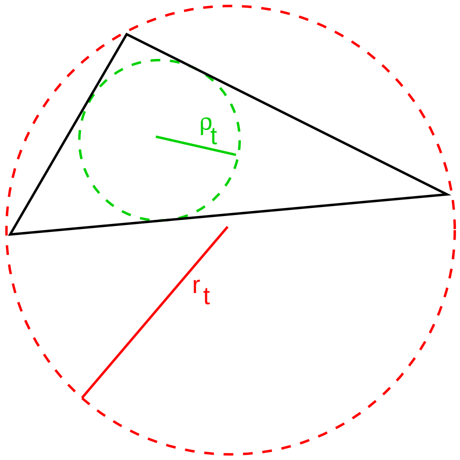


Abbildung 15: Innen- und Umkreisradius ϱ_t und r_t eines Dreiecks.

Das schlechtestmögliche Verhältnis

$$\sigma_h = \max_{t \in \mathcal{T}_h} \frac{r_t}{\varrho_t} \tag{6.8}$$

nennt man *Formregularität* oder kurz *Regularität* von $\mathcal{T}_h$.
Nun können wir unseren zentralen *Approximationssatz* formulieren

Theorem 6.10 *Es sei $m \geq 1$, h klein genug und $u \in H_0^1(\Omega) \cap H^{m+1}(\Omega)$. Dann gilt die a priori Abschätzung*

$$\|u - u_h\|_1 \leq c\,\sigma_h\, h^m\, |u|_{m+1}\,, \qquad |u|_{m+1} = \sum_{|\beta|=m+1} \|\partial\beta u\|_{L^2(\Omega)}$$

mit σ_h aus (6.7).

Beweis:

Der Beweis erfolgt analog zum 1–D Fall. Im Fall m=1 wird nur affine Transformation technisch etwas aufwendiger. Um den allgemeinen Fall $m > 1$ zu behandeln, muß man bei der Abschätzung des Interpolationsfehlers auf dem Referenzdreieck etwas geschickter vorgehen (Stichwort: Bramble–Hilbert–Lemma) Eine ausführliche Darstellung findet sich zum Beispiel auf Kapitel II, Paragraph 6 im Buch von Braess [5]

□.

Wollen wir die Lösung $u \in H_0^1(\Omega) \cap H^m(\Omega)$ durch eine *Folge von Triangulierungen* $\mathcal{T}_h$, $h \in \mathcal{H} = \{h_1 > h_2 > \ldots\}$ immer genauer approximieren, so müssen wir darauf achten, daß σ_h beschränkt bleibt.

Definition 6.11 *Sei $\mathcal{T}_h$, $h \in \mathcal{H} = \{h_1 > h_2 > \ldots\}$ eine Familie von Triangulierungen mit immer kleiner werdenden Schrittweiten.*
Dann heißt die Folge $(\mathcal{T}_h)_{h\in\mathcal{H}}$ <u>formregulär</u> oder kurz <u>regulär</u>, falls es eine Zahl $\sigma > 0$ gibt, so daß

$$\sup_{h\in\mathcal{H}} \sigma_h \leq \sigma$$

gilt.

Offenbar ist $(\mathcal{T}_h)_{h\in\mathcal{H}}$ genau dann regulär, wenn die Innenwinkel der Triangulierungen $\mathcal{T}_h$ gleichmäßig nach unten beschränkt sind. Dann gilt offenbar

$$\|u - u_h\|_1 = \mathcal{O}(h^m),$$

falls $u \in H_0^1(\Omega) \cap H^{m+1}(\Omega)$.
Achtung: Unsere Beweistechnik liefert *nicht*

$$u \in H_0^1(\Omega) \quad \Longrightarrow \quad \|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h) \quad .$$

6.2.2 Fehlerabschätzung in der L^2–Norm

Sei von nun an $m = 1$. Aus Theorem 6.10 können wir nichts über den L^2–Fehler erfahren, denn dort ist $m \geq 1$ vorausgesetzt. Wir zeigen jetzt auf andere Weise, daß unter gewissen Voraussetzungen doch

$$u \in H^2(\Omega) \quad \Longrightarrow \quad \|u - u_h\|_{L^2(\Omega)} = \mathcal{O}\left(h^2\right)$$

gilt.

Satz 6.12 *Sei $\mathcal{T}_h$, $h \in \mathcal{H}$ eine reguläre Familie vom Triangulierungen. Das sogenannte* *<u>duale Problem</u>*

$$w \in H_0^1(\Omega) : \quad a(v, w) = (g, v) \qquad \forall v \in H_0^1(\Omega) \tag{6.9}$$

sei H^2–regulär, d.h. $\|w\|_2 \leq c\, \|g\|_{L^2(\Omega)}$. Dann gilt

$$\boxed{\|u - u_h\|_{L^2(\Omega)} \leq ch\, \|u - u_h\|_1 \,.}$$

Ist das Ausgangsproblem auch H^2–regulär, so gilt sogar

$$\boxed{\|u - u_h\|_{L^2(\Omega)} \leq ch^2\, \|f\|_{L^2(\Omega)} \,.}$$

Beweis:
(Nitsche–Trick, Aubin(67)–Nitsche(68)–Lemma)

1) Duales Problem
   Löse das duale Problem (6.8) für $g = u - u_h$. Dann folgt

$$a(v, w) = (u - u_h, v) \quad \forall v \in H_0^1(\Omega) \quad .$$

   Finite–Elemente–Approximation von w

$$w_h \in S_h : \quad a(v, w_h) = (u - u_h, v) \qquad \forall v \in S_h \quad .$$

   Fehlerabschätzung ($w \in H^2(\Omega) \cap H_0^1(\Omega)$ gilt nach Voraussetzung !)

$$\|w - w_h\|_1 \leq ch\, |w|_2 \leq ch\, \|g\|_{L^2(\Omega)} = ch\, \|u - u_h\|_{L^2(\Omega)} \quad .$$

2) Orthogonalität
   Es gilt bekanntlich

$$a(u - u_h, v) = 0 \qquad \forall v \in S_h \quad .$$

3) Einsetzen von $v = u - u_h$

$$\begin{aligned}
\|u-u_h\|^2_{L^2(\Omega)} &= (u-u_h, u-u_h) = a(u-u_h, w) \\
&= a(u-u_h, w-w_h) \qquad \text{(Orthogonalität)} \\
&\leq \Gamma \|u-u_h\|_1 \|w-w_h\|_1 \\
&\leq \Gamma \|u-u_h\|_1 \, ch \, \|u-u_h\|_{L^2(\Omega)} \quad ,
\end{aligned}$$

und es folgt der erste Teil der Behauptung. Der zweite folgt direkt aus

$$\|u-u_h\|_1 \leq ch \, \|f\|_{L^2(\Omega)}$$

für H^2–reguläre Probleme.

□

6.2.3 Adaptive Gitterverfeinerung

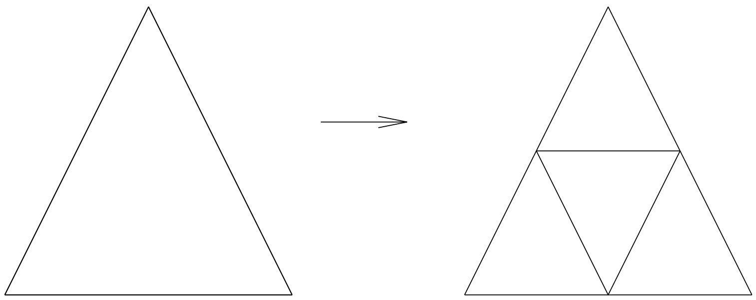


Abbildung 16: Reguläre Verfeinerung eines Dreiecks

Wir wollen kurz die Konstruktion einer Folge von formregulären Triangulierungen $(\mathcal{T}_h)_{h\in\mathcal{H}}$ andiskutieren. Dafür müssen wir als erstes eine

Ausgangstriangulierung $\mathcal{T}_0$

erzeugen. Bei einfachen Geometrien (Einheitsquadrat) kann man das von Hand tun. Für komplexere Fälle stehen Standardtechniken, wie z.B. advancing front methods oder Delaunay–Verfahren zur Verfügung. Einzelheiten finden sich bei George [9, 10]. Es soll nicht verschwiegen werden, daß diese Techniken für spezielle Anwendungen mit Multiskalengebieten (Halbleiter, Menschen mit Adern, poröse Medien mit Klüften,...) unbefriedigende Ergebnisse liefern (viel zu viele Gitterpunkte) und vor allem in drei Raumdimensionen noch viele offene Fragen bestehen.

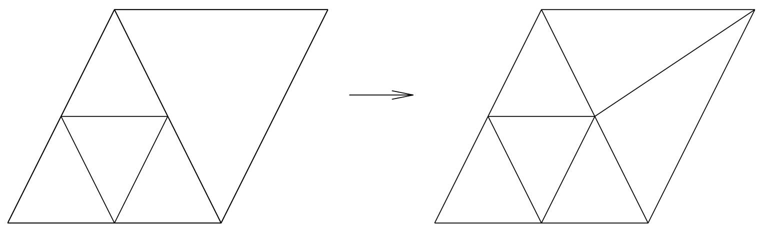


Abbildung 17: Irregulärer Abschluß

Die Ausgangstriangulierung wollen wir verfeinern. Ein zu verfeinerndes Dreieck wollen wir dazu in 4 ähnliche Teildreiecke unterteilen (vgl. Abb. 16). Dabei ändern sich die Innenwinkel nicht. Man spricht daher von einer *regulären* Verfeinerung
Wenn man nicht alle Dreiecke unterteilen will, muß man irgendwo aufhören. Um "hängende Knoten" zu vermeiden, kann man *irreguläre Abschlüsse* einführen, wie in Abb. 17 dargestellt. Offenbar werden dabei die Innenwinkel halbiert.

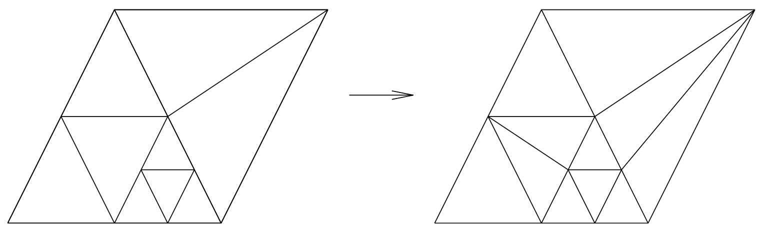


Abbildung 18: Degenerierte Dreiecke durch mehrfache irreguläre Abschlüsse

Fortgesetztes Halbieren führt auf degenerierende Dreiecke (siehe Abb. 18). Aus diesem Grund hat R.E. Bank schon Anfang der achtziger Jahre vorgeschlagen, irreguläre Verfeinerungen vor jedem neuen Verfeinerungsschritt rückgängig zu machen und nie ein Dreieck mehr als einmal irregulär zu verfeinern.
Insgesamt erhält man so folgenden

Verfeinerungsalgorithmus (vergl. Abb. 19)
Gegeben sei die Triangulierung $\mathcal{T}_k$, $k \geq 0$,

1) Falls $k > 0$ lösche alle irregulären Abschlüsse

2) Markiere eine Teilmenge von $\mathcal{T}_h$ zur Verfeinerung

3) Verfeinere alle markierten Dreiecke regulär

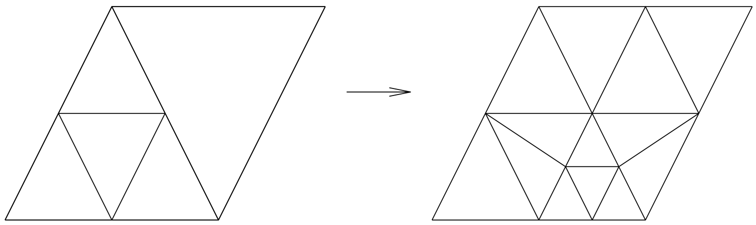


Abbildung 19: Verfeinerung ohne mehrfache irreguläre Abschlüsse

4) Verfeinere alle Dreiecke mit mehr als einer verfeinerten Kante oder mit doppelt verfeinerter Kante regulär

5) Verfeinere alle Dreiecke mit verfeinertem Nachbar irregulär.

Dieser Algorithmus terminiert (Übung).
Natürlich möchte man die zu markierenden Dreiecke so auswählen, daß

möglichst hohe Approximationsgenauigkeit
bei möglichst niedriger Knotenanzahl

erreicht wird. Man möchte also da verfeinern, wo es sich lohnt.
Zur Auswahl der entsprechenden Dreiecke benötigt man geeignete *Verfeinerungsindikatoren*. Ein möglicher Indikator ist der

$$\text{lokale Fehler } \|u - \tilde{u}\|_{H^1(t)},$$

wobei $\tilde{u} \in S_k$ typischerweise eine Approximation von $u_h \in S_k$ darstellt, die man z.B. durch näherungsweise Lösung des entsprechenden lineraren Gleichungssystems erhält. Natürlich ist der lokale Fehler numerisch nicht zugänglich, und muß durch geeignete a posteriori Fehlerschätzungen ersetzt werden, die man auf der Basis von $\tilde{u}$ berechnen kann. Wir verweisen beispielsweise hier auf Kapitel III, Paragraph 7 im Buch von Braess [5]. Der gesamte adaptive Lösungsprozess ist in Abb. 20 dargestellt.

6.3 Kondition der Steifigkeitsmatrix und Fourier–Methode

Wir betrachten wieder das Modellproblem

$$u \in H_0^1(\Omega): \quad a(u,v) = l(v) \qquad \forall v \in H_0^1(\Omega) \quad , \tag{6.10}$$

und setzen voraus, daß $l \in (H_0^1(\Omega))'$ und $a(\cdot,\cdot)$ elliptisch und *symmetrisch* ist, z.B.

$$a(v,w) = \int_\Omega \alpha \nabla v \cdot \nabla w \, dx \quad .$$

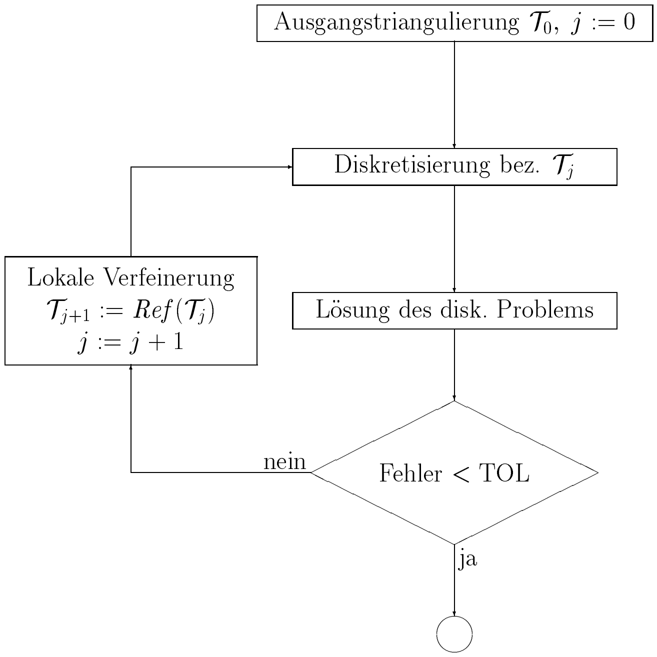


Abbildung 20: Adaptives Finite-Elemente-Verfahren

Diskretisierung mit *linearen* finiten Elementen ($m = 1$) bezüglich einer *regulären* Folge von Triangulierungen $(\mathcal{T}_h)_{h\in\mathcal{H}}$ führt auf das diskrete Problem

$$u_h \in S_h : \quad a(u_h, v) = l(v) \qquad \forall v \in S_h \quad . \tag{6.11}$$

Dies läßt sich als lineares Gleichungssystem

$$AU = b \tag{6.12}$$

für den unbekannten Koeffizientenvektor

$$U = (u_p)_{p\in\mathcal{N}_h}$$

mit

$$u_h = \sum_{p\in\mathcal{N}_h} u_p\lambda_p$$

formulieren. Dabei ist $\Lambda_h = \{\lambda_p,\, p \in \mathcal{N}_h\}$ die Knotenbasis. Steifigkeitsmatrix A und rechte Seite b sind bekanntlich gegeben durch

$$A = (a(\lambda_p, \lambda_q))_{p,q\in\mathcal{N}_h}\ ,\ b = (l(\lambda_p))_{p\in\mathcal{N}_h} \quad .$$

Wir verwenden von nun an den kanonischen Isomorphismus

$$v = \sum_{p\in\mathcal{N}_h} v_p\lambda_p \ \rightarrow\ \underline{v} = (v_p)_{p\in\mathcal{N}_h} \in \mathbb{R}^{n_h}$$

und bezeichnen das euklidische Skalarprodukt mit

$$< \underline{v}, \underline{w} >= \sum_{p\in\mathcal{N}_h} v_p w_p \quad .$$

Dann hat die Steifigkeitsmatrix A die Eigenschaft

$$a(w, v) =< A\underline{w}, \underline{v} > \qquad \forall v \in S_h$$

und $b \in \mathbb{R}^{n_h}$ ist die entsprechende Darstellung von l

$$l(v) =< b, \underline{v} > \quad \forall v \in S_h \quad .$$

Wenn keine Gefahr der Verwechselung besteht, verzichten wir bei Vektoren $\underline{v} \in \mathbb{R}^{n_h}$ auf das Unterstreichen.

Ziel
Bestimme *mit möglichst wenig Rechenoperationen* und *möglichst wenig Speicher* ein $\tilde{u} \in S_h$ oder, äquivalent, ein $\underline{\tilde{u}} \in \mathbb{R}^{n_h}$, so daß

$$\|\tilde{u} - u_h\|_1 \leq c\,h \quad . \tag{6.13}$$

Bemerkung: Optimal wäre ein Verfahren, welches (6.12) nach $\mathcal{O}(n_h)$ Operationen mit $\mathcal{O}(n_h)$ Speicher erreicht, denn selbst für eine Diagonalmatrix A hätte man diesen Aufwand.

Bemerkung: Naive Anwendung des Gauß'schen Algorithmus' erfordert $\mathcal{O}(n_h^3)$ Rechenoperationen.

Bevor wir an die Entwicklung speziell angepaßter linearer Löser gehen, wollen wir in diesem Kapitel einige grundlegende Eigenschaften von A studieren. Vorher erinnern wir kurz an einige Grundlagen aus der linearen Algebra.

Lemma 6.13 *Sei $A \in \mathbb{R}^{n,n}$ symmetrisch und positiv definit (s.p.d.). Dann existiert eine Orthogonalbasis von $\mathbb{R}^n$ aus Eigenvektoren e_i zu positiven Eigenwerten μ_i, $i = 1, \ldots, n$. Weiter gibt es eine s.p.d.–Matrix $A^{\frac{1}{2}} \in R^{n,n}$ mit der Eigenschaft*

$$A^{\frac{1}{2}} A^{\frac{1}{2}} = A \quad .$$

Beweis:

Wir zeigen nur die Existenz von $A^{\frac{1}{2}}$. Es sei $T = (e_1 \ldots e_n)$ (spaltenweise). Dann folgt

$$D = T^{-1} A T, \qquad D = \mathrm{diag}(\mu_1, \ldots, \mu_n) \quad .$$

Wir setzen $D^{\frac{1}{2}} = (\mu_1^{\frac{1}{2}}, \ldots, \mu_n^{\frac{1}{2}})$. Offenbar gilt $D^{\frac{1}{2}} D^{\frac{1}{2}} = D$. Schließlich setzen wir

$$A^{\frac{1}{2}} = T D^{\frac{1}{2}} T^{-1}$$

und es folgt die Behauptung.

