

Mitschrift
Numerik von partiellen Differentialgleichungen
Vortragender: J. M. Melenk

Birgit Pichler

Bernhard Skritek

Ines Stelzer

Wien
SS 2008

Inhaltsverzeichnis

6	Differenzenverfahren für elliptische PDE	1
6.1	Maximumsprinzip	1
6.2	Differenzenverfahren	3
6.2.1	Diskretisierung	3
6.3	Bemerkungen zu den Randbedingungen	10
6.3.1	Neumann-Randbedingungen in 1D	10
6.3.2	Nicht gleichmäßige Gitter in 1D	11
6.3.3	Krummlinige Ränder	11
7	FEM	13
7.1	Variationsformulierung	13
7.2	Steilsturz zu $H^1(\Omega)$	15
7.3	Klassische FEM in 1D	17
7.4	Konvergenzanalyse der FEM im 1D	22
7.5	FEM in 2D	23
7.5.1	$H^1(\Omega)$ und $H_0^1(\Omega)$	23
7.5.2	Schwache Formulierung von (7.28)	25
7.5.3	FEM	26
8	FEM für parabolische Probleme	31
8.1	Variationsformulierung im Ort	31
8.2	Semidiskretisierung im Ort	34
8.3	Volldiskrete Verfahren	37
8.3.1	Zusammenfassendes Beispiel	45

6 Differenzenverfahren für elliptische PDE

Sei $\Omega \subset \mathbb{R}^d$ ein beschränktes Gebiet. Die Aufgabe lautet nun:

$$\left. \begin{aligned} \text{Finde } u \in C^2(\Omega) \cap C(\bar{\Omega}), \text{ sodass} \\ Lu := -\Delta u + b(x) \cdot \nabla u + c(x) \cdot u = f(x), \quad \forall x \in \Omega \\ u(x) = 0, \quad \forall x \in \partial\Omega. \end{aligned} \right\} \quad (6.1)$$

Hier sind b, c, f gegebene glatte Funktionen. Wir machen die Generalvoraussetzung

$$c \geq 0 \text{ auf } \Omega. \quad (6.2)$$

Die Poisson-Gleichung ist gegeben durch $(\Delta u)(x) = \partial_1^2 u(x) + \partial_2^2 u(x) + \dots + \partial_d^2 u(x)$.

6.1 Maximumsprinzip

Satz 6.1 (*Schwaches Maximumprinzip*)

Sei $u \in C^2(\Omega) \cap C(\bar{\Omega})$ mit $Lu \leq 0$ auf Ω . Dann gilt:

- (i) falls $c \equiv 0$, dann ist $\max_{x \in \Omega} u(x) = \max_{x \in \partial\Omega} u(x)$
- (ii) falls $c \geq 0$, dann ist $\max_{x \in \Omega} u(x) \leq \max_{x \in \partial\Omega} u^+(x)$ mit $u^+(x) := \max\{u(x), 0\}$.

Beweis: ad(i) Der Beweis wird in 2 Schritten geführt:

1. *Schritt:* Sei $Lu < 0$ auf Ω . Angenommen es existiert ein $x_0 \in \Omega$ mit

$$u(x_0) = \max_{x \in \Omega} u(x).$$

Dann ist $\nabla u(x_0) = 0$ und die Hess'sche Matrix H der 2. Ableitungen ist symmetrisch und negativ semidefinit. Insbesondere ist $\partial_i^2 u(x_0) \leq 0$, $\forall i = 1, \dots, d$ und damit $\Delta u(x_0) \leq 0$. Also ergibt sich der Widerspruch

$$0 > (Lu)(x_0) = \underbrace{-\Delta u(x_0)}_{\geq 0} + b(x_0) \cdot \underbrace{\nabla u(x_0)}_{=0} + \underbrace{c(x_0)}_{=0} \cdot u(x_0) \geq 0.$$

2. *Schritt:* Sei $Lu \leq 0$ auf Ω . Sei φ eine Funktion mit

- $\varphi \geq 0$ auf $\bar{\Omega}$
- $L\varphi < 0$ auf $\bar{\Omega}$.