□

Lemma 6.14 *Sei $B \in \mathbb{R}^{n,n}$ s.p.d. Dann wird durch*

$$< v, w >_B = < Bv, w >$$

ein Skalarprodukt auf $\mathbb{R}^n$ definiert. Ist weiter $A \in \mathbb{R}^{n,n}$ regulär und symmetrisch bezüglich $< \cdot, \cdot >_B$, d.h.

$$< Av, w >_B = < v, Aw >_B \quad ,$$

so gilt

$$\mu_{\max}(A) = \max_{v \neq 0} \frac{< Av, v >_B}{< v, v >_B}$$

$$\mu_{\min}(A) = \min_{v \neq 0} \frac{< Av, v >_B}{< v, v >_B}$$

Beweis:

Die erste Aussage ist trivial.

Wir betrachten zunächst den Fall $B = I$ (Einheitsmatrix). Einsetzen der Orthonormalbasisdarstellung

$$v = \sum_{i=1}^{n} v_i e_i$$

aus Lemma 6.13 liefert

$$\frac{< Av, v >}{< v, v >} = \frac{\sum_{i=1}^{n} \lambda_i < v_i, v_i >}{< v, v >} \leq \mu_{\max}(A) \quad .$$

Gleichheit erhält man für $v = e_{\max}$. Die Darstellung von $\mu_{\min}(U)$ folgt aus

$$\frac{< Av, v >}{< v, v >} = \frac{< w, w >}{< A^{-1}w, w >} \geq \frac{1}{\mu_{\max}(A^{-1})} = \mu_{\min}(A) \quad .$$

mit $w = A^{-1}v$.
Sei nun B s.p.d. Aus der Symmetrie von A bezüglich $< \cdot, \cdot >_B$ folgt

$$< B^{\frac{1}{2}} AB^{-\frac{1}{2}} v, w > = < AB^{-\frac{1}{2}} v, B^{-\frac{1}{2}} w >_B = < B^{-\frac{1}{2}} v, AB^{-\frac{1}{2}} w >_B = < v, B^{\frac{1}{2}} AB^{-\frac{1}{2}} w > \quad ,$$

also die Symmetrie von $B^{\frac{1}{2}} AB^{-\frac{1}{2}}$ bezüglich des Euklidischen Skalarprodukts. Diese Matrix hat offenbar dieselben Eigenwerte wie A und es gilt

$$\frac{< Av, v >}{< v, v >} = \frac{< B^{\frac{1}{2}} AB^{-\frac{1}{2}} w, w >}{< w, w >}$$

mit $w = B^{\frac{1}{2}} v$. Nun folgen die Behauptungen aus den entsprechenden Aussagen im Falle $B = I$.

□

Bemerkung: Der Quotient

$$\frac{< Av, v >_B}{< v, v >_B}$$

heißt *Rayleigh–Quotient*.

Von nun an bedeutet A wieder die Steifigkeitsmatrix (bezüglich der Knotenbasis).

Satz 6.15 *Die Steifigkeitsmatrix A ist* symmetrisch *und* positiv definit. *Die Anzahl der von* 0 *verschiedenen Elemente in jeder Zeile von A ist gleichmäßig in h beschränkt (A ist* schwach besetzt*).*

Beweis:

Übung.

□

Folgerung: Die Auswertung von Av erfordert $\mathcal{O}(n_h)$ Punktoperationen.

Satz 6.16 *Die Kondition $\kappa(A)$ genügt den Abschätzungen*

$$\boxed{\frac{1}{o(1)} \leq \kappa(A) = \frac{\mu_{\max}(A)}{\mu_{\min}(A)} \leq ch^{-2} \quad .} \tag{6.14}$$

Beweis:

1) Es gilt
$$\mu_{\max}(A) = \max_{\underline{v}\in\mathbb{R}^{n_h}} \frac{<A\underline{v},\underline{v}>}{|\underline{v}|^2} = \max_{v\in S_h} \frac{a(v,v)}{|\underline{v}|^2} \le \Gamma\, \frac{\|v\|_1^2}{|\underline{v}|^2} \le c \quad ,$$
denn
$$\|v\|_{1,t}^2 \le c\,\|\hat{v}\|_{1,T}^2 \le C(v^2(p_1)+v^2(p_2)+v^2(p_3)) \quad ,$$
wobei $t=(p_1,p_2,p_3)$. Aufsummieren gibt (wegen der Regulariät von $\mathcal{T}_h$!) die obige Abschätzung.

2) Es gilt
$$\mu_{\min}(A) = \min_{\underline{v}\in\mathbb{R}^{n_h}} \frac{<A\underline{v},\underline{v}>}{|\underline{v}|^2} = \min_{v\in S_h} \frac{a(v,v)}{|\underline{v}|^2} \ge \gamma \frac{\|v\|_1^2}{|\underline{v}|^2} \quad .$$
Für jedes $t=(p_1,p_2,p_3)\in\mathcal{T}$ gilt
$$v^2(p_1)+v^2(p_2)+v^2(p_3) \le c\,\|\hat{v}\|_{L^2(T)}^2 \le c\,h^{-2}\,\|v\|_{L^2(t)}^2 \le C\,h^{-2}\,\|v\|_{1,t}^2 \quad ,$$
und Aufsummieren liefert
$$|\underline{v}|^2 \le c\,h^{-2}\,\|v\|_1^2 \quad ,$$
also
$$\mu_{\min}(A) \ge ch^2 \quad .$$

3) Im Widerspruch zur Behauptung nehmen wir an, daß es ein $\varepsilon>0$ gibt, so daß
$$a(v,v) \ge \varepsilon\,|\underline{v}|^2 \qquad \forall v\in S_h\ ,\ \forall h\in\mathcal{H}$$
richtig ist. Dann wählen wir $\varphi\in C_0^\infty(\Omega)\cap H_0^1(\Omega)$ so, daß für einen Kreis $K\subset\Omega$ mit positivem Radius
$$\varphi(x)\equiv 1 \quad \forall x\in K$$
gilt und setzen $v:=I_h\varphi\in S_h$ (Interpolation). Dann ist
$$|\underline{v}|^2 \ge |\mathcal{N}_h\cap K| \to \infty \quad \text{für} \quad h\to\infty \quad ,$$
aber
$$a(v,v) \le \Gamma\,\|I_h\varphi\|_1^2 \le 2\Gamma\,(\|I_h\varphi-\varphi\|_1^2+\|\varphi\|_1^2) \le \text{const.} \quad \forall h\in\mathcal{H} \quad .$$
Widerspruch.

□

Bemerkung: Die obere Abschätzung in (6.13) ist scharf. Wir werden weiter unten ein Beispiel angeben, in dem $\kappa(A)=\mathcal{O}(h^2)$ gilt.

Bemerkung: Aus Satz (6.16) folgt insbesondere, daß Rundungsfehler beim direkten Lösen unseres Gleichungssystems mit $h\to 0$ beliebig groß werden können. Dies macht natürlich erhebliche Schwierigkeiten bei der direkten Lösung von (6.11).

Es ist von entscheidender Bedeutung, daß die Kondition der Steifigkeitsmatrix von der Wahl der zugrundeliegenden Basis von S_h abhängt.

Satz 6.17 *Sei* $\hat{\Lambda}_h = \{\hat{\lambda}_p \,|\, p \in \mathcal{N}_h\}$ *eine Basis von* S_h *und*

$$v = \sum_{p \in \mathcal{N}_h} \hat{v}_p \hat{\lambda}_p \to \underline{\hat{v}} = (\hat{v}_p)_{p \in \mathcal{N}_h}$$

der entsprechende Isomorphismus, sowie $\hat{A}$ *die Steifigkeitsmatrix bezüglich* $\hat{\Lambda}_h$*, also*

$$a(v,w) = < \hat{A}\underline{\hat{v}}, \underline{\hat{w}} > \qquad \forall v \in S_h \quad .$$

Sei T^T *die Transformationsmatrix von* Λ_h *nach* $\hat{\Lambda}_h$*, also*

$$\hat{\lambda}_p = \sum_{q \in \mathcal{N}_h} T_{qp} \lambda_q \quad .$$

Dann gilt

$$\hat{A} = T^T A T \quad .$$

Beweis:
Einsetzen liefert

$$\begin{aligned} v = \sum_{p \in \mathcal{N}_h} \hat{v}_p \hat{\lambda}_p &= \sum_{p \in \mathcal{N}_h} \hat{v}_p \sum_{q \in \mathcal{N}_h} T_{qp} \lambda_q \\ &= \sum_{p \in \mathcal{N}_h} \left(\sum_{p \in \mathcal{N}_h} T_{qp} \hat{v}_p \right) \lambda_q \quad , \end{aligned}$$

also

$$\underline{v} = T \underline{\hat{v}} \quad .$$

Aus der Darstellung von $\hat{A}$ folgt

$$< \hat{A}\underline{\hat{v}}, \underline{\hat{w}} >= a(v,w) =< A\underline{v}, \underline{w} >=< T^T A T \underline{\hat{v}}, \underline{\hat{w}} >$$

und damit die Behauptung.

□

Da A symmetrisch und positiv definit ist, gibt es eine Orthonormalbasis von Eigenvektoren $\underline{e}_p$, zu Eigenwerten $\mu_p, p \in \mathcal{N}_h$ von A. Wegen

$$a(e_p, e_q) =< A\underline{e}_p, \underline{e}_q >= \mu_p < \underline{e}_p, \underline{e}_q >$$

sind die entsprechenden Funktionen $e_p \in S_h$, $p \in \mathcal{N}_h$, dann orthogonal bezüglich dem Energie–Skalarprodukt $a(\cdot,\cdot)$, kurz a–orthogonal.

Korollar 6.18 *Wählt man zur Darstellung von* $u_h \in S_h$ *die* a*–Orthonormalbasis*

$$\hat{\lambda}_p = \frac{1}{\sqrt{\mu_p}} e_p$$

von S_h*, so erhält man die Steifigkeitsmatrix*

$$\boxed{\hat{A} = I \quad .}$$

Beweis:

Es gilt

$$\underline{a}_{p,q} = a(\hat{\lambda}_p, \hat{\lambda}_q) = < \frac{1}{\sqrt{\mu_p}} A e_p, \frac{1}{\sqrt{\mu_q}} e_q > = \delta_{p,q} \quad .$$

□

In Spezialfällen kann man die Eigenwerte μ_p und Eigenvektoren e_p tatsächlich explizit angeben. Wir betrachten dazu

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

mit $\Omega = [0,1] \times [0,1]$ und $\mathcal{T}_h$ wie in Abbildung 21. Schrittweite $h = \frac{1}{m}$, Anzahl der Unbekannten $n_h = (m-1)^2$.

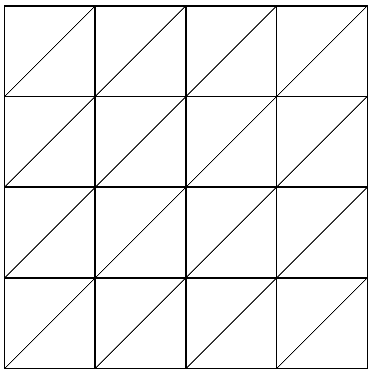


Abbildung 21: Gebiet Ω mit Triangulierung $\mathcal{T}_h$ für $m = 4$

Die Knotenpunkte

$$p_{ij} = (ih, jh)$$

werden *zeilenweise* numeriert. Damit erhält der Unbekanntenvektor U die Gestalt

$$U = (u_p)_{p \in \mathcal{N}_h} = (u_{p_{ij}})_{i,j=1\ldots m-1} = \begin{pmatrix} U_1 \\ \vdots \\ U_{m-1} \end{pmatrix}$$

mit Vektoren $U_j = (u_{p_{ij}})_{i=1,\ldots,m-1}$.

Die Steifigkeitsmatrix A hat die Blocktridiagonalgestalt

$$A = \begin{pmatrix} A_m & -I_m & & & \\ -I_m & A_m & -I_m & & \\ & & \ddots & & \\ & & -I_m & A_m & -I_m \\ & & & -I_m & A_m \end{pmatrix} \tag{6.15}$$

mit den $(m-1)\times(m-1)$-Matrizen

$$A_m = \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & & \ddots & & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{pmatrix} \qquad I_m = \begin{pmatrix} 1 & 0 & & & \\ 0 & 1 & 0 & & \\ & & \ddots & & \\ & & 0 & 1 & 0 \\ & & & 0 & 1 \end{pmatrix}$$

Satz 6.19 *Die Eigenvektoren e_{ij} und die zugehörigen Eigenwerte μ_{ij} von A sind gegeben durch*

$$\begin{aligned} (e_{ij})_{lk} &= \sin(i\pi lh)\cdot\sin(j\pi kh) \\ \mu_{ij} &= 4\left(\sin^2(i\frac{\pi}{2}h)+\sin^2(j\frac{\pi}{2}h)\right) \quad . \end{aligned}$$

Beweis:
Sei $\omega^2=-1$. Mit der Moivre'schen Formel folgt

$$\begin{aligned} \sin\varphi &= \frac{1}{2\omega}\left(e^{\omega\varphi}-e^{-\omega\varphi}\right) \\ \cos\varphi &= \frac{1}{2}\left(e^{\omega\varphi}+e^{-\omega\varphi}\right) \end{aligned}$$

Weiter gilt

$$1-\cos\varphi = 2\sin^2\frac{\varphi}{2} \quad .$$

Wegen

$$\begin{aligned} -e^{\omega i\pi(l-1)h}+2e^{\omega i\pi lh}-e^{\omega i\pi(l+1)h} &= e^{\omega i\pi lh}\left(-e^{-\omega i\pi h}+2-e^{\omega i\pi h}\right) \\ &= 2e^{\omega i\pi lh}\left(1-\cos(i\pi h)\right) \\ &= 4e^{\omega i\pi lh}\sin^2(i\frac{\pi}{2}h) \end{aligned}$$

gilt

$$-\sin(i\pi(l-1)h)+2\sin(i\pi lh)-\sin(i\pi(l+1)h) = 4\sin^2(i\frac{\pi}{2}h)\sin(i\pi lh) \quad .$$

Insbesondere sind damit

$$e_i = (\sin(i\pi lh))_{l=1,\ldots,m-1}\in\mathbb{R}^{m-1}, \qquad \mu_i = 4sin^2(i\frac{\pi}{2}h)$$

Eigenvektoren und Eigenwerte der $(m-1)\times(m-1)$-Matrix B

$$B = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & & \ddots & & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix} \quad .$$

Wir schreiben e_{ij} als Blockvektor

$$e_{ij} = \begin{pmatrix} E_1 \\ \vdots \\ E_{m-1} \end{pmatrix} \quad , \quad E_k = \sin(j\pi kh) \cdot e_i \quad .$$

Blockweise Auswertung von Ae_{ij} liefert dann im k–ten Block

$$\begin{aligned}
(Ae_{ij})_k &= -E_{k-1} + A_m E_k - E_{k+1} \\
&= -E_{k-1} + 2E_k + BE_k - E_{k+1} \\
&= (-\sin(j\pi(k-1)h) + 2\sin(j\pi kh) - \sin(j\pi(k+1)h))\, e_i + \sin(j\pi kh) B e_i \\
&= 4\sin^2(j\frac{\pi}{2}h)\sin(j\pi kh)e_i + 4\sin^2(i\frac{\pi}{2}h)\sin(j\pi kh)e_i \\
&= 4\left(\sin^2(i\frac{\pi}{2}h) + \sin^2(j\frac{\pi}{2}h)\right) E_k \quad .
\end{aligned}$$

Das ist gerade die Behauptung.

□

Bemerkung: Die Kondition der Steifigkeitsmatrix A aus (6.14) ist

$$\boxed{\kappa(A) = \frac{\mu_{\max}}{\mu_{\min}} = \frac{\mu_{m-1,m-1}}{\mu_{1,1}} = \mathcal{O}(h^{-2}) \quad .}$$

Bemerkung: Satz 6.19 wird anschaulich verständlich durch die Tatsache, daß die Eigenfunktionen des Lapalace–Operators

$$-\Delta\varphi_{ij} = \left((i\pi)^2 + (j\pi)^2\right)\varphi_{ij}$$

mit Nullrandbedingungen durch

$$\varphi_{ij}(x,y) = \sin(i\pi x)\cdot\sin(j\pi y) \qquad i,j \in \mathbb{N}$$

gegeben sind. Die Eigenvektoren von A sind also in diesem Fall gerade die punktweisen Restriktionen der entsprechenden Eigenfunktionen.
Übrigens gilt:

$$\frac{1}{h^2}\mu_{ij} \to (i\pi)^2 + (j\pi)^2 \quad \text{für} \quad h \to 0 \quad .$$

Beweis als Übung.

Bemerkung: Beachte, daß die Eigenvektoren sich stark in ihrer *Frequenz* unterscheiden. Es gibt *niederfrequente* Eigenvektoren ($i,j \ll \frac{m}{2}$) und *hochfrequente* Eigenvektoren ($i,j \gg \frac{m}{2}$) und alles dazwischen.

Unter Verwendung von Satz 6.19 können wir die exakte Lösung u_j direkt angeben.

Satz 6.20 *Sei $b = (b_{ij})_{i,j=1,\ldots,m-1}$ die rechte Seite unseres Gleichungssystems. Dann gilt*

$$\boxed{u_{p_{ij}} = \sum_{l,k=1}^{m-1} \frac{(e_{lk})_{ij}}{|e_{lk}|} \mu_{lk}^{-1} \hat{b}_{lk} \quad , \qquad \hat{b}_{lk} := \sum_{r,s=1}^{m-1} \frac{(e_{lk})_{rs}}{|e_{lk}|} b_{rs} \quad .}$$

Beweis:

Bildet man die Transformationsmatrix T durch die spaltenweisen Einträge der zu eins normierten Eigenvektoren $\frac{(e_{lk})}{|e_{lk}|}$, so folgt

$$\left(T^T A T\right) T^T U = T^T b \quad .$$

Also ist wegen $T^T A T = D := \operatorname{diag}(\mu_{ij})$

$$U = T D^{-1} T^T b \quad .$$

Das ist die Matrixschreibweise der obigen Darstellung.

□

Bemerkung: Für die Auswertung der Lösung nach Satz 6.20 benötigt man bei naiver Implementierung der Summation $\mathcal{O}(n_h^2)$ Rechenoperationen. Die Matrix $\hat{A} = T^T A T$ ist eben voll besetzt. Ganz entscheidend ist nun, daß sich die Summen in Satz 6.20 mittels *schneller Fourier–Transformation* (engl. *f*ast *F*ourier–*t*ransform, kurz FFT) auf geschickte Weise mit nur $\mathcal{O}(\log n_h)$ Operationen durchführen lassen. Man beachte z.B. den Unterschied zwischen $n_h = 10^6$ und $\log n_h = 6$. Für die Einzelheiten verweisen wir z.B. auf [20]. Damit erhält man insgesamt einen Aufwand von

$$\boxed{\text{Anzahl der Punktoperationen} = \mathcal{O}(n \log n)}$$

zur Lösung unseres Gleichungssystems. Das eben beschriebene Lösungsverfahren heißt Fourier–Methode.

Bemerkung: Unter der Voraussetzung, daß *die Nebendiagonal–Blöcke mit den Diagonal–Blöcken vertauschbar sind* lässt sich die Fourier–Methode auf andere blocktridiagonale Gleichungssysteme erweitern. Schon im Falle nichtkonstanter Koeffizienten $\alpha(x)$ ist das leider nicht mehr der Fall.

Im allgemeinen kennen wir also die Eigenvektoren $\underline{e}_p$ von A nicht. Bei unserem Modellproblem haben wir gesehen, daß die zugehörigen **a–orthogonalen Finite–Element–Funktionen $\mathbf{e_p} \in \mathbf{S_h}$ eine Skala von Frequenzen** darstellen. Das ist, wie man mathematisch zeigen und physikalisch einsehen kann, tatsächlich auch im allgemeinen der Fall. Wenn man also an a–orthogonalen oder zumindest „fast“ a–orthogonalen Funktionen interessiert ist, so sollte man es mit solchen Funktionen versuchen, die eine möglichst große Skala von Frequenzen überdecken.
Das ist gerade die Grundidee von Mehrgittermethoden.

6.4 Mehrgitter–Methoden

Wir betrachten nach wie vor das Modellproblem (6.9) mit der *symmetrischen* Bilinearform

$$a(v,w) = \int_\Omega \alpha \nabla v \cdot \nabla w \, dx \quad .$$

Gitterhierarchie. Bei der Diskretisierung

$$u_h \in S_h : \quad a(u_h, v) = l(v) \qquad \forall v \in S_h \tag{6.16}$$

nehmen wir an, daß die Triangulierung $\mathcal{T}_h = \mathcal{T}_{h_j}$ durch j Verfeinerungsschritte aus einer Ausgangstriangulierung $\mathcal{T}_{h_0}$ entstanden ist. Der Einfachheit halber gehen wir von uniformer Verfeinerung aus. Daher gilt

$$h_j = \mathcal{O}(2^{-j}), \qquad h_{k-1} = 2h_k \quad k = 1, \dots, j \quad . \tag{6.17}$$

Wir schreiben kurz

$$\mathcal{T}_k = \mathcal{T}_{h_k}, \qquad k = 0, \dots, j \quad .$$

Der geschachtelten Folge von Triangulierungen

$$\mathcal{T}_0 \subset \mathcal{T}_1 \subset \cdots \subset \mathcal{T}_j$$

mit Knotenmengen

$$\mathcal{N}_0 \subset \mathcal{N}_1 \subset \cdots \subset \mathcal{N}_j$$

entspricht dann die geschachtelte Folge von Finite–Element–Räumen

$$S_0 \subset S_1 \subset \cdots \subset S_j = S_{h_j}$$

mit den Knotenbasen

$$\Lambda_k = \{\lambda_p^{(k)} \mid p \in \mathcal{N}_k\}$$

auf jedem Verfeinerungslevel $k = 0, \dots, j$. Wegen der uniformen Verfeinerung wächst die Anzahl der Knoten $n_k = \#\mathcal{N}_k$ geometrisch, d.h. es existiert ein $q > 1$ mit

$$n_k \geq q n_{k-1} \qquad k = 1, \dots, j. \tag{6.18}$$

Bekanntlich ist das diskrete Problem (6.15) äquivalent zum Minimierungsproblem

$$u_j \in S_j : \qquad J(u_j) \leq J(v) \qquad \forall v \in S_j \tag{6.19}$$

mit

$$J(v) = \tfrac{1}{2} a(v,v) - l(v) \qquad v \in S_j \quad .$$

und $u_j = u_{h_j}$.

Teilraumkorrekturmethoden. Wir wollen nun iterative Lösungsverfahren für (6.18) konstruieren. Dazu ersetzen wir das „große“ Minimierungsproblem (6.18) durch eine Folge „kleiner“ Minimierungsprobleme. Dazu wählen wir eine Zerlegung

$$S_j = V_0 + V_1 + \cdots + V_m \tag{6.20}$$

in Teilräume

$$V_k \subset S_j, \qquad k = 0, \ldots, m \quad .$$

Sukzessive Minimierung der Energie J führt dann auf das folgende Iterationsverfahren zur Berechnung einer neuen Iterierten $u_j^{\nu+1} \in S_j$ aus einer gegebenen Iterierten $u_j^\nu \in S_j$.

Algorithmus 6.1: Sukzessive Minimierung

gegeben: $w_{-1} = u_j^\nu \in S_j$.

for $k = 0, \ldots, m$ löse:

$$v_k \in V_k: \quad J(w_{k-1} + v_k) \le J(w_{k-1} + v) \;\; \forall v \in V_k, \qquad w_k = w_{k-1} + v_k, \tag{6.21}$$

neue Iterierte: $u_j^{\nu+1} = w_m$.

Damit erhält man aus jeder Zerlegung (6.19) sofort ein entsprechendes Iterationsverfahren. Wir wollen nun versuchen die Zerlegung (6.19) so zu wählen, daß das entsprechende Iterationsverfahren möglichst schnell konvergiert und mit Aufwand $\mathcal{O}(n_j)$ implementiert werden kann.

Gauß–Seidel–Verfahren. Offenbar kann man die Teilprobleme (6.20) ohne Schwierigkeiten exakt lösen, wenn man eindimensionale Teilräume

$$V_l = \text{spann}\{\lambda_l\}, \qquad l = 1, \ldots, m,$$

wählt. Am naheliegendsten ist dann die Zerlegung

$$S_j = \sum_{l=1}^{n_j} V_l, \qquad V_l = \text{spann}\{\lambda_{p_l}^{(j)}\}, \qquad l = 1, \ldots, n_j. \tag{6.22}$$

Das resultierende Iterationsverfahren ist dann

$$\boxed{u_j^{\nu+1} = u_j^\nu + \sum_{l=1}^{n_j} v_l, \qquad v_l = \frac{l(\lambda_{p_l}^{(j)}) - a(w_{l-1}, \lambda_{p_l}^{(j)})}{a(\lambda_{p_l}^{(j)}, \lambda_{p_l}^{(j)})} \lambda_{p_l}^{(j)},} \tag{6.23}$$

wobei $w_0 = u_j^\nu$ und $w_l = w_{l-1} + v_l$, $l = 1, \ldots, n_j - 1$ gesetzt ist.

Satz 6.21 *Das Iterationsverfahren (6.22) ist gerade das Gauß–Seidel–Verfahren. Die Korrektur*

$$v^{(j)} = v_1 + \cdots + v_{n_j} = u_j^{\nu+1} - u_j^{\nu}$$

ist Lösung des Variationsproblems

$$v^{(j)} \in S_j : \qquad b_j(v^{(j)}, v) = l(v) - a(u_j^{\nu}, v) \qquad \forall v \in S_j \tag{6.24}$$

wobei

$$b_j(v,w) = \sum_{\substack{i,l=1\\ i\le l}}^{n_j} v(p_i) a\big(\lambda_{p_i}^{(j)}, \lambda_{p_l}^{(j)}\big) w(p_l)\ , \qquad v, w \in S_j\ , \tag{6.25}$$

gesetzt ist.

Beweis:

Sei $l = 1, \ldots, n_j$ beliebig aber fest gewählt. Aus $w_{l-1} = u_j^{\nu} + \sum_{i=0}^{l-1} v_i$ (6.22) folgt sofort

$$\begin{aligned}
l\big(\lambda_{p_l}^{(j)}\big) - a\big(u_j^{\nu}, \lambda_{p_l}^{(j)}\big) &= v_l(p_l) a\big(\lambda_{p_l}^{(j)}, \lambda_{p_l}^{(j)}\big) + \sum_{i=0}^{l-1} a\big(v_i, \lambda_{p_l}^{(j)}\big)\\
&= \sum_{i=0}^{l} v^{(j)}(p_i) a\big(\lambda_{p_i}^{(j)}, \lambda_{p_l}^{(j)}\big)\\
&= b\big(v^{(j)}, \lambda_{p_l}^{(j)}\big) \quad .
\end{aligned}$$

Da l beliebig war, folgt die Behauptung.
Die Matrixschreibweise von (6.23) lautet

$$B\underline{v}^{\nu} = b - A\underline{u}^{\nu} \quad . \tag{6.26}$$

Dabei ist

$$A = (a_{p_i,p_l})_{i,l=1}^{n_j}, \quad a_{p_i,p_l} = a\big(\lambda_{p_i}^{(j)}, \lambda_{p_l}^{(j)}\big), \qquad b = \big(l(\lambda_{p_i}^{(j)})\big)_{i=1}^{n_j}$$

und die Matrix

$$B = L + D$$

resultiert aus der Zerlegung von $A = L+D+R$ in Subdiagonalanteil L, Diagonale D und Superdiagonalanteil R. Damit ist (6.25) gerade das bekannte Gauß–Seidel–Verfahren (vgl. z.B. Braess [5][Kapitel 4, Paragraph 1]).

□

Die Konvergenzrate des Gauß–Seidel–Verfahrens degeneriert exponentiell mit j:

Satz 6.22 *Es gilt für jedes $u_j^0 \in S_j$*

$$\boxed{\|u_j - u_j^{\nu+1}\| \le (1 - ch^2)\|u_j - u_j^{\nu}\| \quad \forall \nu \ge 0 \quad .}$$

mit der Energienorm $\|\cdot\| = a(\cdot,\cdot)^{1/2}$ und einer j–unabhängigen Zahl $c > 0$.

Beweis:

In Lehrbüchern wird meist nur die Anwendung auf das diskrete Modellproblem (6.14) betrachtet. Ein allgemeiner Beweis findet sich bei Kornhuber [16][Satz 7.32].

□

Multilevel–Gauß–Seidel. Nach Satz 6.22 ist das Gauß–Seidel–Verfahren zwar global konvergent, die Konvergenzgeschwindigkeit nimmt allerdings mit wachsender Anzahl von Unbekannten sehr schnell ab.
Woran liegt das? Die Antwort ergibt sich aus folgender Bemerkung.

Satz 6.23 *Ist (6.19) eine a–orthogonale Zerlegung, d.h. gilt*

$$a(v,w) = 0, \qquad \forall v \in V_i, w \in V_k, \quad i \neq k \quad ,$$

so liefert der entsprechende Algorithmus 6.1 für jeden Startwert u_j^0 in einem Schritt die exakte Lösung u_j.

Beweis:

Übung.

□

In Kenntnis der Eigenvektoren $\underline{e}_p$ der Steifigkeitsmatrix A wäre also

$$V_l = \text{spann}\{e_{p_l}\} \quad l = 1, \dots, n$$

eine optimale, weil a-orthogonale Wahl der Teilräume. Wir erinnern uns: e_{p_l} überdeckt eine große Skala von Frequenzen. Umgekehrt könnten vielleicht Funktionen λ_l, die eine große Skala von Frequenzen überdecken, auf ein schnell konvergentes Verfahren führen. Jedenfalls haben die Knotenbasiselemente von S_j diese Eigenschaft *nicht*: alle $\lambda_p^{(j)}$ sind „hochfrequent" (Träger mit Durchmesser $\mathcal{O}(h_j)$).
Im Vergleich dazu kann man Knotenbasisfunktionen von groben Gittern als „niederfrequent" einstufen (siehe Abbildung 22).
Wir definieren daher die sogenannte *Multilevel–Knotenbasis* Λ,

$$\Lambda = \bigcup_{k=0}^{j} \Lambda_k = \{\lambda_l \mid l = 1, \dots, m_S\} , \qquad m_S = n_0 + \dots + n_j ,$$

als Vereinigung aller Knotenbasis–Funktionen von allen Verfeinerungsleveln. Die Numerierung $\lambda_l = \lambda_{p_l}^{(k_l)}$ erfolgt von fein nach grob, d.h. aus $l > l'$ folgt $k_l \leq k_{l'}$. Von unserem *rein heuristischen Standpunkt* aus überdeckt die entsprechende Zerlegung

$$S_j = \sum_{l=1}^{m_S} V_l, \qquad V_l = \text{spann}\{\lambda_l\}, \qquad l = 1, \dots, m_S, \tag{6.27}$$

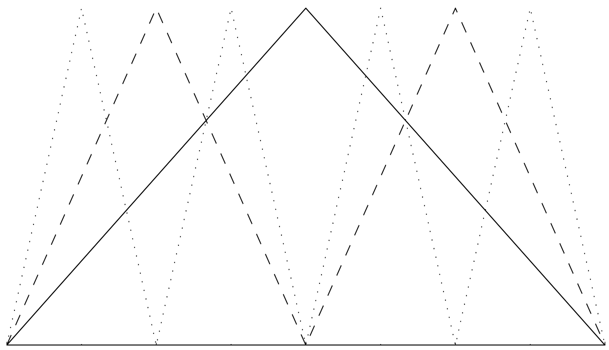


Abbildung 22: Multilevel–Knotenbasis als Skala von Frequenzen

also eine Skala von Frequenzen, die von niederfrequenten Funktionen $\lambda_l \in \Lambda_0$ bis zu hochfrequenten Funktionen $\lambda_l \in \Lambda_j$ reicht. Die erweiterte Zerlegung (6.26) erzeugt entsprechend Algorithmus 6.1 das sogenannte *Multilevel–Gauß–Seidel–Verfahren*

$$u_j^{\nu+1} = u_j^\nu + \sum_{l=1}^{m_S} v_l, \qquad v_l = \frac{l(\lambda_l) - a(w_{l-1}, \lambda_l)}{a(\lambda_l, \lambda_l)}\lambda_l, \tag{6.28}$$

wobei $w_0 = u_j^\nu$ und $w_l = w_{l-1} + v_l$, $l = 1, \ldots, m_S - 1$ gesetzt ist.

Mehrgitter–V–Zyklus. Das Verfahren 6.27 lässt sich äquivalent umschreiben, indem man jeweils alle Korrekturen auf Level k zu einer Korrektur $v^{(k)}$ zusammenfasst. Auf diese Weise erhält man das folgende *Mehrgitter–Verfahren.*

Algorithmus 6.2: Mehrgitter V–Zyklus

gegeben: u_j^ν
initialisiere: $r_j = \ell - a(u_j^\nu, \cdot), \quad a_j(\cdot,\cdot) = a(\cdot,\cdot)$

for $k = j, \dots, 1$ do:
$\{$
löse:
$v^{(k)} \in S_k : \quad b_k(v^{(k)}, v) = r_k(v) \quad \forall v \in S_k$ (Vorglättung)

$r_k = r_k - a_k(v^{(k)}, \cdot)$ (Aktualisierung des Residuums)

$r_{k-1} = r_k|_{S_{k-1}}$
$a_{k-1}(\cdot,\cdot) = a_k(\cdot,\cdot)|_{S_{k-1}\times S_{k-1}}$ (kanonische Restriktion)
$\}$

löse:
$v_0 \in S_0 : \quad b_0(v_0, v) = r_0(v) \quad \forall v \in S_0$ (näherungsweise Grobgitterlösung)

for $k = 1, \dots, j$ do:
$\{$
$v^{(k)} = v^{(k)} + v^{(k-1)}$ (kanonische Interpolation)
$\}$

neue Iterierte: $u_j^{\nu+1} = u_j^\nu + v^{(j)}$

Die Bilinearformen

$$b_k(v,w) = \sum_{\substack{i,l=1\\ i\le l}}^{n_k} v(p_i)a\big(\lambda_{p_i}^{(k)}, \lambda_{p_l}^{(k)}\big)w(p_l)\ , \qquad v,w \in S_k\ . \tag{6.29}$$

entsprechen einem Gauß–Seidel–Schritt auf Level k. Die Knoten $p_l \in \mathcal{N}_k$ sind dabei in der gleichen Reihenfolge zu durchlaufen wie die Teilräume $V_l = \text{spann}\{\lambda_{p_l}^{(k_l)}\}$ auf Level $k_l = k$ in der Formulierung (6.1). Da Gauß–Seidel–Schritte die hochfrequenten Anteile auf S_k schnell reduziert, spricht man von einem *Glätter*.
Die *kanonischen Restriktionen* des Residuums $r_k \in S_k'$ und der Bilinearform $a_k(\cdot,\cdot)$ auf S_j sind durch

$$r_{k-1}(v) = r_k(v), \quad a_{k-1}(w,v) = a_k(v,w), \qquad v,w \in S_{k-1} \subset S_k$$

definiert.
Die Bezeichnung *V–Zyklus* rührt daher, daß man zunächst vom feinen auf das grobe Gitter absteigend Korrekturen $v^{(k)}$ berechnet, die man anschließend aufsteigend vom groben auf das feine Gitter wieder einsammelt. Auf diese Weise kommt man wegen (6.17) in jedem Iterationschritt mit $\mathcal{O}(n_j)$ Punktoperationen aus.

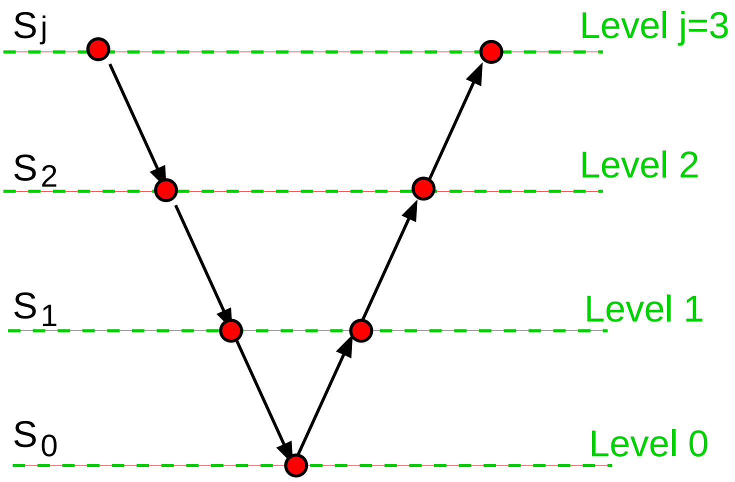


Abbildung 23: Ab– und Aufstieg durch die Gitter im V–Zyklus für $j = 3$.

Implementierung. Bevor man an die Implementierung des Verfahrens geht, hat man in der üblichen Weise Finite–Element–Funktionen $v \in S_k$ durch Vektoren $\underline{v} \in \mathbb{R}^{n_k}$ sowie Bilinearformen $a_k(\cdot,\cdot)$, $b_k(\cdot,\cdot)$ durch Matrizen $A_k, B_k \in \mathbb{R}^{n_k,n_k}$ auszudrücken. Residuen $r \in S_k'$ stellt man durch ihre Werte auf der Knotenbasis dar, also

$$\underline{r} = \left(r(\lambda_{p_i}^{(k)})\right)_{i=1}^{n_k} \in \mathbb{R}^{n_k} \quad .$$

Aus der kanonischen Interpolation von S_{k-1} nach S_k wird dann eine Matrix $I_k \in \mathbb{R}^{n_k,n_{k-1}}$ und die Restriktion der Residuen von S_k nach S_{k-1} lässt sich durch $R_k = I_k^T \in \mathbb{R}^{n_{k-1},n_k}$ darstellen. Allerdings stellt man diese Matrizen bei der praktischen Implementierung nicht explizit auf. Zur Interpolation nutzt man die Beziehung

$$\boxed{v^{(k-1)}(p) = \sum_{q\in\mathcal{N}_{k-1}} v^{(k-1)}(q)\lambda_q^{(k-1)}(p) \quad \forall p \in \mathcal{N}_k}$$

und zur Restriktion die Beziehung

$$\boxed{r_k(\lambda_p^{(k-1)}) = \sum_{q\in\mathcal{N}_k} \lambda_p^{(k-1)}(q) r_k(\lambda_q^{(k)}) \quad \forall p \in \mathcal{N}_{k-1} \quad .}$$

direkt aus. Auf diese Weise erhält man, ähnlich wie bei Differenzenverfahren, lokale Interpolations– und Restriktionssterne (Übung).