So eine Funktion kann mit dem Ansatz $\varphi(x_1, \dots, x_d) = e^{\lambda x_1}$ konstruiert werden. Damit ist

$$L\varphi = -\lambda^2 e^{\lambda x_1} + (b(x) \mathbf{e}_1) \lambda e^{\lambda x_1},$$

wobei $\mathbf{e}_1 = (1, 0, \dots, 0)^T$. Weil Ω beschränkt ist, ergibt sich φ mit den gewünschten Eigenschaften für λ hinreichend groß. Für jedes $\varepsilon > 0$ gilt dann $L(u + \varepsilon\varphi) = \underbrace{Lu}_{\leq 0} + \varepsilon L\varphi \leq \varepsilon L\varphi < 0$.

$$1. \text{ Schritt} \Rightarrow \max_{x \in \Omega} (u(x) + \varepsilon\varphi(x)) = \max_{x \in \partial\Omega} (u(x) + \varepsilon\varphi(x)).$$

Für $\varepsilon \rightarrow 0$ folgt somit, dass $\max_{x \in \tilde{\Omega}}(u(x)) = \max_{x \in \partial\Omega}(u(x))$.

ad(ii): Sei $\Omega^+ := \{x \in \tilde{\Omega} \mid u(x) > 0\}$. Dann gilt für den Operator \tilde{L} definiert durch $\tilde{L}u := -\Delta u + \mathbf{b}\nabla u$, dass $\tilde{L}u = Lu - cu \leq -cu \leq 0$ auf Ω^+ . Nach (i) folgt also für $\Omega^+ \neq \emptyset$, dass

$$\begin{aligned} \max_{x \in \tilde{\Omega}} u(x) &\stackrel{u(x) \leq 0 \text{ für } x \in \Omega \setminus \Omega^+}{\leq} \max_{x \in \Omega^+} u(x) \stackrel{(i)}{=} \max_{x \in \partial\Omega^+} u(x) \leq \max_{x \in \partial\Omega^+} u^+ \leq \\ &\leq \max\{0, \max_{(x \in \partial\Omega^+ \cap \partial\Omega)} u^+\} \leq \max_{x \in \partial\Omega} u^+. \end{aligned}$$

Falls $\Omega^+ = \emptyset$, dann ist $\max_{x \in \tilde{\Omega}} u \leq \max_{x \in \partial\Omega} u^+$ trivialerweise erfüllt. \square

Korollar 6.2 (*Vergleichsprinzip*) Seien $u, v \in C^2(\Omega) \cap C(\bar{\Omega})$ mit

- $Lu \leq Lv$ auf Ω
- $u \leq v$ auf $\partial\Omega$

Dann ist $u \leq v$ auf $\bar{\Omega}$.

Beweis: Übung (Wende Satz 6.1 auf $u - v$ an.) \square

Korollar 6.3 (*Stabilität in $\|\cdot\|_C$*) Es existiert eine Konstante $C > 0$ sodass für jedes $u \in C^2(\Omega) \cap C(\bar{\Omega})$ gilt:

$$\|u\|_{C(\bar{\Omega})} \leq \|u\|_{C(\partial\Omega)} + C\|Lu\|_{C(\bar{\Omega})}$$

Beweis: Sei $\varphi \geq 0$ mit $L\varphi \leq -1$ auf $\bar{\Omega}$. (Dass so eine Funktion φ existiert, sieht man aus dem Beweis von Satz 6.1.) Definiere die Funktion

$$u^\pm := \pm u + \|Lu\|_{C(\bar{\Omega})}\varphi.$$

Dann ist $Lu^\pm = \pm Lu + \|Lu\|_{C(\bar{\Omega})}L\varphi \leq \pm Lu - \|Lu\|_{C(\bar{\Omega})} \leq 0$ auf Ω .

$$\begin{aligned} \stackrel{\text{Satz 6.1}}{\Rightarrow} \max_{\bar{\Omega}} u^\pm &\leq \max_{\partial\Omega} u^\pm \leq \max_{\partial\Omega} (\pm u) + \|Lu\|_{C(\bar{\Omega})}\|\varphi\|_{C(\bar{\Omega})} \leq \\ &\leq \max_{\partial\Omega} (\pm u) + C\|Lu\|_{C(\bar{\Omega})} \leq \|u\|_{C(\partial\Omega)} + C\|Lu\|_{C(\bar{\Omega})} \end{aligned}$$

Wegen $\pm u = u^\pm - \|Lu\|_{C(\bar{\Omega})}\varphi \leq u^\pm$ folgt also $\|u\|_{C(\bar{\Omega})} \leq \|u\|_{C(\partial\Omega)} + C\|Lu\|_{C(\bar{\Omega})}$. \square

Korollar 6.3 impliziert die Eindeutigkeit der Lösung von $Lu = f$ auf Ω , $u = g$ auf $\partial\Omega$ für gegebenen Funktionen f und g . Der Beweis davon wird dem Leser überlassen. Die Existenz einer Lösung von (6.1) ist jedoch noch nicht gesichert. Hierzu benötigt man typischerweise Regularität von $\partial\Omega$ und f .