Konvergenz. Unsere heuristische Motivation wird durch folgendes Konvergenzresultat gerechtfertigt.

Theorem 6.24 *Es existiert ein $\varrho < 1$, das nur von der Regularität von $\mathcal{T}_0$ und den Elliptizitätskonstanten γ, Γ von $a(\cdot, \cdot)$ abhängt, sodaß für beliebiges $u_j^0 \in S_j$ die Fehlerabschätzung*

$$\boxed{\|u_j - u_j^{\nu+1}\| \leq \varrho \|u_j - u_j^{\nu}\| \qquad \forall \nu \geq 0} \tag{6.30}$$

besteht.

Ein Beweis würde den Rahmen dieser Vorlesung sprengen. Wir verweisen auf die gefeierten Übersichtsartikel von Xu [29] und Yserentant [30] oder (als Lesehilfe) auf das Skript von Kornhuber [16]. Eine sehr elegante Darstellung der klassischen Mehrgitterkonvergenztheorie gibt Braess [5][Kapitel V]. Das obige Resultat (unsymmetrisches Gauß–Seidel–Verfahren als Glätter) folgt aus Ergebnissen von Neuss [18].

Geschachtelte Iteration. Bei der praktischen Anwendung von Mehrgitterverfahren sind wir daran interessiert, die exakte Finite–Element–Approximation u_j in der Energienorm bis auf einen Fehler der Ordnung $\mathcal{O}(h_j)$ zu approximieren. Die dafür erforderliche Anzahl von Iterationsschritten hängt natürlich stark von der Startiterierten ab. Mehrgitter–Schritte auf feinen Gittern sind teurer als auf groben Gittern. Daher ist es naheliegend, durch eine gewisse Anzahl ν^* von Mehrgitter–Schritte auf $\mathcal{T}_k$ gute Startiterierte für die Iteration auf $\mathcal{T}_{k+1}$ zu berechnen. Diese Vorgehensweise heißt *geschachtelte Iteration* (siehe Hackbusch [12]).

Algorithmus 6.3: Mehrgitter mit geschachtelter Iteration

gegeben: $\tilde{u}_0 \in S_0$

for $k = 1, \ldots, j$ do:
{
Startiterierte: $u_k^0 = \tilde{u}_{k-1}$
ν^* Mehrgitterschritte: $\tilde{u}_k = u_k^{\nu^*}$
}

Endergebnis: $\tilde{u}_j$

Es stellt sich die Frage, wie man ν^* zu wählen hat. Die Antwort ergibt sich aus folgendem Satz.

Satz 6.25 *Wir setzen voraus, daß die Finite–Element–Näherungen* u_k *die Abschätzungen*

$$\|u - u_k\| \leq c_1 h_k \ , \qquad k = 0, 1, \ldots \ , \tag{6.31}$$

mit einer k–unabhängigen Zahl c_1 *erfüllen, daß weiter*

$$\tilde{u}_0 = u_0 \tag{6.32}$$

und daß schließlich ν^* *so gewählt ist, daß das Abbruchkriterium*

$$\|u_k - \tilde{u}_k\| \leq \tfrac{\sigma}{2} \|u_k - u_k^0\| \ , \qquad k = 1, 2, \ldots \ , \tag{6.33}$$

mit einem k–unabhängigen $\sigma < 1$ *erfüllt ist. Dann gibt es ein j–unabhängiges* C*, so daß*

$$\boxed{\|u - \tilde{u}_j\| \leq C h_j \ .} \tag{6.34}$$

Beweis:

Unter Verwendung der Voraussetzungen (6.32), (6.30), (6.31) und (6.16) berechnet man

$$\begin{aligned}
\|u - \tilde{u}_j\| &\leq \|u - u_j\| + \|u_j - \tilde{u}_j\| \leq \|u - u_j\| + \tfrac{1}{2}\sigma \|u_j - \tilde{u}_{j-1}\| \\
&\leq (1 + \tfrac{1}{2}\sigma)\|u - u_j\| + \tfrac{1}{2}\sigma \|u - \tilde{u}_{j-1}\| \\
&\leq (\tfrac{\sigma}{2})^j \|u - u_0\| + (1 + \tfrac{1}{2}\sigma) \sum_{i=0}^{j-1} (\tfrac{\sigma}{2})^i \|u - u_{j-i}\| \\
&\leq (1 + \tfrac{1}{2}\sigma) c_1 \sum_{i=0}^{j} \sigma^i h_j \\
&\leq \tfrac{3}{2} c_1 (1 - \sigma)^{-1} h_j \quad .
\end{aligned}$$

Es folgt (6.33) mit $C = \frac{5}{2} c_1 (1 - \sigma)^{-1}$.

□

Natürlich hängt im allgemeinen c_1 somit auch C von u ab. Bekanntlich ist die Regularitätsvoraussetzung $u \in H \cap H^2(\Omega)$ hinreichend für die Diskretisierungsgenauigkeit (6.30). Die exakte Lösung u_0 auf dem (hoffentlich) groben Gitter $\mathcal{T}_0$ berechnet man mit einem direkten Lösungsverfahren. Um das Abbruchkriterium (6.32) nachprüfen zu können, benötigen wir eine *a posteriori Schätzung des algebraischen Fehlers* $\|u_k - u_k^\nu\|$. Aus der Dreiecksungleichung folgt mit ϱ aus Theorem 6.24 (Übung)

$$(1 + \varrho)^{-1} \|u_k^{\nu+1} - u_k^\nu\| \leq \|u_k - u_k^\nu\| \leq (1 - \varrho)^{-1} \|u_k^{\nu+1} - u_k^\nu\| \ . \tag{6.35}$$

Die Mehrgitter–Korrekturen liefern also gleichzeitig untere und obere Schranken für den algebraischen Fehler.

Korollar 6.26 *Für die Berechnung einer Näherung $\tilde{u}_j$ mit Genauigkeit*

$$\|u - \tilde{u}_j\| = \mathcal{O}(h_j)$$

benötigt das Mehrgitterverfahren mit geschachtelter Iteration (siehe Theorem 6.25)

$$\boxed{\mathcal{O}(n_j) \text{ Punktoperationen.}}$$

Beweis:

Die Anzahl der Punktoperationen auf einem festen Level k ist durch cn_k beschränkt. Dabei ist c unabhängig von k. Da die Konvergenzrate ϱ unabhängig von j ist, ist auch die Anzahl ν^* von Mehrgitterschritten, die zum Erfüllen des Abbruchkriteriums (6.32) benötigt werden, unabhängig von j. Insbesondere kann man zu jedem vorgegebenen $\sigma < 1$ die benötigte Anzahl von Mehrgitter–Schritten ν^* aus $\varrho^{\nu^*} < \frac{1}{2}\sigma$ berechnen. Mit Blick auf (6.17) ist damit der Gesamtaufwand des Mehrgitterverfahrens mit gschachtelter Iteration beschränkt durch

$$c\nu^* \sum_{k=1}^{j} n_k \leq c\nu^*(1 - q^{-1})^{-1} n_j = \mathcal{O}(n_j) \ .$$

□

Bemerkung: Unser eigentliches Ziel ist die Berechnung der unbekannten Funktion u bis auf vorgegebene Genauigkeit TOL. Daher sollte die Anzahl j der Verfeinerungsschritte nicht a priori festgelegt, sondern im Rahmen eines adaptiven Multilevel–Verfahrens *a posteriori* bestimmt werden. Dazu sind entsprechende a posteriori Schätzungen des Diskretisierungsfehlers erforderlich.

Bemerkung: In der Praxis werden Mehrgitterverfahren meist als *Vorkonditionierer* für das *Verfahren der konjugierten Gradienten* (cg-Verfahren) eingesetzt. Auf diese Weise erhält man eine weitere Verbesserung der Konvergenzgeschwindigkeit. Für Einzelheiten zu cg–Verfahren und Vorkonditionierung verweisen wir beispielsweise auf Deuflhard und Hohmann [7][Kapitel 8] oder Braess [5][Kapitel IV].

7 Parabolische Differentialgleichungen

7.1 Klassische Lösungen

Zu den nachfolgenden zwei Abschnitten können u.a. bei John [14, Chapter 7] Einzelheiten nachgelesen werden.

7.1.1 Das Cauchy-Problem für die Wärmeleitungsgleichung

Wir betrachten die Wärmeleitungsgleichung

$$\begin{aligned} u_t &= \alpha u_{xx} \qquad && x \in \mathbb{R}\,,\ t \in (0,T] \\ u(x,0) &= u_0(x) && x \in \mathbb{R} \qquad , \end{aligned} \tag{7.1}$$

wobei zunächst $\alpha \equiv 1$ gesetzt wird. Gesucht wird eine *klassische* Lösung u, d.h.

$$u \in C(\overline{Q}) \quad , \quad u_{xx}, u_t \in C(Q) \quad \text{mit} \quad Q = \mathbb{R} \times (0,T] \quad .$$

Durch „qualifiziertes Raten“ wollen wir eine geschlossene Lösung von (7.1) ermitteln. Dazu definieren wir für $v \in C_0^\infty(\mathbb{R})$ die *Fouriertransformierte*

$$\boxed{\hat{v}(\xi) = (2\pi)^{-\frac{1}{2}} \int\limits_{\mathbb{R}} e^{-ix\xi} v(x)\, dx \quad \text{mit} \quad i^2 = -1 \quad .} \tag{7.2}$$

In einem etwas größeren Raum als $C_0^\infty(\mathbb{R})$ (einem Raum „stark fallender“ C^∞-Funktionen) stellt sich die Fouriertransformation $v \mapsto \hat{v}$ als Bijektion heraus, und es gilt die inverse Beziehung

$$v(x) = (2\pi)^{-\frac{1}{2}} \int\limits_{\mathbb{R}} e^{ix\xi} \hat{v}(\xi)\, d\xi \quad . \tag{7.3}$$

Eine fundamentale Eigenschaft der Fouriertransformation ist es, Ableitungen in Multiplikationen zu überführen. Es ist nämlich unter Verwendung der Produktregel (mit $|v(x)| \to 0$ für $|x| \to \infty$)

$$\begin{aligned} -i\xi\hat{v}(\xi) &= (2\pi)^{-\frac{1}{2}} \int\limits_{\mathbb{R}} \frac{\partial}{\partial x} e^{-ix\xi} v(x)\, dx \\ &= -(2\pi)^{-\frac{1}{2}} \int\limits_{\mathbb{R}} e^{-ix\xi} \frac{d}{dx} v(x)\, dx \quad , \end{aligned}$$

also

$$\boxed{-i\xi\hat{v} = \widehat{\tfrac{d}{dx}v} \quad .}$$

Mit dieser Eigenschaft läßt sich unser Problem (7.1) wesentlich vereinfachen und durch weitere Umformungen lösen. formal durchgeführt und erst nachträglich gerechtfertigt werden. Dazu wenden die Fouriertransformation auf den Ortsanteils von u in der Wärmeleitungsgleichung an und erhalten für jedes feste $t > 0$

$$\hat{u}_t = \widehat{u_t} = \widehat{u_{xx}} = (i\xi)^2\hat{u} = -\xi^2\hat{u} \quad . \tag{7.4}$$

Die Anfangsbedingung geht über in

$$\hat{u}(0) = \widehat{u_0} \quad .$$

Die Lösung dieser gewöhnlichen Differentialgleichung im Fourierraum ist gegeben durch

$$\hat{u}(\xi, t) = \widehat{u_0}(\xi)e^{-\xi^2 t} \quad . \tag{7.5}$$

Nun müssen wir zurücktransformieren. Dies könnte man mit mit der oben genannten Inversen der Fouriertransformation auch tun. (Äußerste Vorsicht allerdings bei der Herleitung der Lösungsdarstellung (1.10) bei John [14, Chapter 7], insbesondere bei der Durchführung der Variablentransformation.) Arbeitssparender ist es, die folgende Formel (siehe z.B. Werner [28]) zu verwenden, welche wieder für „stark fallende“ C^∞-Funktionen v und w zutrifft:

$$(2\pi)^{-\frac{1}{2}}\widehat{v * w} = \hat{v}\hat{w} \quad . \tag{7.6}$$

Dabei ist $v * w$ die Faltung von v und w, definiert durch

$$v * w(x) \int\limits_{\mathbb{R}} v(y)w(x-y)\,dy \quad \forall x \in \mathbb{R} \quad .$$

Die Fouriertransformation überführt also Faltungen in Produkte von Funktionen. Um (7.5) auf (7.4) anwenden zu können, müssen wir noch die Funktion $w : \xi \mapsto e^{-\xi^2 t}$ als eine Fouriertransformierte identifizieren. In der Tat ist diese Funktion bis auf eine Variablentransformation die Fouriertransformierte ihrer selbst. Es ist nämlich $x \mapsto e^{-x^2/2}$ eine Eigenfunktion der Fouriertransformation und daher w die Fouriertransformierte von

$$x \mapsto (2t)^{-\frac{1}{2}}e^{-x^2/4t} \tag{7.7}$$

(Beweis zur Übung). Daraus folgt nun mit (7.5) für unsere Lösung des Cauchyproblems

$$u(x,t) = (4\pi t)^{-\frac{1}{2}} \int\limits_{\mathbb{R}} e^{-\frac{(x-\xi)^2}{4t}} u_0(\xi)\,d\xi. \tag{7.8}$$

Man sieht, daß unsere Umformungen offenbar zumindest für die hier nicht näher definierten „stark fallenden“ C^∞-Anfangsdaten u_0 eine geschlossene Darstellung für eine Lösung u des Cauchyproblems (7.1) liefern. Wir werden nun zeigen, daß durch (7.7) sogar für jedes stetige und beschränkte u_0 eine Lösung gegeben ist, die noch dazu stetig von u_0 abhängt. Zu diesem Zweck definieren wir wie schon im elliptischen Fall die *Greensche Funktion* G durch

$$\boxed{G(x,\xi,t) = (4\pi t)^{-\frac{1}{2}}\, e^{-\frac{(x-\xi)^2}{4t}} \quad , \quad t>0\ ,\ x,\xi\in\mathbb{R} \quad .}$$

Satz 7.1 *Für jede Anfangsbedingung* $u_0 \in C(\mathbb{R})$ *mit* $\|u_0\|_\infty = \sup_{x\in\mathbb{R}} |u_0(x)| < \infty$ *definiert*

$$\boxed{u(x,t) = \int_{\mathbb{R}} G(x,\xi,t)u_0(\xi)\,d\xi \quad ,\ x\in\mathbb{R}\ ,\ t>0}$$

eine klassische Lösung des Cauchy-Problems (7.1). *Darüberhinaus gilt für jees* $T>0$

$$\max_{t\in[0,T]} \|u(\cdot,t)\|_\infty \le \|u_0\|_\infty \quad . \tag{7.9}$$

Beweis:
Es gilt

$$\begin{aligned}
\frac{\partial}{\partial t}G(x,\xi,t) &= \frac{1}{t\sqrt{4\pi t}}\, e^{-\frac{(x-\xi)^2}{4t}} \left[-\frac{1}{2} + \frac{(x-\xi)^2}{4t}\right] \quad , \\
\frac{\partial}{\partial x}G(x,\xi,t) &= -\frac{(x-\xi)}{2t\sqrt{4\pi t}}\, e^{-\frac{(x-\xi)^2}{4t}} \quad , \\
\frac{\partial^2}{\partial x^2}G(x,\xi,t) &= \frac{1}{t\sqrt{4\pi t}}\, e^{-\frac{(x-\xi)^2}{4t}} \left[\frac{(x-\xi)^2}{4t} - \frac{1}{2}\right] \quad ,
\end{aligned}$$

so daß

$$u_t(x,t) = \int_{\mathbb{R}} \frac{\partial}{\partial t}G(x,\xi,t)u_0(\xi)\,d\xi = \int_{\mathbb{R}} \frac{\partial^2}{\partial x^2}G(x,\xi,t)u_0(\xi)\,d\xi = u_{xx}(x,t) \quad .$$

Dabei kann man Integration und Differentiation vertauschen, da für $t>0$ die partiellen Ableitungen von G nach x bzw. t gleichgradig stetig bezüglich ξ sind und zudem gleichmäßig in einer Umgebung von x bzw. t für $\xi\to\infty$ genügend stark gegen 0 gehen. Konkret: Sei $\varepsilon>0$. Wir betrachten $\frac{\partial}{\partial t}f(t,\xi)$ für $f(t,\xi) = G(x,\xi,t)u_0(\xi)$, $x\in\mathbb{R}$, und ein $t>0$. Unter Verwendung des Mittelwertsatzes erhält man

$$\left|\frac{f(t+h,\xi)-f(t,\xi)}{h} - \frac{\partial}{\partial t}f(t,\xi)\right| \le \varepsilon,$$

falls $h\le\delta$ für ein $\delta>0$, welches unabhängig von ξ gewählt werden kann. Hieraus folgt

$$\frac{\partial}{\partial t}\int_{I_n} f(t,\xi)d\xi = \int_{I_n} \frac{\partial}{\partial t}f(t,\xi)d\xi$$

für jedes Intervall $I_n = [-n, n]$, $n \in \mathbb{N}$. (Warum?) Weiter sieht man für eine gewisse Umgebung U_t von t und ein $N \in \mathbb{N}$

$$\left| \int\limits_{\mathbb{R}\setminus I_n} \frac{\partial}{\partial t'} f(t', \xi) d\xi \right| \leq \varepsilon$$

für $n \geq N$ gleichmäßig in $t' \in U_t$. Da auch f uneigentlich integrierbar ist, Daraus kann man nun, da f ebenfalls uneigentlich integrierbar ist, die Vertauschbarkeit der partiellen Ableitung und des Integrals auch über dem unbeschränkten Integrationsbereich $\mathbb{R}$ schließen. (Wie?)
Jetzt untersuchen wir den Grenzwert von $u(\cdot, t)$ für $t \to 0$. Es gilt offenbar

$$x \neq \xi \quad \Longrightarrow \quad \lim_{t \to 0} G(x, \xi, t) = 0$$

gleichmäßig für $|x - \xi| \geq \delta > 0$ und

$$G(x, \xi, t) > 0 \quad \forall x, \xi \in \mathbb{R}\,,\ t > 0 \quad .$$

Zusammen liefert das für jedes $\delta > 0$

$$\left| \int\limits_{|x-\xi|\geq\delta} G(x, \xi, t) u_0(\xi)\, d\xi \right| \ \leq\ \|u_0\|_\infty \int\limits_{|x-\xi|\geq\delta} G(x, \xi, t)\, d\xi \ \to\ 0 \quad , \quad t \to 0 \quad .$$

Außerdem ist wegen

$$\int\limits_{\mathbb{R}} G(x, \xi, t)\, d\xi = 1 \qquad x \in \mathbb{R}\,,\ t > 0 \tag{7.10}$$

(Nachrechnen!) die folgende Abschätzung richtig:

$$\begin{aligned}
&\left| u_0(x) - \int\limits_{|x-\xi|\leq\delta} G(x, \xi, t) u_0(\xi)\, d\xi \right| \\
&\leq\ |u_0(x)| \int\limits_{|x-\xi|\geq\delta} G(x, \xi, t)\, d\xi \ +\ \int\limits_{|x-\xi|\leq\delta} G(x, \xi, t)\, |u_0(x) - u_0(\xi)|\ d\xi \\
&\leq\ \|u_0\|_\infty \int\limits_{|x-\xi|\geq\delta} G(x, \xi, t)\, d\xi \ +\ \max_{|x-\xi|\leq\delta} |u_0(x) - u_0(\xi)| \quad \to 0 \\
&\quad \text{für}\ \ t, \delta \to 0 \quad .
\end{aligned}$$

Da u_0 auf kompakten Teilmengen von $\mathbb{R}$ gleichmäßig stetig ist, folgt aus den obigen Abschätzungen die gleichmäßige Konvergenz von $u(\cdot, t) \to u_0$ für $t \to 0$ auf kompakten Teilmengen von $\mathbb{R}$ und damit $u \in C(\overline{Q})$. (Warum reicht die punktweise Konvergenz dafür nicht aus?)
Die a-priori-Abschätzung (7.8) folgt direkt aus der Mittelungseigenschaft (7.9).

□

Bemerkung:

1) Die Eigenschaft

$$G(x, \xi, t) > 0 \quad \forall x, \xi \in \mathbb{R}\,,\ t > 0$$

bedeutet, daß eine Veränderung von u_0 in einem beliebigen Punkt x_0 zu Änderungen der Lösung $u(x,t)$ in *jedem* $x \in \mathbb{R}$ zu *jeder* Zeit $t > 0$ bewirkt. Damit besteht das *Abhängigkeitsgebiet* von jedem $(x,t) \in Q$ aus ganz $\mathbb{R}$.

Anders ausgedrückt:

Störungen der Anfangsbedingungen breiten sich mit *unendlicher Geschwindigkeit* aus.

2) Unabhängig von der Wahl der Anfangsbedingung $u_0 \in C(\mathbb{R})$ ist $u(\cdot,t) \in C^\infty(\mathbb{R})$ für jedes $t > 0$. (Warum?) Diese Glättungseigenschaft ist typisch für parabolische Gleichungen.

3) u genügt sogar dem *Maximumprinzip*

$$\inf_{y\in\mathbb{R}} u_0(y) \le u(x,t) \le \sup_{y\in\mathbb{R}} u_0(y) \quad \forall x \in \mathbb{R}\ ,\ t > 0 \quad .$$

4) Die Lösung des Cauchy-Problems *ist nicht eindeutig*. Insbesondere gibt es "unphysikalische" Lösungen mit $|u(x,t)| \to \infty$ für $t \to 0$. Eine Konstruktion solcher Lösungen findet man z.B. bei John [14, p. 211].

   Hinreichende Bedingungen für die Eindeutigkeit sind zum Beispiel

$$\begin{array}{rcllll} |u(x,t)| & \le & M\, e^{\alpha x} & , & x \in \mathbb{R}\ ,\ 0 < t < T & \text{John, p. 217}\ , \\ u(x,t) & \ge & 0 & , & x \in \mathbb{R}\ ,\ 0 < t < T & \text{John, p. 222}\ . \end{array}$$

7.1.2 Anfangs-Randwertprobleme

Wir betrachten auf einem beschränkten Gebiet $\Omega \subset \mathbb{R}^n$ das Anfangs-Randwertproblem für die Wärmeleitungsgleichung

$$\begin{aligned} u_t - \Delta u &= f && \text{in } Q = \Omega \times (0,T] \\ u(x,0) &= u_0(x) && x \in \overline{\Omega} \\ u(x,t) &= g(t) && (x,t) \in \partial\Omega \times (0,T] \quad . \end{aligned} \tag{7.11}$$

Um $u \in C(\overline{\Omega})$ zu gewährleisten, müssen die Daten f, u_0 und g als stetig und zudem die Konsistenzbedingungen

$$u_0(x) = g(0) \quad \forall x \in \partial\Omega$$

vorausgesetzt werden. Entscheidendes Hilfsmittel zum Beweis von Eindeutigkeit und a-priori-Abschätzung ist folgendes Maximumprinzip.

Satz 7.2 *Es sei $u \in C(\overline{Q})$, $u_t, u_{xx} \in C(Q)$, und es gelte*

$$u_t - \Delta u \ \le\ 0 \quad \text{in } Q \quad .$$

Dann folgt

$$\max_{(x,t)\in\overline{Q}} u(x,t) \leq \max_{(x,t)\in\partial Q\setminus\Gamma_T} u(x,t) \quad ,$$

wobei $\Gamma_T := \{(x,t)\,|\, x\in\Omega \text{ und } t = T\}$.

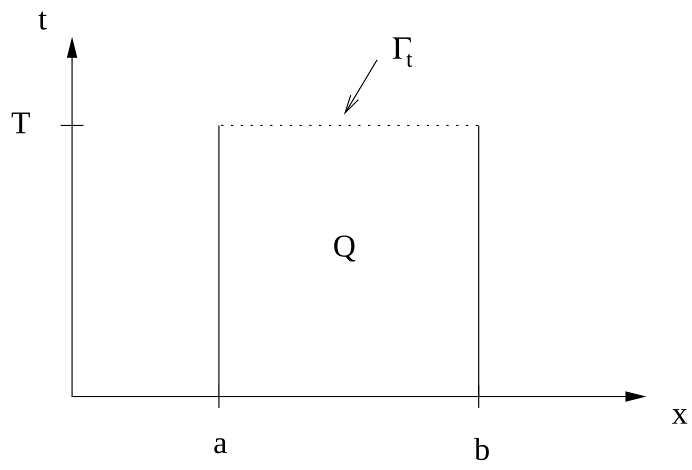


Abbildung 24: Rechengebiet Q

Beweis:
Wähle $\varepsilon > 0$ und betrachte die Funktion

$$v(x,t) = u(x,t) - \varepsilon\, t \quad .$$

Die Funktion v nehme in (x_0,t_0) ihr Maximum an. Wir zeigen $(x_0,t_0) \notin Q \cup \Gamma_T$.
Wäre das doch der Fall, müßte wegen der notwendigen Bedingungen für ein Maximum

$$v_t(x_0,t_0) \geq 0 \quad , \quad \Delta v(x_0,t_0) \leq 0$$

sein. Das steht im Widerspruch zu

$$v_t(x_0,t_0) = u_t(x_0,t_0) - \varepsilon \leq \Delta u(x_0,t_0) - \varepsilon = \Delta v(x_0,t_0) - \varepsilon \leq -\varepsilon < 0 \quad .$$

Würde nun u sein Maximum nicht in $\Gamma_D := \partial Q\setminus\Gamma_T$ annehmen, so gäbe es ein $(x_0,t_0) \in Q\cup\Gamma_T$ mit

$$u(x_0,t_0) > u(x,t) \qquad \forall (x,t) \in \Gamma_D \quad ,$$

und damit ein $\varepsilon > 0$ mit

$$u(x_0,t_0) - \varepsilon\, t_0 > u(x,t) - \varepsilon\, t \qquad \forall\,(x,t)\in\Gamma_D$$

im Widerspruch zu dem oben Gezeigten.

$\square$

Satz 7.3 *Das Anfangswertproblem* (7.11) *hat höchstens eine klassische Lösung* $u \in C(\overline{Q}) \cap C^2(Q)$.

Beweis:

Angenommen u_1, u_2 wären zwei verschiedene klassische Lösungen von (7.11), so wäre $w = u_1 - u_2$ die Lösung des homogenen Problems (mit $u_0 = 0$ und $g = 0$) und

$$w_t - \Delta w \leq 0 \quad .$$

Also muß gelten

$$\max_{(x,t)\in\overline{Q}} w(x,t) \leq \max_{(x,t)\in\Gamma_D} w(x,t) = 0 \quad .$$

Analog folgt aus $(-w)_t - (-w)_{xx} \leq 0$ sofort $-w \leq 0$. Damit ist $u_1 = u_2$ gezeigt.

□

Satz 7.4 *Eine klassische Lösung von* (7.11) *hängt, falls existent, bezüglich der* $\|\cdot\|_\infty$*-Norm stetig von den Daten* u_0 *und* g *ab. Es gilt sogar*

$$\sup_{0<t\leq T} \|u(\cdot,t)\|_\infty = \max\{\|u_0\|_\infty \, , \|g\|_\infty\} \quad .$$

Beweis:

Übung.

□

Den Nachweis der Existenz einer Lösung wollen wir für den homogenen Fall $f = 0$ auf $\Omega = (0,1)$ erbringen. Hierzu verwenden wir die *Fouriersche Methode*. Grundidee ist der Produktansatz

$$\boxed{u(x,t) = v(x)w(t) \quad .}$$

Einsetzen in die Differentialgleichung liefert

$$\frac{w'}{w} = \frac{v''}{v} = \text{const.} =: \lambda$$

und damit das *Eigenwertproblem*

$$\begin{aligned} v'' &= -\lambda v \\ w' &= -\lambda w \quad . \end{aligned} \tag{7.12}$$

Lösung von $v'' = -\lambda v$:

Ansatz: $v(x) = \sin(\lambda^{\frac{1}{2}} x)$
Der Ansatz erfüllt die Differentialgleichung, und mit $v(0) = v(1) = 0$ folgt

$$\lambda_k = (k\pi)^2 \quad , \quad k = 0, 1, 2, \ldots \quad .$$

Lösung von $w' = -\lambda_k w$:

$$w(t) = a_k e^{-\lambda_k t} \quad .$$

Lösung der Wärmeleitungsgleichung mit Nullrandbedingung:

$$u_k(x,t) = a_k e^{-(k\pi)^2 t} \sin(k\pi x) \qquad k = 0, 1, 2, \ldots \quad .$$

Wegen der Linearität der Differentialgleichung ist

$$u_N(x,t) = \sum_{k=0}^{N} u_k(x,t) = \sum_{k=0}^{N} a_k e^{-(k\pi)^2 t} \sin(k\pi x)$$

für jedes N eine Lösung, welche die Randbedingung erfüllt.
Stetige Annahme der Anfangsbedingung:
Bestimme die Koeffizienten a_k so, daß

$$u_0(x) = \sum_{k=0}^{\infty} a_k \sin(k\pi x) \quad .$$

Erweitert man u_0 zu einer ungeraden Funktion auf $[-1, 1]$ und entwickelt diese in ihre Fourierreihe, so erhält man gerade die *Fourierkoeffizienten*

$$a_k = 2 \int_0^1 u_0(\xi) \sin(k\pi\xi) \, d\xi \quad .$$

Ist u_0 stetig und stückweise stetig differenzierbar, so konvergiert die Fourierreihe punktweise gegen u_0 (siehe z.B. Endl/Luh: Analysis II, Satz 4.5.2).
Als Lösung der Differentialgleichung erwartet man nun

$$u(x,t) = \sum_{k=0}^{\infty} a_k e^{-(k\pi)^2 t} \sin(k\pi x) \quad . \tag{7.13}$$

Zur Rechtfertigung dieser Vorgehensweise haben wir folgenden Fragen nachzugehen.

(i) Konvergiert die Reihe (7.12) für jedes $(x,t) \in \Omega$?

(ii) Darf man gliedweise differenzieren? (Nur dann erfüllt u die Differentialgleichung.)

(iii) Wird die Anfangsbedingung in stetiger Weise angenommen?

Unter gewissen Voraussetzungen lassen sich alle drei Fragen mit „ja“ beantworten:

(i) Es gelte $\sup_{x\in(0,1)} |u_0(x)| \le M$. Dann folgt

$$|a_k| \le 2M \quad ,$$

und daher ist für jedes $t \ge \delta > 0$

$$\left| a_k e^{-(k\pi)^2 t} \sin(k\pi x) \right| \le 2M \left(e^{-\delta\pi^2}\right)^{k^2} \quad , \quad k = 0, 1, 2, \ldots \quad .$$

Das Majorantenkriterium liefert also sogar *gleichmäßige* Konvergenz für jedes feste $t > 0$.

(ii) Ähnlich sieht man, daß

$$2M(k\pi)^2 \left(e^{-\delta\pi^2}\right)^{k^2} \quad , \quad k = 0, 1, 2, \ldots$$

eine Majorante für die gliedweise nach t differenzierte Reihe liefert. Damit konvergiert die gliedweise differenzierte Reihe gleichmäßig und stimmt mit der entsprechenden Ableitung der Reihe überein (vgl. z.B. Endl/Luh, Analysis II, Satz 3.7.2). Analoges gilt für alle weiteren partiellen Ableitungen.

(iii) Es gelte

$$\sum_{k=1}^{\infty} |a_k| < \infty \quad . \tag{7.14}$$

Dann konvergiert nach dem Majorantenkriterium die Reihe

$$\sum_{k=1}^{\infty} a_k e^{-(k\pi)^2 t} \sin(k\pi x)$$

gleichmäßig in $x \in [0,1]$ und $t \ge 0$. Damit ist die Grenzfunktion stetig in $x \in [0,1]$, $t \ge 0$ (vgl. Endl/Luh Analysis II, Satz 3.7.1). Darüberhinaus können wir Grenzwert und Summation vertauschen und erhalten

$$\lim_{t\to 0} u(x,t) = \sum_{k=0}^{\infty} \lim_{t\to 0} a_k e^{-(k\pi)^2 t} \sin(k\pi x) = u_0(x) \qquad \forall x \in [0,1] \quad .$$

Hinreichend für die Bedingung (7.13) ist, daß u_0 stetig, stückweise differenzierbar und $u(0) = u(1) = 0$ ist (vgl. Heuser, Lehrbuch zur Analysis, Teil 2, 1981, Satz 136.5).

Wir haben (zusammen mit Satz 7.3) folgenden Existenzsatz bewiesen.

Satz 7.5 *Es sei u_0 stetig, stückweise differenzierbar und $u(0) = u(1) = 0$. Dann hat das Anfangs-Randwertproblem* (7.11) *eine eindeutig bestimmte klassische Lösung u gegeben durch*

$$\boxed{\sum_{k=1}^{\infty} a_k e^{-(k\pi)^2 t} \sin(k\pi x)}$$

mit den Fourierkoeffizienten

$$\boxed{a_k = 2 \int_0^1 u_0(\xi) \sin(k\pi\xi)\, d\xi \quad .}$$

Bemerkung: Setzt man die Formel für die Fourierkoeffizienten in der Reihe (7.12) ein, so kann man wegen der gleichmäßigen Konvergenz der entstehenden Reihe für $t > 0$ Integration und Summation vertauschen und erhält die Darstellung

$$u(x,t) = \int_0^1 G(x,\xi,t) u_0(\xi)\, d\xi$$

mit der Greenschen Funktion

$$G(x,\xi,t) = 2 \sum_{k=1}^{\infty} \sin(k\pi x) \sin(k\pi\xi) e^{-(k\pi)^2 t} \quad .$$

Beachte für $x \in (0,1)$

$$G(x,x,t) = 2 \sum_{k=1}^{\infty} (\sin(k\pi x))^2 e^{-(k\pi)^2 t} \to \infty, \quad t \to 0$$

und damit die Analogie zur Darstellung der Lösung des Cauchyproblems in Satz 7.1.

7.2 Schwache Lösungen

Für ein beschränktes Gebiet $\Omega \subset \mathbb{R}^n$ betrachten wir das Anfangs-Randwertproblem

$$\begin{array}{rcll} u_t &=& \operatorname{div}(\alpha(x)\nabla u) + f(x,t) & , \quad x \in \Omega\ ,\ t > 0 \\ u(x,0) &=& u_0(x) & , \quad x \in \Omega \\ u(x,t) &=& 0 & , \quad (x,t) \in \partial\Omega \times (0,T] \quad . \end{array} \tag{7.15}$$

Dabei sei $0 < \alpha_0 \leq \alpha(x) < \alpha_1$ fast überall in Ω vorausgesetzt. Wie schon bei elliptischen

Problemen können wir die Existenz klassischer Lösungen nicht erwarten, wenn $\alpha \notin C^1(\Omega)$ ist. Wir wollen daher den schwachen Lösungsbegriff von elliptischen (= stationäre parabolische Probleme) auf instationäre parabolische Probleme erweitern.
Wie schon im Falle elliptischer partieller Differentialgleichungen können wir auch hier eine schwache Formulierung finden, indem wir die Differentialgleichung (7.15) für festes $t > 0$ mit Testfunktionen $v \in H_0^1(\Omega)$ multiplizieren, anschließend integrieren, die Greensche Formel anwenden und die Ableitung nach t mit der Integration vertauschen.
Dann erhält man für jedes $t \in (0, T)$ das Variationsproblem:
Finde $u : \ (0, T) \to H_0^1(\Omega)$ so, daß

$$\frac{d}{dt}(u(t), v) + a(u(t), v) = (f(t), v) \qquad \forall v \in H_0^1(\Omega) \tag{7.16}$$

falls $t \in (0, T)$ und $u(0) = u_0$.
Als Abkürzungen wurden dabei

$$\begin{aligned} (u(t), v) &= \int_\Omega u(x,t)v(x)\,dx \\ a(v, w) &= \int_\Omega \alpha \nabla v \cdot \nabla w_x\,dx \\ (f(t), v) &= \int_\Omega f(x,t)v(x)\,dx \end{aligned}$$

verwendet. Um dieser Formulierung Sinn zu geben, müssen folgende Fragen geklärt werden.

1) Welche Abbildungen $u : \ (0, T) \to H_0^1(\Omega)$ kommen in Frage?

2) Wie soll die Anfangsbedingung angenommen werden?

3) Wie soll die Ableitung $\frac{d}{dt}$ interpretiert werden?

Wir gehen zunächst die erste Frage an und betrachten dazu den Raum $C([0, T], W)$ aller stetigen Funktionen v

$$[0, T] \ni t \mapsto v(t) \in W \quad .$$

Dabei sei W ein Hilbertraum mit Skalarprodukt $(\cdot, \cdot)_W$ und zugehöriger Norm $\|\cdot\|_W$. Versehen mit der Norm

$$\|v\|_{C([0,T],W)} := \max_{t\in[0,T]} \|v(t)\|_W$$

wird $C([0, T], W)$ ein Banachraum.

Beispiel: $W = \mathbb{R} \quad \Longrightarrow \quad C([0, T], \mathbb{R}) = C([0, T])$

Beispiel: $W = L^2(\Omega)$
Beachte, daß der resultierende Raum $C([0,T], L^2(\Omega))$ einerseits ein Raum stetiger Abbildungen ist, andererseits aber aus Funktionen $v(x,t)$ besteht, die in x-Richtung nicht stetig zu sein brauchen (Äquivalenzklassen!). Zum Beispiel ist

$$v(x,t) = v_1(x)v_2(t) \in C([0,T], L^2(\Omega)) \quad ,$$

falls $v_1(x) \in L^2(\Omega)$ und $v_2(t) \in C([0,T])$ gilt.

Mit dem Skalarprodukt

$$(v,w)_{L^2((0,T),W)} := \int_0^T (v(t), w(t))_W \, dt \tag{7.17}$$

wird der lineare Raum

$$L_C((0,T),W) := \{v \in C([0,T],W) \mid \int_0^T \|v(t)\|_W^2 \, dt < \infty\}$$

zum Prä-Hilbertraum. Die Vervollständigung bezüglich der durch (7.16) induzierten Norm bezeichnen wir mit

$$L^2((0,T),W) \quad .$$

Dabei haben wir es wie im Spezialfall

$$L^2((0,T),\mathbb{R}) = L^2(0,T)$$

wieder mit Äquivalenzklassen $[v]$ von Funktionen $v : (0,T) \to W$ zu tun, die sich auf Teilmengen von $(0,T)$ mit Lebesgue-Maß Null unterscheiden können.
Als Lösungsraum verwenden wir nun

$$C([0,T], L^2(\Omega)) \;\cap\; L^2((0,T), H_0^1(\Omega)) \quad .$$

Die Bedingung $u \in C([0,T], L^2(\Omega))$ sichert dabei, daß Anfangsbedingungen $u_0 \in L^2(\Omega)$ stetig angenommen werden können.