6.2 Differenzenverfahren

Differenzenverfahren sind die einfachsten numerischen Verfahren zur Lösung von Differentialgleichungen. Dabei wird zunächst das Gebiet, für das die Gleichung gelten soll, in eine endliche Zahl von Gitterpunkten zerlegt. Dies geschieht meist durch ein Gitter von senkrecht aufeinander stehenden Linien, den Gitterpunkten entsprechen dann die Kreuzungspunkte. Die Ableitungen an den Gitterpunkten werden dann durch Differenzen approximiert. Die partiellen Differentialgleichungen werden so in ein System von Differenzgleichungen umformuliert und mittels verschiedener Algorithmen entweder implizit oder explizit gelöst.

Betrachte das Modellproblem

$$\text{Finde } u \in C^2(\Omega) \cap C(\bar{\Omega}), \text{ sodass } -\Delta u = f \text{ auf } \Omega = (0,1)^2, u = 0 \text{ auf } \partial\Omega \quad (6.3)$$

für gegebenes $f \in C(\bar{\Omega})$.

6.2.1 Diskretisierung

Der Einfachheit halber wird ein uniformes Gitter $x_{ij} := (x_i, x_j)$ verwendet, wobei $x_i = ih$, $i = 0, \dots, n$, $n \in \mathbb{N}$ und $h = \frac{1}{n}$.

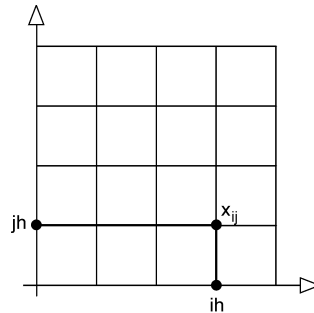


Abbildung 6.1: Ausschnitt eines Gitters mit Gitterpunkt x_{ij}

Wir definieren weiters

- $\bar{\Omega}_h := \{x_{ij} \mid 0 \leq i, j \leq n\}$ = Menge aller Gitterpunkte
- $\Omega_h := \{x_{ij} \mid 1 \leq i, j \leq n-1\}$ = Menge der inneren Gitterpunkte
- $\partial\Omega_h = \bar{\Omega}_h \setminus \Omega_h$ = Menge der Gitterpunkte am Rand

Die Gitterfunktionen $u^h : \bar{\Omega}_h \rightarrow \mathbb{R}$ sind Abbildungen, die auf dem Gitter definiert sind. Wir verwenden austauschbar die Notation $u_{ij}^h = u^h(x_{ij})$ für $x_{ij} \in \bar{\Omega}_h$.

Weil im 1D (also im Eindimensionalen)

$$u''(x) \approx \left[\frac{u(x+h) - u(x)}{h} - \frac{u(x) - u(x-h)}{h} \right] \cdot \frac{1}{h} = \frac{1}{h^2} (-2u(x) - u(x-h) - u(x+h))$$

ist, ist eine Approximation an $-\Delta u(x_{ij})$ durch

$$-\Delta u(x_{ij}) \approx \frac{1}{h^2} (4u(x_{ij}) - u(x_{i-1,j}) - u(x_{i+1,j}) - u(x_{i,j-1}) - u(x_{i,j+1}))$$

gegeben. Zur Diskretisierung von $L = -\Delta$ verwenden wir deshalb

$$(L_h u^h)_{ij} := \frac{4u_{ij}^h - u_{i+1,j}^h - u_{i-1,j}^h - u_{i,j+1}^h - u_{i,j-1}^h}{h^2} \quad \text{für } i, j = 1, \dots, n-1.$$

Diese Diskretisierung von $-\Delta = -\partial_1^2 - \partial_2^2$ verwendet also jeweils die unmittelbar benachbarten Gitterpunkte von x_{ij} und wird kurz als sogenannter Differenzenstern notiert, vgl. Abbildung 6.2.

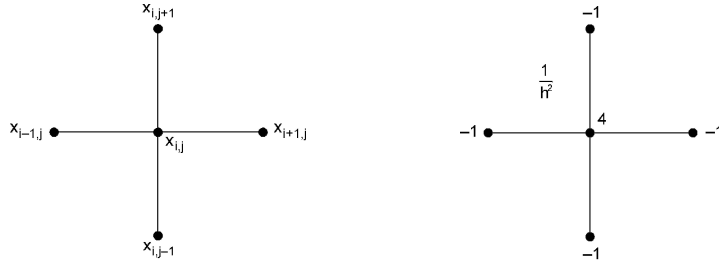


Abbildung 6.2: x_{ij} mit seinen unmittelbaren Nachbarn (links) und der dazugehörige Differenzenstern (rechts)

Damit ergibt sich als Differenzenverfahren für (6.3)

$$\left. \begin{aligned} &\text{Finde eine Gitterfunktion } u^h : \bar{\Omega}_h \rightarrow \mathbb{R}, \text{ sodass} \\ &L_h u^h(x_{ij}) = f_{ij} := f(x_{ij}), \quad x_{ij} \in \Omega_h \\ &R_h u^h(x_{ij}) = 0 \quad \forall x_{ij} \in \partial\Omega_h \end{aligned} \right\} \quad (6.4)$$

wobei $R_h u^h(x_{ij}) := u^h(x_{ij})$

(6.4) lässt sich als LGS für die Funktionswerte u_{ij}^h in den inneren Gitterpunkten schreiben und man erhält eine Lösung der Form

$$\mathbf{A}_h \mathbf{u} = \mathbf{f} \quad (6.5)$$

wobei $(\mathbf{I} = \mathbf{Id}_{n-1})$

$$\mathbf{A}_h = \frac{1}{h^2} \begin{pmatrix} \mathbf{T} & -\mathbf{I} & & & \\ -\mathbf{I} & \mathbf{T} & -\mathbf{I} & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -\mathbf{I} \\ & & & -\mathbf{I} & \mathbf{T} \end{pmatrix} \in \mathbb{R}^{(n-1)^2 \times (n-1)^2} \quad (6.6)$$

mit

$$\mathbf{T} = \begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 4 \end{pmatrix} \in \mathbb{R}^{(n-1) \times (n-1)} \quad (6.7)$$

Dabei haben wir die Randbedingung $u^h(x_{ij}) = 0$ für $x_{ij} \in \partial\Omega_h$ explizit berücksichtigt und die „Doppelindizes“ (i, j) durch „lexikographische“ Nummerierung („zeilenweise“) in Einfachindizes umgewandelt, d.h. die Anordnung ist $(1, 1), (1, 2), (1, 3), \dots, (1, n-1), (2, 1), (2, 2), \dots, (2, n-1), \dots$. In der rechten Graphik von Abb. 6.3 ist nochmals ein Gitter zu sehen, in dem beispielhaft einige Indizes eingetragen sind.

Zum Messen der Konsistenz definieren wir die Funktionen

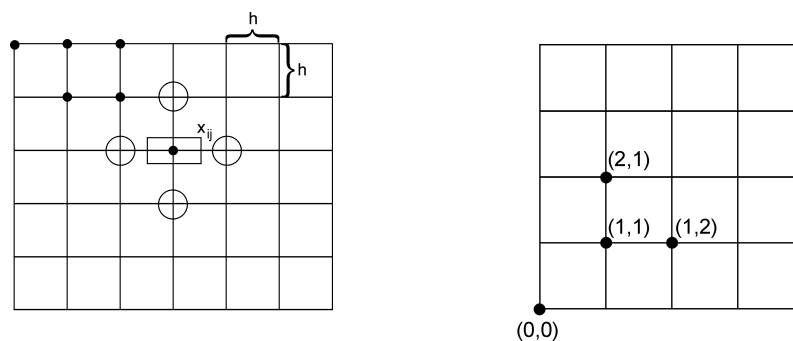


Abbildung 6.3: Links im Bild ist wieder ein uniformes Gitter dargestellt, rechts ist ein Gitter mit einigen beispielhaft eingetragenen Doppelindizes zu sehen

$$[\cdot]_h : C(\bar{\Omega}) \rightarrow \text{Gitterfunktion auf } \bar{\Omega}_h$$

$$u \mapsto ([u]_h)(x_{ij}) = u(x_{ij}) \quad \forall x_{ij} \in \bar{\Omega}_h$$

$$[L \cdot]_h : C^2(\Omega) \cap C(\bar{\Omega}) \rightarrow \text{Gitterfunktion auf } \Omega_h$$

$$u \mapsto ([Lu]_h)(x_{ij}) = (Lu)(x_{ij}) \quad \forall x_{ij} \in \Omega_h$$

$$[R \cdot]_h : C(\bar{\Omega}) \rightarrow \text{Gitterfunktion auf } \partial\Omega_h$$

$$u \mapsto ([Ru]_h)(x_{ij}) = u(x_{ij}) \quad \forall x_{ij} \in \partial\Omega_h$$

Wir definieren nun den *Konsistenzfehler* wieder durch die Vorschrift „Einsetzen der exakten Lösung in das numerische Verfahren“:

$$\tau(h) = \|L_h[u]_h - [Lu]_h\|_\infty + \|R_h([u]_h) - [Ru]_h\|_\infty = \|[Lu]_h - L_h[u]_h\|_\infty$$

Lemma 6.4 (Konsistenz) Sei $u \in C^4(\bar{\Omega})$. Dann gilt:

$$\|[Lu]_h - L_h[u]_h\|_\infty \leq Ch^2 \max_{|\alpha|=4} \|D^\alpha u\|_{C(\bar{\Omega})}$$

Beweis: Taylorentwicklung um jeden Gitterpunkt $x_{ij} \in \Omega_h$ □

Sei u eine Lösung von (6.3) und u^h eine Lösung von (6.4). Aus dem Satz 5.17 „Stabilität + Konsistenz \Rightarrow Konvergenz“ folgt dann für den Fehler $[u]_h - u^h$

$$\|[u]_h - u^h\|_\infty \leq Ch^2, \text{ falls } u \in C^4(\bar{\Omega}) \text{ und } \|\mathbf{A}_h^{-1}\|_\infty \leq C$$

für $C > 0$ unabhängig von h .