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

bedeutet nämlich

$$\lim_{t \to 0} \|u(\cdot,t) - u_0\|_{L^2(\Omega)} = 0 \quad . \tag{7.18}$$

Damit ist insbesondere Frage 2) geklärt.
Wir kommen zu Frage 3). Die Eigenschaft $u \in L^2((0,T), H_0^1(\Omega))$ liefert

$$(u(\cdot), v)\ ,\ a(u(\cdot), v) \in L^2(0,T) \qquad \forall v \in H_0^1(\Omega) \quad , \tag{7.19}$$

denn es gilt ja

$$\int_0^T a(u(t), v)^2\, dt \leq \alpha_1^2 \int_0^T \|u(t)\|_{H^1(\Omega)}^2\, \|v\|_{H^1(\Omega)}^2\, dt < \infty \quad .$$

Ebenso ist im Falle $f \in L^2((0,T), L^2(\Omega))$

$$(f(t), v) \ \in\ L^2(0,T)$$

für alle $v \in L^2(\Omega)$, denn

$$\int_0^T (f(t), v)^2\, dt \ \leq\ \int_0^T \|f(t)\|_{L^2(\Omega)}^2\, \|v\|_{L^2(\Omega)}^2\, dt < \infty$$

Die Zeitableitung $\frac{d}{dt}$ läßt sich daher in bekannter Weise im schwachen Sinn, nunmehr in $L^2(0,T)$, interpretieren.

Erinnerung:

$\frac{dw}{dt} \in L^2(0,T)$ heißt schwache Ableitung von $w \in L^2(0,T)$, falls

$$\int_0^T w(t)\varphi'(t)\, dt = -\int_0^T \frac{dw}{dt}\, \varphi(t)\, dt$$

für alle $\varphi \in C_0^\infty(0,T)$ gilt.

Die schwache Formulierung (7.15) von (7.15) lautet damit:

Finde $u \in L^2((0,T), H_0^1(\Omega))$, so daß

$$-\int_0^T (u(t), v)\varphi'(t)\, dt = \int_0^T \left(-a(u(t), v) + (f(t), v)\right) \varphi(t)\, dt \tag{7.20}$$

für alle $v \in H_0^1(\Omega)$ und alle $\varphi \in C_0^\infty(0,T)$ gilt.

Es sei von nun an $u_0 \in L^2(\Omega)$ und $f \in L^2((0,T), L^2(\Omega))$.

Definition 7.6 *$u \in C([0,T], L^2(\Omega)) \cap L^2((0,T), H_0^1(\Omega))$ heißt schwache Lösung des Anfangs-Randwertproblems (7.15), wenn u für fast alle $t \in (0,T)$ im Sinne von (7.19) der Variationsgleichung*

$$\frac{d}{dt}(u(t), v) + a(u(t), v) = (f(t), v) \qquad \forall v \in H_0^1(\Omega) \tag{7.21}$$

genügt und im Sinne von (7.17) die Anfangsbedingung $u(0) = u_0$ annimmt.

Als Vorbereitung für den Nachweis von Existenz und Eindeutigkeit brauchen wir

Satz 7.7 *Es existieren Funktionen $\varphi_k \in H_0^1(\Omega)$, $k = 1, 2, \ldots$, und zugehörige $\mu_k \in \mathbb{R}$ mit*

$$0 < \mu_1 \leq \mu_2 \leq \ldots \quad , \quad \lim_{k\to\infty} \mu_k = \infty \quad ,$$

so daß gilt

$$a(\varphi_k, v) = \mu_k(\varphi_k, v) \qquad \forall v \in H_0^1(\Omega) \tag{7.22}$$

und

$$\{\varphi_k\}_{k\in\mathbb{N}} \quad \text{ist Orthonormalbasis von } L^2(\Omega) \quad . \tag{7.23}$$

Beweis:

Da $a(\cdot,\cdot)$ stetig und $H_0^1(\Omega)$-elliptisch ist, hat das Variationsproblem

$$u \in H_0^1(\Omega) : \quad a(u,v) = (f,v) \quad \forall v \in H_0^1(\Omega)$$

für $f \in L^2(\Omega)$ nach Theorem 5.31 eine eindeutig bestimmte Lösung $Lu \in H_0^1(\Omega)$. Die Stabilitätsabschätzung liefert gerade die Stetigkeit des Lösungsoperators $L : L^2(\Omega) \to H_0^1(\Omega)$. Wir versehen nun $H_0^1(\Omega)$ mit dem neuen Skalarprodukt $a(\cdot,\cdot)$ und betrachten den Operator $T : H_0^1(\Omega) \to H_0^1(\Omega)$, definiert durch $Tv = Lv \ \forall v \in H_0^1(\Omega)$. Aus dem Rellichschen Einbettungssatz (siehe Satz 5.28) folgt jetzt mit der Stetigkeit von L die Kompaktheit von T. Aufgrund der Symmetrie bzw. $H_0^1(\Omega)$-Elliptizität von $a(\cdot,\cdot)$ ist T auf $(H_0^1(\Omega), a(\cdot,\cdot))$ selbstadjungiert und positiv (letzteres besagt $a(Tv, v) > 0 \ \forall v \in H_0^1(\Omega)$, $v \neq 0$). Der Spektralsatz für kompakte Operatoren in Hilberträumen (siehe Werner, Funktionalanalysis, 2. Auflage, Theorem VI.3.2) liefert nun eine $a(\cdot,\cdot)$-Orthonormalbasis von Eigenvektoren $\{\psi_k\}_{k\in\mathbb{N}}$ von T in $(H_0^1(\Omega), a(\cdot,\cdot))$ zu einer Nullfolge positiver Eigenwerte λ_k. Mit $\mu_k := \lambda_k^{-1}$ und $\varphi_k := \lambda_k^{1/2}\psi_k$ folgt die Behauptung.

□

Bemerkung:

1) Die Aussage von Satz 7.7 bleibt richtig, wenn wir $H_0^1(\Omega)$ durch einen anderen kompakt in $L^2(\Omega)$ eingebetteten Hilbertraum H mit einer H-elliptischen Bilinearform $a(\cdot,\cdot)$ ersetzen.

2) Aus (7.22) folgt insbesondere

$$\begin{aligned} v &= \textstyle\sum_{k=1}^{\infty}(v,\varphi_k)\varphi_k && \text{(Fourier-Entwicklung)} \\ \|v\|^2_{L^2(\Omega)} &= \textstyle\sum_{k=1}^{\infty}(v,\varphi_k)^2 && \text{(Parsevalsche Gleichung)} \end{aligned}$$

für alle $v \in L^2(\Omega)$.

3) Im Fall $\alpha \equiv 1$ und $\Omega = (0,1)$ haben wir

$$\begin{aligned} \varphi_k(x) &= \sqrt{2}\sin(k\pi x) \quad , \quad x \in \Omega = (0,1) \\ \mu_k &= (k\pi)^2 \end{aligned}$$

für $k = 1, 2, \ldots$ (vgl. [2], p. 228).

Man beachte auch die Eigenfunktionen und Eigenwerte des Laplaceoperators zu unserem Modellbeispiel auf $\Omega = (0,1)\times(0,1)$ (siehe Satz 6.19). Man kann übrigens auch allgemeineren Fällen die Eigenwerte μ_k mit einer Skala von Frequenzen der Eigenfunktionen in Verbindung bringen.

Nun können wir folgenden Darstellungssatz beweisen.

Satz 7.8 *Sei* $u \in C([0,T], L^2(\Omega)) \cap L^2((0,T), H_0^1(\Omega))$ *eine Lösung des Variationsproblems* (7.20). *Dann hat* u *die Darstellung*

$$u(t) = \sum_{k=1}^{\infty}\left((u_0,\varphi_k)e^{-\mu_k t} + \int_0^t (f(s),\varphi_k)e^{-\mu_k(t-s)}\,ds\right)\varphi_k \quad .$$

Beweis:

Sei also u eine Lösung. Dann ist insbesondere

$$u(t) \in L^2(\Omega) \qquad \forall t > 0$$

und daher

$$u(t) = \sum_{k=1}^{\infty}(u(t),\varphi_k)\,\varphi_k \qquad \forall t > 0 \quad . \tag{7.24}$$

Einsetzen in die Variationsgleichung ergibt unter Verwendung von (7.20) und (7.21) zusammen mit der Stetigkeit von $(\cdot,\cdot)$ und $a(\cdot,\cdot)$ in $L^2(\Omega)$

$$\frac{d}{dt}\sum_{k=1}^{\infty}(u(t),\varphi_k)(\varphi_k,v)+\sum_{k=1}^{\infty}(u(t),\varphi_k)\mu_k(\varphi_k,v)=(f(t),v)\qquad \forall v\in H_0^1(\Omega)\quad .$$

Im nächsten Schritt setzen wir $v=\varphi_k$, $k=1,2\ldots$, und erhalten die gewöhnlichen Differentialgleichungen

$$u_k'(t)+\mu_k u_k(t)=(f(t),\varphi_k) \tag{7.25}$$

für die (noch) unbekannten Fourierkoeffizienten

$$u_k(t)=(u(t),\varphi_k)\quad ,\quad k=1,2,\ldots\quad .$$

Die Anfangsbedingungen

$$u_k(0)=(u_0,\varphi_k) \tag{7.26}$$

folgen durch Fourier-Entwicklung von u_0 und Koeffizientenvergleich.
Die Lösung des Anfangswertproblems (7.24), (7.25) ist

$$u_k(t)=(u_0,\varphi_k)e^{-\mu_k t}+\int_0^t (f(s),\varphi_k)e^{-\mu_k(t-s)}\,ds\quad .$$

Diese Formel, die sich für C^1-Lösungen u_k leicht verifizieren läßt, kann man für stetige rechte Seiten in (7.24) durch *Variation der Konstanten* herleiten. Man beachte jedoch, daß das Anfangswertproblem hier im schwachen Sinne für beliebige rechte Seiten aus $L^2(0,T)$ zu lösen ist. Der Nachweis, daß (7.24), (7.25) auch in diesem Sinne eine eindeutige Lösung hat, die durch die genannte Formel gegeben ist, sei den Lesern zur Übung überlassen.

Schließlich liefert Einsetzen der Formel für die u_k in die Fourierentwicklung (7.23) die Behauptung.

□

Folgerung

Die Lösung des Variationsproblems ist eindeutig bestimmt, denn wir kennen ja (im Existenzfall!) die Fourierkoeffizienten.

Ziel unserer Untersuchungen ist der abschließende Existenzsatz.

Satz 7.9 *Es sei $u_0\in L^2(\Omega)$ und $f\in L^2((0,T),L^2(\Omega))$. Dann hat das Anfangsrandwertproblem* (7.15) *die eindeutig bestimmte Lösung $u\in C([0,T],L^2(\Omega))\cap L^2((0,T),H_0^1(\Omega))$ gegeben durch*

$$u(t)=\sum_{k=1}^{\infty}\left((u_0,\varphi_k)e^{-\mu_k t}+\int_0^t (f(s),\varphi_k)e^{-\mu_k(t-s)}\,ds\right)\varphi_k\quad . \tag{7.27}$$

Beweis:

1. Schritt
   Wir lösen das Problem für die Approximation
   $$u_{0,m} = \sum_{k=1}^{m} (u_0, \varphi_k)\varphi_k$$
   von u_0. Die Lösung u_m ist dann gegeben durch
   $$u_m(t) = \sum_{k=1}^{m} \left((u_0, \varphi_k) e^{-\mu_k t} + \int_0^t (f(s), \varphi_k) e^{-\mu_k (t-s)}\, ds \right) \varphi_k \quad . \tag{7.28}$$
   Das bestätigt man einfach durch Einsetzen unter Verwendung der Linearität der schwachen Ableitung. Wieder ist u_m als schwache Lösung des Anfangswertproblems aufzufassen.

2. Schritt
   Wir führen den Grenzübergang $m \to \infty$ durch, und zwar
   a) in $C([0,T], L^2(\Omega))$
   b) in $L^2((0,T), H_0^1(\Omega))$.

   a) Wir zeigen, daß u_m eine Cauchy-Folge in $C([0,T], L^2(\Omega))$ ist.
   Seien dazu $m < n$ beliebig, aber fest gewählt. Dann folgt aus der Parsevalschen Gleichung und der Cauchy-Schwarzschen Ungleichung für jedes $t \in [0,T]$
   $$\begin{aligned} \|u_n(t) - u_m(t)\|_{L^2(\Omega)} &\le \left(\sum_{k=m+1}^{n} (u_0, \varphi_k)^2 \right)^{\frac{1}{2}} + \left(\sum_{k=m+1}^{n} \left(\int_0^t (f(s), \varphi_k) e^{-\mu_k (t-s)}\, ds \right)^2 \right)^{\frac{1}{2}} \\ &\le \left(\sum_{k=m+1}^{n} (u_0, \varphi_k)^2 \right)^{\frac{1}{2}} + \left(\sum_{k=m+1}^{n} \frac{1}{2\mu_k} \int_0^T (f(s), \varphi_k)^2\, ds \right)^{\frac{1}{2}} \quad . \end{aligned}$$
   Für $n, m \to \infty$ geht die rechte Seite gegen 0 (und dies gleichmäßig in $t \in [0,T]$), denn es gilt ja
   $$\|u_0\|_{L^2(\Omega)} = \left(\sum_{k=1}^{\infty} (u_0, \varphi_k)^2 \right)^{\frac{1}{2}}$$
   und
   $$\begin{aligned} \|f\|_{L^2((0,T), L^2(\Omega))} &= \left(\int_0^T \|f(t)\|_{L^2(\Omega)}^2\, dt \right)^{\frac{1}{2}} \\ &= \left(\int_0^T \sum_{k=1}^{\infty} (f(t), \varphi_k)^2\, dt \right)^{\frac{1}{2}} \quad . \end{aligned}$$

Es ist dabei eine Anwendung des Lebesgueschen Integralsatzes fällig. Insgesamt ist damit u_m eine Cauchy-Folge in $C([0,T],L^2(\Omega))$, also

$$u_m \to u_1^* \quad \text{in} \quad C([0,T],L^2(\Omega)) \quad . \tag{7.29}$$

b) Wir zeigen, daß u_m eine Cauchy-Folge in $L^2((0,T),H_0^1(\Omega))$ ist.
Einsetzen der Darstellung (7.27) liefert zunächst

$$a(u_n(t)-u_m(t),u_n(t)-u_m(t)) = \sum_{k=m+1}^{n} \mu_k \left((u_0,\varphi_k)e^{-\mu_k t} + \int_0^t (f(s),\varphi_k)e^{-\mu_k(t-s)}\,ds \right)^2 \quad .$$

Verwenden wir nun die $H_0^1(\Omega)$-Elliptizität von $a(\cdot,\cdot)$ und $(a+b)^2 \leq 2(a^2+b^2)$, so ergibt sich für ein $c>0$ und jedes $t \in [0,T]$

$$\begin{aligned} \|u_n(t)-u_m(t)\|_{H^1(\Omega)}^2 &\leq \frac{1}{c}\,a(u_n(t)-u_m(t),u_n(t)-u_m(t)) \\ &\leq \frac{2}{c}\sum_{k=m+1}^{n}\mu_k \left((u_0,\varphi_k)^2 e^{-2\mu_k t} + \left(\int_0^t (f(s),\varphi_k)e^{-\mu_k(t-s)}\,ds \right)^2 \right) \quad . \end{aligned}$$

Wir haben diese Abschätzung in t über $[0,T]$ zu integrieren. Zur weiteren Abschätzung nach oben stellen wir

$$\mu_k \int_0^T e^{-2\mu_k t}\,dt = \frac{1}{2}(1-e^{-2\mu_k T}) < \frac{1}{2}$$

und

$$\begin{aligned} \mu_k \int_0^T \left(\int_0^t (f(s),\varphi_k)e^{-\mu_k(t-s)}\,ds \right)^2 dt &\leq \mu_k \int_0^T \int_0^t e^{-2\mu_k(t-s)}\,ds \int_0^t (f(s),\varphi_k)^2\,ds\,dt \\ &\leq \frac{T}{2}\int_0^T (f(t),\varphi_k)^2\,dt \end{aligned}$$

bereit. Daraus folgt

$$\int_0^T \|u_n(t)-u_m(t)\|_{H^1(\Omega)}^2\,dt \leq \frac{1}{c}\sum_{k=m+1}^{n}\left((u_0,\varphi_k)^2 + T\int_0^T (f(t),\varphi_k)^2\,dt \right) \quad .$$

Wie oben gilt, daß die rechte Seite dieser Abschätzung für $n,m \to \infty$ verschwindet. Damit ist u_m eine Cauchy-Folge in $L^2((0,T),H_0^1(\Omega))$, und aus der Vollständigkeit dieses Raums folgt die Konvergenz

$$u_m \to u_2^* \quad \text{in} \quad L^2((0,T),H_0^1(\Omega)) \quad . \tag{7.30}$$

3. Schritt ($u_1^* = u_2^*$)
   Aus (7.28) folgt

$$u_m \to u_1^* \quad \text{in} \quad L^2((0,T),L^2(\Omega)) \quad ,$$

und (7.29) liefert

$$u_m \to u_2^* \quad \text{in} \quad L^2((0,T),L^2(\Omega)) \quad .$$

Also muß $u_1^* = u_2^* = u^* \;\in\; L^2((0,T),H_0^1(\Omega)) \;\cap\; C([0,T],L^2(\Omega))$ sein.

4. Schritt (u^* ist Lösung)
   Aus der Konvergenz in $L^2((0,T),H_0^1(\Omega))$ folgt für fast alle $t \in (0,T)$ die Konvergenz

$$u_m(t) \to u^*(t) \quad \text{in} \quad H_0^1(\Omega) \quad .$$

   Daraus erhält man die Konvergenz von

$$a(u_m(t),v) \to a(u^*(t),v)$$

   für fast alle $t \in (0,T)$ und beliebiges $v \in H_0^1(\Omega)$). Schließlich folgt aus

$$\frac{d}{dt}(u_m(t),v) = (f(t),v) - a(u_m(t),v) \to (f(t),v) - a(u^*(t),v)$$

   und

$$(u_m(t),v) \to (u^*(t),v) \quad ,$$

   daß

$$\frac{d}{dt}(u_m(t),v) \to \frac{d}{dt}(u^*(t),v)$$

   gelten muß. Daher ist die Variationsgleichung

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

   für fast alle $t \in (0,T)$ und beliebiges $v \in H_0^1(\Omega)$) erfüllt.
   Zur Anfangsbedingung: Aus (7.28) erhält man

$$u_m(0) \to u^*(0) \quad \text{in} \quad L^2(\Omega) \quad .$$

   Andererseits ist aber

$$u_m(0) = \sum_{k=1}^{m} (u_0,\varphi_k)\varphi_k \to u_0 \quad \text{in} \quad L^2(\Omega) \quad .$$

   Also muß $u^*(0) = u_0$ sein. Damit ist $u = u^*$ die Lösung unseres Problems.

□

Aus der Lösungsdarstellung (7.26) folgt direkt die stetige Abhängigkeit der Lösung von den Daten u_0 und f. Dies ergibt sich nämlich aus der folgenden a-priori-Abschätzung.