Um $\|\mathbf{A}_h^{-1}\|_\infty \leq C$ zu zeigen, kann man entweder versuchen \mathbf{A}_h^{-1} auszurechnen, was auf die „diskrete Green'sche Funktion“ führen würde, oder man verwendet das Maximumsprinzip, was wir im Folgenden tun werden.

Lemma 6.5 (diskretes Maximumsprinzip)

Sei $u^h : \bar{\Omega}_h \rightarrow \mathbb{R}$ eine Gitterfunktion mit $(L_h u^h)(x_{ij}) \leq 0 \forall x_{ij} \in \Omega_h$. Dann gilt:

$$\max_{x_{ij} \in \Omega_h} u^h(x_{ij}) \leq \max_{x_{ij} \in \partial\Omega_h} u^h(x_{ij})$$

Beweis: Angenommen u^h nimmt sein Maximum in $x_{ij} \in \Omega_h$ an. Sei $x_{ij}^* \in \Omega_h$ oBdA so, dass ein unmittelbarer Nachbar $\tilde{x}_{ij} \in \{x_{i-1,j}^*, x_{i+1,j}^*, x_{i,j-1}^*, x_{i,j+1}^*\}$ mit $u^h(\tilde{x}_{ij}) < u^h(x_{ij}^*)$ existiert. Dann erhält man den Widerspruch

$$0 \geq h^2(L_h u^h)(x_{ij}^*) = 4u^h(x_{ij}^*) - u^h(x_{i-1,j}^*) - u^h(x_{i+1,j}^*) - u^h(x_{i,j-1}^*) - u^h(x_{i,j+1}^*) > 0$$

□

Satz 6.6 (diskretes Vergleichsprinzip) Seien u^h, v^h Gitterfunktionen mit

- $L_h u^h \leq L_h v^h$ auf Ω_h
- $u^h \leq v^h$ auf $\partial\Omega_h$

Dann ist $u^h \leq v^h$ auf $\bar{\Omega}_h$.

Beweis: Wende Lemma 6.5 (diskretes Maximumsprinzip) auf $u^h - v^h$ an. □

Satz 6.7 (diskrete Stabilität)

Definiere die lineare Abbildung $F_h : v^h \mapsto (L_h v^h, R_h v^h)$. Dann existiert eine Konstante $C > 0$ unabhängig von h , sodass für jede Gitterfunktion $v^h : \bar{\Omega}_h \rightarrow \mathbb{R}$ gilt:

$$\|v^h\|_\infty \leq \|R_h v^h\|_\infty + C \|L_h v^h\|_\infty \leq 2C \|F_h v^h\|_\infty$$

Beweis: Der Beweis basiert auf dem diskreten Maximumsprinzip. Sei $\varphi_h : \bar{\Omega}_h \rightarrow \mathbb{R}$ eine Gitterfunktion mit folgenden Eigenschaften

- a) $\varphi_h \geq 0$ (komponentenweise)
- b) $L_h \varphi_h \geq 1$ (komponentenweise)
- c) $\|\varphi_h\|_\infty \leq C$ für eine Konstante C unabhängig von h

Diese Eigenschaften erfüllt z.B. $\varphi_h = [x(1-x)]_h$. Definiere nun die Funktion

$$v^\pm = \pm v^h - \|L_h v^h\|_\infty \varphi_h$$

Dann folgt aus $L_h v^\pm \leq 0$ mit Lemma 6.5

$$\begin{aligned} \Rightarrow \max_{x \in \Omega_h} v^\pm(x) &\leq \max_{x \in \partial\Omega_h} v^\pm(x) = \max_{x \in \partial\Omega_h} \pm v^h(x) - \|L_h v^h\|_\infty \underbrace{\varphi_h}_{\geq 0} \leq \max_{x \in \partial\Omega_h} \pm v^h(x) \leq \\ &\leq \max_{x \in \partial\Omega_h} |v^h(x)| = \|R_h v^h\|_\infty \end{aligned}$$

Wegen $\pm v^h = v^\pm + \|L_h v^h\|_\infty \varphi_h$ folgt also

$$\|v^h\|_\infty \leq \|v^\pm\|_\infty + \|L_h v^h\|_\infty \|\varphi_h\|_\infty = \|R_h v^h\|_\infty + \|L_h v^h\|_\infty \|\varphi_h\|_\infty$$