Satz 7.10 *Es gilt für alle* $t \in [0,T]$

$$\|u(t)\|_{L^2(\Omega)} \le \|u_0\|_{L^2(\Omega)}\; e^{-\mu_1 t} + \int_0^t \|f(s)\|_{L^2(\Omega)}\; e^{-\mu_1 (t-s)}\, ds \quad .$$

Beweis:

Zur Übung.

□

8 Numerische Verfahren für parabolische Probleme

8.1 Zeitintegration

Wir betrachten das Anfangswertproblem

$$u \in C^1([0,T],\mathbb{R}) \ : \quad u' = \varphi(t,u) \quad \text{für} \quad t \in (0,T] \quad , \quad u(0) = u_0 \quad .$$

Zur Diskretisierung gehen wir der Einfachheit halber von einem äquidistanten Gitter

$$t_i \;=\; i\Delta t \qquad i = 0,\dots,N$$

aus. Wir betrachten das Einschrittverfahren

$$\frac{1}{\Delta t}(u_{i+1} - u_i) \;=\; \Theta\,\varphi(t_{i+1},u_{i+1}) \;+\; (1-\Theta)\,\varphi(t_i,u_i)$$

für $\Theta \in [0,1]$.
Anschaulich ersetzen wir die exakte Lösung zwischen u_i und u_{i+1} durch ein Gerade mit der gemittelten Steigung $\Theta u'(t_{i+1}) + (1-\Theta)u'(t_i)$.

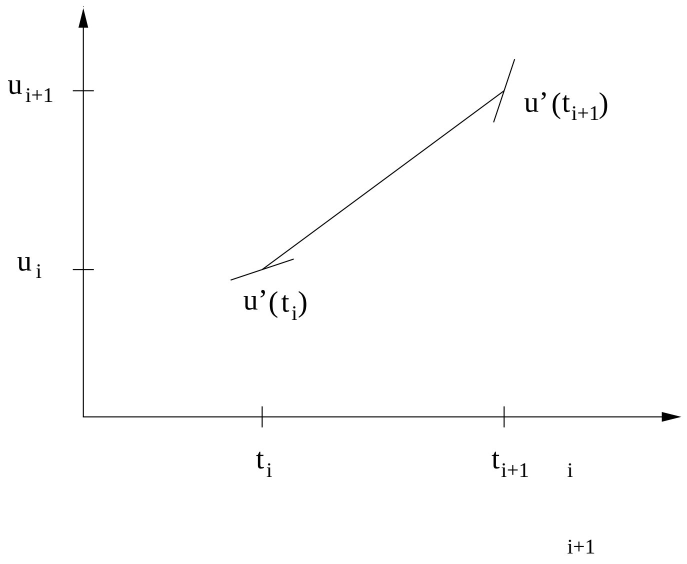


Abbildung 25: Integrationsschritt

Die Werte $\Theta = 0, \frac{1}{2}, 1$ sind von besonderer Bedeutung. Man erhält:

$\Theta = 0$	Eulersche Polygonzugmethode (explizites Eulerverfahren)
$\Theta = \frac{1}{2}$	Mittelpunktsregel
$\Theta = 1$	implizites Eulerverfahren

Man beache, daß für $\Theta = \frac{1}{2}, 1$ die Berechnung von u_{i+1} auf die Lösung einer (u. U. nichtlinearen) Gleichung hinausläuft. Solche Verfahren nennt man *implizit*. Für $\Theta = 0$ ist u_{i+1} *explizit* gegeben.
Wir untersuchen die *Konsistenzordnung* der Verfahren, welche man (wie schon bei Differenzenverfahren für elliptische Differentialgleichungen) durch Untersuchung des *Abschneidefehlers* erhält. Der Abschneidefehler ergibt sich durch Einsetzen der exakten Lösung u in die Differenzengleichung.
Wir erhalten durch Taylorentwicklung

$$\begin{aligned}
\tau(t_i) &= \frac{1}{\Delta t}(u(t_{i+1}) - u(t_i)) - \Theta\,\varphi(t_{i+1}, u_{i+1}) - (1-\Theta)\,\varphi(t_i, u_i) \\
&= u'(t_i) + \frac{\Delta t}{2} u''(t_i) + \mathcal{O}\left((\Delta t)^2\right) - (\Theta(u'(t_{i+1}) - u'(t_i)) + u'(t_i)) \\
&= (\frac{1}{2} - \Theta)\,\Delta t\, u''(t_i) + \mathcal{O}\left((\Delta t)^2\right) \\
&= \begin{cases} \mathcal{O}\,(\Delta t) & \Theta \neq \frac{1}{2} \\ \mathcal{O}\,((\Delta t)^2) & \Theta = \frac{1}{2} \end{cases} .
\end{aligned}$$

Damit ist das Verfahren von erster Ordnung für $\Theta \neq \frac{1}{2}$ und von zweiter Ordnung für $\Theta = \frac{1}{2}$.
Wir betrachten die *Stabilität* der Verfahren. Entsprechend dem Stabilitätskonzept von Dahlquist (vgl. [8, Kap. 6, 7]) betrachten wir das Anfangswertproblem

$$u' = -\lambda u \quad t > 0\,,\ \lambda > 0 \quad \text{mit} \quad u(0) = u_0 \tag{8.1}$$

mit der Lösung $u(t) = u_0\, e^{-\lambda t}$ und verlangen, daß die entsprechenden Näherungen wie u abklingen oder zumindest nicht wachsen. Die Anwendung unserer Diskretisierung liefert

$$\frac{1}{\Delta t}(u_{i+1} - u_i) = \Theta(-\lambda u_{i+1}) + (1-\Theta)(-\lambda u_i) \quad ,$$

also

$$\begin{aligned}
u_{i+1} &= \frac{1 - (1-\Theta)\,\Delta t\,\lambda}{1 + \Theta\,\Delta t\,\lambda}\, u_i \\
u_{i+1} &= \left(\frac{1 - (1-\Theta)\,\Delta t\,\lambda}{1 + \Theta\,\Delta t\,\lambda}\right)^i u_0 \quad .
\end{aligned}$$

Beschränktheit von $|u_i|$ ist also äquivalent mit

$$|R(\lambda\,\Delta t)| := \left|\frac{1 - (1-\Theta)\,\Delta t\,\lambda}{1 + \Theta\,\Delta t\,\lambda}\right| \leq 1 \quad , \tag{8.2}$$

und das „$<$“-Zeichen sorgt für Abklingen.

Im Falle $0 \le \Theta < \frac{1}{2}$ ist (8.1) äquivalent zu der Schrittweitenbeschränkung

$$\lambda \, \Delta t \le \frac{2}{1 - 2\Theta} \quad , \quad 0 \le \Theta < \frac{1}{2} \quad ,$$

wohingegen (8.1) für $\frac{1}{2} \le \Theta \le 1$ unbedingt erfüllt ist. Damit ergibt sich

$\Theta = 0$	explizites Eulerverfahren stabil, falls $\Delta t \le \frac{2}{\lambda}$
$\Theta = \frac{1}{2}, 1$	Mittelpunktsregel, implizites Eulerverfahren unbedingt stabil

Zwischen $\Theta = \frac{1}{2}$ und $\Theta = 1$ gibt es noch einen kleinen Unterschied. Es gilt offenbar

$$\lim_{\Delta t \to 0} R(\Delta t \, \lambda) = |1 - \Theta^{-1}| \quad \begin{cases} < 1 & \text{falls } \Theta > \frac{1}{2} \\ \\ \ge 1 & \text{falls } \Theta \le \frac{1}{2} \end{cases} \quad .$$

Damit wirkt das implizite Euler-Verfahren auch für beliebig große Zeitschritte noch dämpfend! Solche Verfahren heißen *stark stabil*. Die Mittelpunktsregel hat diese Eigenschaft nicht und ist damit, wie sich zeigen wird, für Rechnungen über lange Zeitintervalle, bei denen sich viele Fehler akkumulieren können, nicht so vorteilhaft.

8.2 Semidiskrete Verfahren

8.2.1 Linienmethode

Analog zum Galerkin-Verfahren für elliptische Probleme wählen wir nun eine Folge endlichdimensionaler Teilräume

$$S_h \subset H_0^1(\Omega) \quad .$$

Der Diskretisierungsparameter $h = h_j$, $j = 0, 1, \ldots$, steht dabei zum Beispiel für den maximalen Durchmesser aller Dreiecke einer zugrundeliegenden Triangulierung.
Wir wollen nun für alle $t \in [0, T]$ die kontinuierliche Lösung $u(t)$ durch ein $u_h(t) \in S_h$ approximieren. Das entspricht einer Diskretisierung im Ort. In Analogie zum kontinuierlichen Problem formulieren wir nun das entsprechende semidiskrete Problem:
Finde $u_h \in C([0, T], S_h)$, so daß

$$\begin{aligned} \frac{d}{dt}(u_h(t), v) + a(u_h(t), v) &= (f(t), v) \qquad \forall v \in S_h \\ u_h(0) &= u_{0,h} \in S_h \quad . \end{aligned} \tag{8.3}$$

Beachte, daß die Normen $\|\cdot\|_{L^2(\Omega)}$ und $\|\cdot\|_{H^1(\Omega)}$ auf dem endlichdimensionalen Raum S_h äquivalent sind, d.h. es gilt mit einem $c > 0$

$$c \, \|v\|_{H^1(\Omega)} \le \|v\|_{L^2(\Omega)} \le \|v\|_{H^1(\Omega)} \qquad \forall v \in S_h \quad .$$

Daher ist

$$C([0,T],S_{h,L^2(\Omega)}) = C([0,T],S_{h,H^1(\Omega)}) \subset L^2((0,T),S_{h,H^1(\Omega)}) \quad ,$$

wobei jeweils die Norm im Index in S_h zugrunde gelegt werden soll. Im Gegensatz zum kontinuierlichen Fall (siehe Satz 7.9) ist damit der Raum $C([0,T],S_h)$ bereits der Lösungsraum für das Problem (8.3).

Lemma 8.1 *Es existieren Funktionen* $\varphi_{kh} \in S_h$, $k = 1,\dots,N$, *und zugehörige* $\mu_{k,h} \in \mathbb{R}$ *mit*

$$0 < \mu_1 \le \mu_{1,h} \le \mu_{2,h} \le \ldots \le \mu_{N,h}$$

(*wobei* μ_1 *aus Satz 7.7 stammt*), *so daß gilt*

$$\begin{aligned} a(\varphi_{k,h},v) &= \mu_{k,h}(\varphi_{k,h},v) \quad \forall v \in S_h \quad , \\ \varphi_{k,h} &\text{ ist Orthonormalbasis von } S_{h,L^2(\Omega)} \quad . \end{aligned}$$

Beweis:

Siehe Satz 7.7 und anschließende Bemerkung 1). Zur Behauptung $\mu_1 \le \mu_{1,h}$ siehe [21, p. 146].

□

Satz 8.2 *Das semidiskrete Problem* (8.3) *hat eine eindeutig bestimme Lösung* u_h, *und es gilt*

$$u_h(t) = \sum_{k=1}^{N} \left((u_{0,h},\varphi_{k,h})e^{-\mu_{k,h}t} + \int_0^t (f(s),\varphi_{k,h})e^{-\mu_{k,h}(t-s)}\,ds \right) \varphi_{k,h} \quad .$$

Beweis:

Man übertrage den Beweis zu Satz 7.9.

□

Wir wollen die Konvergenz von u_h gegen u zeigen. Ein wesentlicher Schritt in diese Richtung ist die nachfolgende a-priori-Fehlerabschätzung. Hierfür definieren wir noch $C^1([0,T],H_0^1(\Omega))$ als den Raum aller $u \in C([0,T],H_0^1(\Omega))$, deren schwache Ableitung $\frac{d}{dt}u(t) = u'(t)$ für jedes $t \in [0,T]$ als Grenzwert

$$\lim_{\Delta t \to 0} \frac{u(t+\Delta t) - u(t)}{\Delta t}$$

in der $H^1(\Omega)$-Norm auftritt und in $C([0,T],H_0^1(\Omega))$ liegt. Damit ist für festes v die in (8.3) auftretende Ableitung sogar klassisch (siehe (8.3) im nachfolgenden Beweis).

Satz 8.3 *Es sei $u \in C^1([0,T], H_0^1(\Omega))$. Dann gilt für alle $t \in [0,T]$*

$$\begin{aligned}\|u_h(t) - u(t)\|_{L^2(\Omega)} &\leq \|u_{0,h} - P_h u_0\|_{L^2(\Omega)}\, e^{-\mu_1 t} + \|(I - P_h)u(t)\|_{L^2(\Omega)} \\ &\quad + \int_0^t \left\|(I - P_h)\frac{du}{dt}(s)\right\|_{L^2(\Omega)} e^{-\mu_1 (t-s)}\, ds \quad .\end{aligned}$$

Dabei ist I die Identität auf $H_0^1(\Omega)$ und $P_h : H_0^1(\Omega) \to S_h$ die Ritz-Projektion, also

$$P_h w \in S_h : \quad a(P_h w, v) = a(w, v) \qquad \forall v \in S_h$$

für alle $w \in H_0^1(\Omega)$.

Beweis:

1. Schritt: (Defektproblem)
Da u und u_h für alle $S_h \subset H_0^1(\Omega)$ die Variationsgleichung in (8.3) erfüllen, folgt nach Definition von P_h

$$\frac{d}{dt}(u_h(t) - P_h\, u(t), v) + a(u_h(t) - P_h\, u(t), v) = \frac{d}{dt}(u(t) - P_h\, u(t), v) \quad \forall v \in S_h \quad .$$

Aus der $H_0^1(\Omega)$-Elliptizität von $a(\cdot,\cdot)$ folgt die Stetigkeit von P_h, da P_h eine $a(\cdot,\cdot)$-Orthogonalprojektion ist. Weiter gilt wegen $u \in C^1([0,T], H_0^1(\Omega))$ gerade

$$\begin{aligned}\left\|P_h\, u(t+\Delta t) - P_h\, u(t) - \Delta t\, P_h \frac{du}{dt}(t)\right\|_{H^1(\Omega)} &\leq \text{const.} \left\|u(t+\Delta t) - u(t) - \Delta t\, \frac{du}{dt}(t)\right\|_{H^1(\Omega)} \\ &= o(|\Delta t|)\end{aligned}$$

Damit haben wir $P_h u \in C^1([0,T], H_0^1(\Omega))$ mit

$$\frac{d}{dt} P_h u = P_h \frac{du}{dt}$$

nachgewiesen. Einsetzen liefert

$$\frac{d}{dt}(u_h(t) - P_h\, u(t), v) + a(u_h(t) - P_h\, u(t), v) \;=\; (\frac{du}{dt} - P_h \frac{du}{dt}, v) \quad \forall v \in S_h \quad . \tag{8.4}$$

(Wie rechtfertigt man $\frac{d}{dt}(u(t), v) = (\frac{du}{dt}, v)$?)
Weiter gilt natürlich

$$u_h(0) - P_h u(0) = u_{0,h} - P_h u_0 \quad . \tag{8.5}$$

2. Schritt (a-priori-Abschätzung)
$u_h(t) - P_h u(t)$ ist Lösung des Defektproblems (8.3), (8.4) und genügt daher der a-priori-Abschätzung aus Satz 7.10. Dabei hat man zunächst statt μ_1 den diskreten Eigenwert $\mu_{1,h}$ einzusetzen, und aus $\mu_1 \leq \mu_{1,h}$ folgt dann die Behauptung.

□

Nun können wir den angestrebten Konvergenzsatz beweisen.

Satz 8.4 *Es gelte $u \in C^1([0,T], H_0^1(\Omega))$ und für alle $v \in H_0^1(\Omega)$*

$$\inf_{v_h \in S_h} \|v - v_h\|_{H^1(\Omega)} \to 0$$

für alle $v \in H_0^1(\Omega)$ sowie

$$\|u_{0,h} - u_0\|_{L^2(\Omega)} \to 0 \quad .$$

Dann folgt

$$u_h \to u \qquad \text{in} \ \ C([0,T], L^2(\Omega)) \quad .$$

Beweis:

1. Schritt
Wir betrachten die Folge $h_i \to 0$. Für ein festes $v \in C([0,T], H_0^1(\Omega))$ setzen wir

$$g_i := \|(I - P_{h_i})\, v(\cdot)\|_{H^1(\Omega)} \ \in \ C([0,T]) \quad .$$

Wir wollen nun zeigen, daß

$$\max_{t \in [0,T]} g_i(t) \to 0 \quad \text{für} \quad i \to \infty \quad . \tag{8.6}$$

Das Céa-Lemma 5.16 liefert für jedes feste $t \in [0,T]$

$$\|(I - P_h)\, v(t)\|_{H^1(\Omega)} \leq \frac{\Gamma}{\gamma} \inf_{v_h \in S_h} \|v(t) - v_h\|_{H^1(\Omega)}$$

also nach Voraussetzung

$$g_i(t) \to 0 \qquad \forall t \in [0,T] \quad .$$

Da alle P_h $a(\cdot,\cdot)$-Orthogonalprojektionen sind, gilt die Abschätzung

$$\|P_h v(t)\|_{H^1(\Omega)} \ \leq \ \frac{\Gamma}{\gamma}\ \|v(t)\|_{H^1(\Omega)}$$

gleichmäßig in h. Dies liefert zusammen mit der Stetigkeit von $v : [0,T] \to H_0^1(\Omega)$ die gleichgradige Stetigkeit der Menge $\{g_i \,|\, i \in \mathbb{N}\}$, d. h. zu jedem $t^* \in [0,T]$ und jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß

$$|t - t^*| < \delta \Longrightarrow |g_i(t) - g_i(t^*)| < \varepsilon \qquad \forall i \in \mathbb{N} \quad .$$

Nun folgt (8.5) durch einen einfachen Widerspruchsbeweis oder direkt aus dem Satz von Arzelà-Ascoli.

2. Schritt
Wir verwenden das Resultat (8.5) im Falle $v = u$ und $v = \frac{du}{dt}$ und erhalten

$$\begin{aligned} \|u_{0,h} - P_h u_0\|_{L^2(\Omega)} &\leq \|u_{0,h} - u_0\|_{L^2(\Omega)} + \|(I - P_h)u(0)\|_{L^2(\Omega)} && \to 0 \\ \max\nolimits_{t \in [0,T]} \|(I - P_h)u(t)\|_{L^2(\Omega)} &\leq \max\nolimits_{t \in [0,T]} \|(I - P_h)u(t)\|_{H^1(\Omega)} && \to 0 \end{aligned}$$

sowie

$$\max_{t \in [0,T]} \int_0^t \left\|(I - P_h)\frac{du}{dt}(s)\right\|_{L^2(\Omega)} e^{-\mu_1(t-s)}\, ds \ \leq \ \frac{1}{\mu_1} \max_{s \in [0,T]} \left\|(I - P_h)\frac{du}{dt}(s)\right\|_{L^2(\Omega)} \to 0 \quad .$$

Nun folgt die Behauptung aus der a-priori-Fehlerabschätzung in Satz 8.3.