□

Satz 6.8 Sei $u \in C^4(\bar{\Omega})$ eine Lösung von (6.1). Sei u^h die nach Satz 6.7 existierende Lösung von (6.4). Dann gilt mit einer von h unabhängigen Konstanten $C > 0$

$$\|u^h - [u]_h\|_\infty \leq Ch^2 \max_{|\alpha|=4} \|D^\alpha u\|_{C(\bar{\Omega})}.$$

Beweis: Der Beweis wird ähnlich zu jenem von Satz 5.17 „Konsistenz + Stabilität \Rightarrow Konvergenz“ geführt. Sei $v^h := u^h - [u]_h$. Dann folgt mit Satz 6.7

$$\begin{aligned} \|u^h - [u]_h\|_\infty &\leq \underbrace{\|R_h u^h - R_h [u]_h\|_\infty}_{=0} + C \underbrace{\|L_h u^h - L_h [u]_h\|_\infty}_{=[f]_h} \leq \\ &\leq C \underbrace{\|[f]_h - [Lu]_h}_{=[f]_h} + [Lu]_h - L_h [u]_h\|_\infty \stackrel{\text{Lemma 6.4}}{=} Ch^2 \max_{|\alpha|=4} \|D^\alpha u\|_{C(\bar{\Omega})}. \end{aligned}$$

□

Die Matrix \mathbf{A}_h aus (6.5) erfüllt die Abschätzung $\|\mathbf{A}_h^{-1}\|_\infty \leq C$ gleichmäßig in h nach dem Stabilitätsresultat Satz 6.7. Der Beweis dazu beruht auf dem Lemma 6.5, welches folgendes Monotonieprinzip ausdrückt:

$$\mathbf{A}_h \mathbf{u} \leq \mathbf{0} \text{ (komponentenweise)} \Rightarrow \mathbf{u} \leq \mathbf{0} \text{ (komponentenweise)}.$$

Diese Eigenschaft kann durch den Begriff der M -Matrix verallgemeinert werden:

Definition 6.9 Eine Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ heißt

- (i) L_0 -Matrix, falls $\mathbf{A}_{ij} \leq 0 \quad \forall i \neq j$
- (ii) L -Matrix, falls $\mathbf{A}_{ii} > 0 \quad \forall i$ und $\mathbf{A}_{ij} \leq 0 \quad \forall i \neq j$
- (iii) inversmonoton, falls $\mathbf{A}\mathbf{x} \leq \mathbf{A}\mathbf{y}$ (komponentenweise) $\Rightarrow \mathbf{x} \leq \mathbf{y}$ (komponentenweise)
- (iv) M -Matrix, falls \mathbf{A} eine L_0 -Matrix ist, für die \mathbf{A}^{-1} existiert und $\mathbf{A}^{-1} \geq 0$ ist (elementweise)

Übung 6.10 Zeigen Sie, dass eine inversmonotone Matrix \mathbf{A} invertierbar ist und $\mathbf{A}^{-1} \geq 0$ elementweise gilt. Daraus folgt, dass M -Matrizen gerade die inversmonotonen L_0 -Matrizen sind.

Wir sehen, dass \mathbf{A}_h eine M -Matrix ist. Der Kern des Beweises von Satz 6.7 (d.h. $\|\mathbf{A}_h^{-1}\|_\infty \leq C$) war die Existenz der Gitterfunktionen $\varphi_h \geq 0$ mit $L_h \varphi_h \geq 1$. Die Existenz einer solchen Funktion ist charakteristisch für M -Matrizen:

Satz 6.11 (M -Kriterium)

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ eine L_0 -Matrix. Dann ist \mathbf{A} inversmonoton genau dann, wenn ein Vektor $\varphi > \mathbf{0}$, $\varphi \in \mathbb{R}^n$, mit $\mathbf{A}\varphi > \mathbf{0}$ (komponentenweise) existiert. In diesem Fall gilt:

$$\|\mathbf{A}^{-1}\|_\infty \leq \frac{\|\varphi\|_\infty}{\min_{i=1, \dots, n} (\mathbf{A}\varphi)_i}$$

Beweis:

„ \Rightarrow “ Übung 6.10 zeigt die Existenz von \mathbf{A}^{-1} und $\mathbf{A}^{-1} \geq 0$. Also erfüllt $\varphi = \mathbf{A}^{-1}(1 \dots 1)^T$ das Gewünschte.

„ \Leftarrow “ Zu zeigen ist

(1) \mathbf{A}^{-1} existiert und

(2) $\mathbf{A}^{-1} \geq 0$ (elementweise)

Sei $\varphi > 0$ mit $\mathbf{A}\varphi > 0$. Dann gilt:

$$\mathbf{A}_{ii} = \frac{\mathbf{A}_{ii}\varphi_i}{\varphi_i} = \frac{1}{\varphi_i} \left(\underbrace{\sum_{j=1}^n \mathbf{A}_{ij}\varphi_j}_{>0 \text{ nach Vorr.}} - \sum_{\substack{j=1 \\ j \neq i}}^n \underbrace{\mathbf{A}_{ij}}_{\leq 0} \underbrace{\varphi_j}_{> 0} \right)$$

also existiert \mathbf{D}^{-1} , wenn $\mathbf{D} := \text{diag}(\mathbf{A})$ und $\mathbf{D}^{-1} \geq 0$ komponentenweise. Setze $\mathbf{P} := \mathbf{D}^{-1}(\mathbf{D} - \mathbf{A}) = \mathbf{I} - \mathbf{D}^{-1}\mathbf{A}$. Dann gilt:

(i) $\mathbf{P} \geq 0$ komponentenweise

(ii) $(\mathbf{I} - \mathbf{P})\varphi = \mathbf{D}^{-1}\mathbf{A}\varphi > 0$ komponentenweise

(iii) $\mathbf{P}\varphi < \varphi$ komponentenweise

Als nächstes zeigen wir, dass $(\mathbf{I} - \mathbf{P})^{-1}$ existiert. Dazu reicht es eine Norm $\|\cdot\|_*$ auf \mathbb{R}^n zu finden, sodass $\|\mathbf{P}\|_* < 1$. Definiere $\|\cdot\|_*$ durch $\|x\|_* := \max_i \frac{|x_i|}{\varphi_i}$ und setze $x_i := \varphi_i \tilde{x}_i$ sowie $\tilde{\mathbf{P}}_{ij} := \mathbf{P}_{ij} \frac{\varphi_j}{\varphi_i}$. Dann gilt:

$$\begin{aligned} \|\mathbf{P}\|_* &= \max_{x \neq 0} \frac{\|\mathbf{P}x\|_*}{\|x\|_*} = \max_{\tilde{x} \neq 0} \frac{\|\sum_{j=1}^n \mathbf{P}_{ij}\varphi_j \tilde{x}_j\|_*}{\|\tilde{x}\|_\infty} = \max_{\tilde{x} \neq 0} \frac{\max_{i=1, \dots, n} |\sum_{j=1}^n \mathbf{P}_{ij}\varphi_j \tilde{x}_j \frac{1}{\varphi_i}|}{\|\tilde{x}\|_\infty} = \\ &= \|\tilde{\mathbf{P}}\|_\infty = \max_{i=1, \dots, n} \sum_{j=1}^n |\tilde{\mathbf{P}}_{ij}| = \max_{i=1, \dots, n} \sum_{j=1}^n \left| \mathbf{P}_{ij} \frac{\varphi_j}{\varphi_i} \right| \stackrel{(i)}{=} \max_{i=1, \dots, n} \sum_{j=1}^n \mathbf{P}_{ij} \frac{\varphi_j}{\varphi_i} = \\ &= \|\mathbf{P}\varphi\|_* \stackrel{(iii)}{<} 1 \end{aligned}$$

Damit ist (1) gezeigt, d.h. \mathbf{A}^{-1} existiert.

Zum Beweis von (2) bemerke, dass $\mathbf{A}^{-1} \geq 0$ (komponentenweise) äquivalent ist zu $\mathbf{A}^{-1}\mathbf{D} \geq 0$ (komponentenweise), denn die Diagonalmatrix \mathbf{D} hat positive Diagonaleinträge. Nun ist $\mathbf{A}^{-1}\mathbf{D} = (\mathbf{D}^{-1}\mathbf{A})^{-1} = (\mathbf{I} - \mathbf{P})^{-1} = \sum_{j=0}^{\infty} \mathbf{P}^j \geq 0$ weil $\mathbf{P} \geq 0$ elementweise. Dass $\sum_{j=0}^{\infty} \mathbf{P}^j$ konvergiert, folgt aus dem Beweis von (1).

Als letztes bleibt noch $\|\mathbf{A}^{-1}\|_\infty$ abzuschätzen. Betrachte dazu, dass aus $\mathbf{A}\varphi > 0$ folgt, dass $\mathbf{A}\varphi \geq \left[\min_{i=1, \dots, n} (\mathbf{A}\varphi)_i \right] (1, \dots, 1)^T$. Somit gilt für die Lösung \mathbf{u} von $\mathbf{A}\mathbf{u} = \mathbf{f}$, also $\mathbf{u} = \mathbf{A}^{-1}\mathbf{f}$:

$$\begin{aligned} \Rightarrow |\mathbf{u}| &\leq |\mathbf{A}^{-1}\mathbf{f}| \leq |\mathbf{A}^{-1}| |\mathbf{f}| = \mathbf{A}^{-1}|\mathbf{f}| \leq \mathbf{A}^{-1}(1, \dots, 1)^T \|\mathbf{f}\|_\infty \leq \\ &\leq \mathbf{A}^{-1}\mathbf{A}\varphi \frac{1}{\min_{i=1, \dots, n} (\mathbf{A}\varphi)_i} \|\mathbf{f}\|_\infty \leq \frac{\|\mathbf{f}\|_\infty}{\min_{i=1, \dots, n} (\mathbf{A}\varphi)_i} \varphi \end{aligned}$$

wobei die Anwendung von $|\cdot|$ immer komponentenweise erfolgt.