□

Das semidiskrete Problem (8.3) ist ein Anfangswertproblem für ein System gewöhnlicher Differentialgleichungen. Wählt man nämlich eine Basis $\{\lambda_i\ ,\ i = 1,\dots,N\}$ von S_h, so führt die Darstellung

$$u_h(t) = \sum_{i=1}^{N} u_i(t)\lambda_i$$

auf

$$\begin{aligned} M\dot{U} + AU &= F \qquad t \in (0,T) \\ U(0) &= U_0 \quad . \end{aligned} \tag{8.7}$$

Dabei ist

$$\begin{aligned}
M &= ((\lambda_i,\lambda_j))_{i,j=1}^N && \text{Massenmatrix} \\
A &= (a(\lambda_i,\lambda_j))_{i,j=1}^N && \text{Steifigkeitsmatrix} \\
F &= ((f(t),\lambda_i))_{i=1}^N && \text{rechte Seite} \\
U &= (u_i(t))_{i=1}^N && \text{Unbekanntenvektor} \\
U_0 &= (u_i(0))_{i=1}^N && u_{0,h} = \textstyle\sum_{i=1}^N u_i(0)\lambda_i \quad .
\end{aligned}$$

Zur Lösung von Problemen der Form (8.7) stehen eine Vielzahl von Verfahren (und Codes) zur Verfügung. Es sei jedoch darauf hingewiesen, daß es sich wegen der stark unterschiedlich abklingenden Komponenten der Lösung (beachte die Darstellung in Satz 8.2) um ein *steifes System* handelt, bei dem gewisse Stabilitätsprobleme auftauchen.

Beispiel: Man betrachte das Problem (8.3) mit dem Raum S_h der linearen finiten Elemente auf $\Omega = (0,1)$ und $h = (N+1)^{-1}$ für ein $N \in \mathbb{N}$. Stellt man dafür das System in (8.7) auf und vereinigt die h-Abhängigkeit in der Steifigkeitsmatrix, indem man durch h teilt, so erhält man die $N \times N$-Tridiagonalmatrizen

$$\frac{1}{h}M = \frac{1}{6}\begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & & \ddots & & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{pmatrix} \quad \text{und} \quad \frac{1}{h}A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & & \ddots & & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix} \quad .$$

Die Matrix M/h wird in der Praxis durch sogenanntes „Lumping“ zur Einheitsmatrix gemacht: Man addiert im wesentlichen einfach die Zeileneinträge und schreibt die Summe in die Diagonale. Dies ist so, als ob man zur Berechnung der Einträge von M die Trapezregel als Quadraturformel verwenden würde. Damit erspart man sich das Bilden von M^{-1}, und erstaunlicherweise büßt man durch diese Vereinfachung nichts an Approximationsgüte ein (siehe [26, Kap. 15])! Verständlich wird dies durch die Tatsache, daß die Massenmatrix M zu einer Korrelation der Zeitableitungen an verschiedenen Punkten führt – ein Phänomen, das nur durch die Diskretisierung auftritt und im kontinuierlichen Problem gar nicht vorhanden ist. In diesem Sinne kann man Lumping sogar als eine Richtigstellung der Diskrepanz zwischen dem diskreten und kontinuierlichen Problem ansehen.

Die Matrix A/h^2 hat (siehe den Beweis zu Satz 6.19) die Eigenwerte $\lambda_i = 4h^{-2}\sin^2(i\frac{\pi}{2}h)$, $i = 1, \ldots, N$, also $\lambda_i \approx (i\frac{\pi}{2})^2$ für kleine h, d.h. man hat es bei feiner werdender Ortsdiskretisierung mit einer immer gößeren Bandbreite von Eigenwerten zu tun. Eine numerische Behandlung des entstandenen Anfangswertproblems mit expliziten Verfahren würde nun, ähnlich wie in Abschnitt 8.1, eine Schrittweitenbeschränkung in der Zeit nötig machen. Konkret hätte man, da λ_N in der Größenordnung h^{-2} liegt, die Beschränkung

$$\Delta t \leq \frac{1}{2(1-\Theta)}\, h^2$$

zu beachten. Man ist also bei solchen Problemen gezwungen, implizite Verfahren zu verwenden, für die diese Stabilitätsanforderung nicht erfüllt sein muß (das ist hier sogar nur für das voll implizite Verfahren mit $\Theta = 1$ der Fall).

Daß für $\Theta = 0$ (explizit) eine Zeitschrittbeschränkung *notwendig* ist, zeigt die Untersuchung des *numerischen Abhängigkeitsgebiets*.

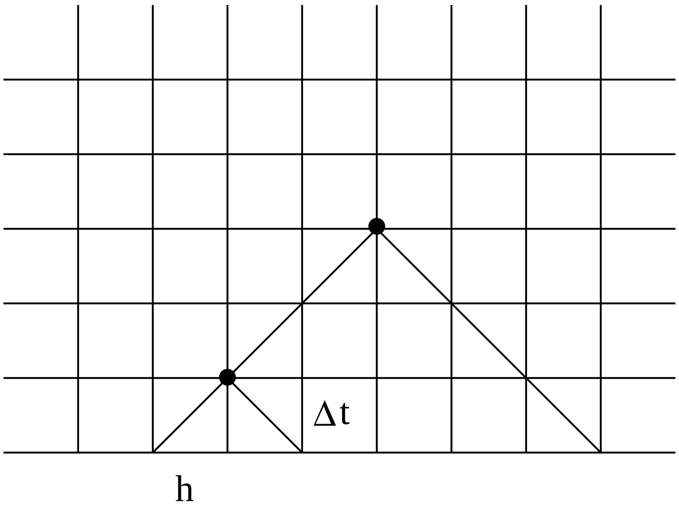


Abbildung 26: Numerisches Abhängigkeitsgebiet

Man hätte nach dem ersten Zeitschritt als eine Näherung von $(u_i(\Delta t))_{i=1}^N$ den Vektor

$U_1 = (U_{i,1})_{i=1}^N$ mit

$$U_1 = U_0 - (\Delta t/h^2)AU_0 + \Delta t\, F(0)$$

zu berechnen. Während jedoch, wie in Abschnitt 7.1.2 gezeigt (siehe die letzte Bemerkung: Greensche Funktion!), das Abhängigkeitsgebiet von $u(x_i, \Delta t)$ aus dem ganzen Intervall $(0, 1)$ besteht, ist das numerische Abhängigkeitsgebiet von $U_{i,1}$ nur $u_{i-1}(0)$, $u_i(0)$ und $u_{i+1}(0)$. Damit sich dieser Kegel asymptotisch aufweitet, muß $\frac{\Delta t}{h} \to 0$ gehen, das heißt Δt muß schneller verschwinden als h. Andernfalls könnte man ja U_0 außerhalb von $x_i \pm h$ ($x_i = i/N$) verändern, ohne daß $U_{i,1}$ als Approximation von $u(x_i, \Delta t)$ davon etwas merken würde.

<u>Stabilitätsregel</u> (Courant, Friedrichs, Levi):
Das numerische Abhängigkeitsgebiet muß das kontinuierliche approximieren.

Diese CFL-Bedingung ist eine wichtige Faustregel zur Einschätzung der Stabilitätseigenschaften von numerischen Verfahren.

Beim voll impliziten Verfahren kommt man ohne Zeitschrittbegrenzung aus. Das numerische Abhängigkeitsgebiet ist in diesem Fall aber auch das ganze vorherige Ortsgitter. Man muß nämlich dann das Gleichungssystem

$$(I + (\Delta t/h^2)A)U_1 = U_0 + \Delta t\, F(0)$$

lösen, und es läßt sich zeigen, daß die Matrix $(I + (\Delta t/h^2)A)^{-1}$, die man auf die vorherige Zeitschicht anwenden muß, mit positiven Koeffizienten vollbesetzt ist.

Für weitere Betrachtungen im Hinblick auf Stablitätsfragen zu Problemen der Art (8.7) verweisen wir auf das Lehrbuch von P. Deuflhard und F. Bornemann [8].

Bemerkung: Die Änderung der Dimension von S_h, also der Approximationsgüte, ändert die Anzahl der Unbekannten in (8.7) und macht einen Restart der Zeitintegration erforderlich. Das bedeutet Einschränkungen bei *adaptiver Wahl* der Ortsdiskretisierung, die vor allem in 2 und mehr Raumdimensionen schmerzlich spürbar werden.

8.2.2 Rothe-Methode

Beim Existenzbeweis in Abschnitt 7.1.2 wie auch bei der Linienmethode haben wir das Anfangsrandwertproblem für die Wärmeleitungsgleichung durch ein Anfangswertproblem für ein großes System gewöhnlicher Differentialgleichungen ersetzt. Wir wollen nun die schwache Formulierung von 7.15 (siehe (7.19)) direkt als Anfangswertproblem in einem *unendlichdimensionalen* Funktionenraum formulieren. Dazu bedarf es einiger Vorbereitungen.

Jedes $g \in L^2(\Omega)$ definiert durch $v \mapsto (g, v)$, $v \in H_0^1(\Omega)$, ein beschränktes lineares Funktional auf $H_0^1(\Omega)$, also ist in diesem Sinne

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H_0^1(\Omega)' \;\;!$$

Für jedes $w \in H_0^1(\Omega)$ ist $a(w, \cdot) \in H_0^1(\Omega)'$! Damit ist durch

$$Aw \in L^2(\Omega) : \quad (Aw, v) = a(w, v) \qquad \forall v \in H_0^1(\Omega) \tag{8.8}$$

eine lineare (unbeschränkte!) Abbildung

$$A : \quad D(A) \subset H_0^1(\Omega) \to L^2(\Omega)$$

definiert, wobei

$$D(A) := \{w \in H_0^1(\Omega) \,|\; \text{(8.7) hat eine Lösung } \}$$

gesetzt ist. A heißt L^2-*Darstellung* der Bilinearform $a(\cdot, \cdot)$.

Beispiel:

Im Falle $a(v, w) = \int_\Omega \nabla v \, \nabla w \, dx$ ist

$$D(A) = H^2(\Omega) \cap H_0^1(\Omega)$$

und

$$Aw = -\Delta w \quad , \quad w \in D(A) \quad .$$

Durch Vervollständigung von $C^1([0, T], L^2(\Omega))$ bezüglich des Skalarprodukts

$$(v, w)_{H^1((0,T),L^2(\Omega))} := \int_0^T (v(t), w(t))\, dt + \int_0^T (v'(t), w'(t))\, dt$$

erhalten wir den Raum $H^1((0, T), L^2(\Omega))$. Dieser Raum besteht wie im reellwertigen Fall aus allen (Äquivalenzklassen von) Funktionen $v \in L^2((0, T), L^2(\Omega))$, deren schwache Ableitung v' wieder in $L^2((0, T), L^2(\Omega))$ liegt. Dabei heißt v' schwache Ableitung, wenn

$$\int_0^T (v(t), \varphi'(t))\, dt = -\int_0^T (v'(t), \varphi(t))\, dt$$

für alle $\varphi \in C_0^\infty((0, T), L^2(\Omega))$ richtig ist. (Man definiert den Raums $C_0^\infty((0, T), L^2(\Omega))$ analog zum Fall reellwertiger Funktionen.)
Wir betrachten nun das folgende Anfangswertproblem in $L^2(\Omega)$:
Finde $u \in L^2((0, T), H_0^1(\Omega)) \cap H^1((0, T), L^2(\Omega))$ so, daß

$$\begin{aligned} u' - \Delta u &= f \\ u(0) &= u_0 \end{aligned} \tag{8.9}$$

in $L^2(\Omega)$. Die Anfangsbedingung wird stetig in $L^2(\Omega)$ angenommen, denn es gilt wie bei reellwertigen Funktionen

Lemma 8.5

$$H^1((0,T), L^2(\Omega)) \subset C([0,T], L^2(\Omega)) \quad .$$

Den Zusammenhang mit schwachen Lösungen klärt folgender Satz.

Satz 8.6 *Jede Lösung von* (8.9) *ist eine schwache Lösung von* (7.15). *Umgekehrt ist jede schwache Lösung* $u \in H^1((0,T), L^2(\Omega))$ *von* (7.15) *eine Lösung von* (8.9).

Beweis:

Sei u eine Lösung von (8.9). Multiplikation mit einem beliebigen $v \in H_0^1(\Omega)$ und Integration über Ω liefert

$$(u'(t), v) + (Au(t), v) = (f(t), v) \qquad \forall t \in (0,T) \quad . \tag{8.10}$$

Sei $\varphi \in C_0^\infty([0,T])$. Dann gilt $v\varphi \in C_0^\infty((0,T), L^2(\Omega))$ mit $(v\varphi)' = v\varphi'$ (warum?), und wegen $u \in H^1((0,T), L^2(\Omega))$ ist dann

$$\begin{aligned} \int_0^T (u'(t), v)\varphi(t)\,dt &= \int_0^T (u'(t), v\varphi(t))\,dt \\ &= -\int_0^T (u(t), v\varphi'(t))\,dt \\ &= -\int_0^T (u(t), v)\varphi'(t)\,dt \quad . \end{aligned}$$

Damit ist $(u(\cdot), v)$ schwach differenzierbar mit

$$(u'(t), v) = \frac{d}{dt}(u(t), v) \qquad \forall v \in H_0^1(\Omega) \quad . \tag{8.11}$$

Definitionsgemäß ist auch

$$(Au, v) = a(u, v) \qquad \forall v \in H_0^1(\Omega) \quad .$$

Einsetzen in (8.9) liefert

$$\frac{d}{dt}(u(t), v) + (Au(t), v) = (f(t), v) \qquad \forall v \in (0,T) \quad .$$

Mit Blick auf Lemma 8.5 ist u also eine schwache Lösung.
Ist umgekehrt $u \in H^1((0,T), L^2(\Omega))$ schwache Lösung, so gilt wieder (8.10) und somit

$$a(u(t), v) = (f(t) - u'(t), v) \qquad \forall v \in H_0^1(\Omega) \quad .$$

Da $f(t) - u'(t) \in L^2(\Omega)$ ist, folgt $u(t) \in D(A)$ für fast alle $t \in (0,T)$, und wir erhalten

$$(u'(t) + Au(t) - f(t), v) = 0 \qquad \forall v \in H_0^1(\Omega) \quad .$$

Schließlich ist $H_0^1(\Omega) \subset L^2(\Omega)$ dicht und somit

$$u'(t) + Au(t) = f(t) \quad \text{ f.ü. in } (0,T) \quad .$$

□

Die Formulierung (8.9) ist Ausgangspunkt für Halbgruppen-Methoden für Evolutionsprobleme. Wir verweisen auf [22, Chapter 111] und die dort zitierte Literatur.

Grundidee der Rothe-Methode:

(Originalartikel [23])

a) Semi-Diskretisierung in t

   Übertrage bekannte Techniken zur Lösung von Anfangswertproblemen für gewöhnliche Differentialgleichungen im $\mathbb{R}^m$ auf das Anfangswertproblem (8.9) (Ordnungs- und Schrittweitensteuerung)

b) Diskretisierung im Ort

   Die resultierenden Ortsprobleme müssen schließlich so genau gelöst werden, daß die Eigenschaften der Zeitdiskretisierung erhalten bleiben.

Beispiel: Implizites Eulerverfahren zur Schrittweite Δt

In jedem Zeitschritt hat man das Ortsproblem

$$U(t_i) - U(t_{i-1}) + \Delta t \, AU(t_i) = \Delta t \, f(t_i)$$

oder äquivalent das Variationsproblem

$$(U(t_i), v) + \Delta t \, a(U(t_i), v) = (U(t_{i-1}) + \Delta t \, f(t_i), v) \qquad \forall v \in H_0^1(\Omega) \tag{8.12}$$

zu lösen. Zur Lösung vom (8.11) kann man jetzt vorhandene Verfahren (und Codes) für elliptische Probleme verwenden. Insbesondere ist in jedem Zeitschritt ein anderes Ortsgitter möglich. Um die Genauigkeitsanforderungen an die Ortsdiskretisierung so niedrig wie möglich zu halten, sollten möglichst robuste Zeitintegratoren verwendet werden, Zum Beispiel scheiden Extrapolationsverfahren aus, da das zugrundeliegende Problem für wachsende Ordnung schlecht konditioniert ist. Einzelheiten finden sich in den jüngsten Arbeiten von F. Bornemann [4].

Literatur

[1] Adams, R.A. :*Sobolev Spaces*, Academic Press, 1975.

[2] Alt, H. : *Lineare Funktionalanalysis*, (2.Auflage), Springer 1992.

[3] Bjørstad, P.E./ Gropp, W.D./ Smith, B.F. : *Domain Decomposition*, Cambridge University Press, New York 1996.

[4] Bornemann, F. : *An adaptive multilevel approach to parabolic equations I-III*, IMPACT Comput. Sci. Engrg. **2-4**, 1990–1992.

[5] Braess, D. : *Finite Elemente*, (2.Auflage), Springer, 1997.

[6] Ciarlet, Ph. : *The Finite Element Method for Elliptic Problems*, North-Holland, 1978.

[7] Deuflhard, P. / Hohmann, A. : *Numerische Mathematik I*, (2. Auflage) de Gruyter, Berlin 1993.

[8] Deuflhard, P. / Bornemann, F : *Numerische Mathematik II*, de Gruyter, Berlin 1994.

[9] George, P.L. : *Automatic Mesh Generation*, Wiley, 1991.

[10] George, P.L. : *Delaunay triangulation and meshing*, Hermes, Paris 1998.

[11] Hackbusch, W. : *Multi-Grid Methods and Applications*, Springer, 1985.

[12] Hackbusch, W. : *Theorie und Numerik elliptischer Differentialgleichungen*, Teubner, 1986.

[13] Hackbusch, W. : *Iterative Solution of Large Sparse Systems of Equations*, Springer 1994.

[14] John, F. : *Partial Differential Equations*, (4. Auflage), Springer, 1982.

[15] Jost, J. : *Partielle Differentialgleichungen*, Springer, 1998.

[16] Kornhuber, R. : *Theorie und Numerik partieller Differentialgleichungen*, Skript WS 98/99, FU Berlin, 1999.

[17] Meyers, N.G., Serrin, J. : $H = W$, Proc. Nat. Acad. Sci USA, **51**, p. 1055–1056, 1964.

[18] Neuss, N. : *V-cycle convergence with unsymmetric smoothers and application to an anisotropic model problem*, SIAM J. Num. Anal., **35**, p. 1201–1212, 1998.

[19] Oswald, P.: *Multilevel Finite Element Approximation*, Teubner, 1994.

[20] Press, W.H.;Teukolsky, S.A.; Vetterling, W.T. und Flannery, B.P. : *Numerical Recipes in C* Cambridge University Press, 1992.

[21] Raviart, P.A. und Thomas, J.M. : *Introduction à l'analyse numérique des équations aux dérivées partielles.* Masson, 1992.

[22] Renardy, M. und Rogers, R. C. : *An Introduction to Partial Differential Equations*, Springer, 1993.

[23] Rothe, E. : *Zweidimensionale Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben.*, Math. Ann. **102**, p. 650–670, 1930.

[24] Stoer, J. und Bulirsch, R. : *Numerische Mathematik II*, Springer, 1990.

[25] Strauss, W. A. : *Partielle Differentialgleichungen*, Vieweg, 1995.

[26] Thomée, V. : *Galerkin Finite Element Methods for Parabolic Problems*, Springer, 1997.

[27] Verfürth, R. : *A Review of A Posteriori Error Estimation and Mesh–Refinement Techniques*, Wiley–Teubner, 1996.

[28] Werner, D.: *Funktionalanalysis*, (3. Auflage) Springer, Berlin, 2000.

[29] Xu, J. : *Iterative methods by space decomposition and subspace correction*, *SIAM Review*, **34**, p. 581–613, 1992.

[30] Yserentant, H. : *Old and new convergence proofs for multigrid methods*, Acta Numerica 1993, p. 285–326. Cambridge University Press, 1993