Also gilt $\|\mathbf{u}\|_\infty \leq \|\mathbf{f}\|_\infty \frac{\|\varphi\|_\infty}{\min_{i=1, \dots, n} (\mathbf{A}\varphi)_i}$ und damit $\|\mathbf{A}^{-1}\|_\infty \leq \frac{\|\varphi\|_\infty}{\min_{i=1, \dots, n} (\mathbf{A}\varphi)_i}$. \square

Satz 6.11 liefert also eine explizite Abschätzung für $\|\mathbf{A}^{-1}\|_\infty$. Wir illustrieren das am folgenden Beispiel.

Beispiel 6.12 Die Matrix \mathbf{A}_h aus (6.5) ist eine M -Matrix. Interpretiert man die Gitterfunktion $\varphi_h = [x(1-x)]_h$ als Vektor φ (für die inversen Knoten), dann gilt $\mathbf{A}\varphi = 2(1, \dots, 1)^T$, $\varphi \geq 0$, $\|\varphi\|_\infty \leq \frac{1}{4}$. Damit ergibt sich $\|\mathbf{A}^{-1}\|_\infty \leq \frac{\|\varphi\|_\infty}{\min_{i=1, \dots, n} (\mathbf{A}\varphi)_i} \leq \frac{\frac{1}{4}}{2} = \frac{1}{8}$.

Es ist sinnvoll, die M -Matrix-Eigenschaft durch rein algebraische Kriterien zu überprüfen. Betrachte hierzu

Definition 6.13 Eine Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ heißt

- (i) *strikt diagonal dominant*, falls $|\mathbf{A}_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |\mathbf{A}_{ij}|$, $\forall i = 1, \dots, n$.
- (ii) *schwach diagonal dominant*, falls $|\mathbf{A}_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^n |\mathbf{A}_{ij}|$, $\forall i = 1, \dots, n$ und eine echte Ungleichung für mindestens einen Index gilt.
- (iii) *reduzibel*, falls es eine Permutationsmatrix \mathbf{P} gibt, sodass $\mathbf{P}^T \mathbf{A} \mathbf{P} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{pmatrix}$ mit nichttrivialen quadratischen Matrizen \mathbf{A}_{11} und \mathbf{A}_{22} .
- (iv) *irreduzibel*, falls \mathbf{A} nicht reduzibel ist.
- (v) \mathbf{A} hat die *Ketteneigenschaft*, falls es für jedes Paar $(i, j) \in \{1, \dots, n\}^2$ eine Folge $i = i_0, i_1, \dots, i_l = j$ gibt, sodass $\mathbf{A}_{i_0, i_1}, \mathbf{A}_{i_1, i_2}, \dots, \mathbf{A}_{i_{l-1}, i_l} \neq 0$.
- (vi) \mathbf{A} heißt *irreduzibel diagonal dominant*, falls \mathbf{A} irreduzibel und schwach diagonal dominant ist.

Übung 6.14 Es gilt die Äquivalenz: \mathbf{A} ist irreduzibel $\Leftrightarrow \mathbf{A}$ hat die Ketteneigenschaft

Bemerkung 6.15 Die Ketteneigenschaft ist eine Aussage über die Besetzungsstruktur von \mathbf{A} , die mit dem Besetzungsgraphen G beschrieben werden kann. G ist wie folgt definiert:

- Die Knoten von G sind die Indizes $1, \dots, n$.
- Eine Kante von G verbindet Knoten i mit Knoten j und ist von i nach j gerichtet, falls der Matrixeintrag $\mathbf{A}_{ij} \neq 0$.

Beispiel 6.16 Sei die Besetzungsstruktur der Matrix \mathbf{A} gegeben durch

$$\mathbf{A} = \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{pmatrix}$$

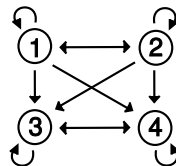


Abbildung 6.4: Besetzungsstruktur der Matrix \mathbf{A} (links) und der zugehörige Besetzungsgraph (rechts).

wobei \textcircled{i} den i -ten Knoten, $i = 1, \dots, 4$, bezeichnet.

und u_{N+1} somit aus der vorletzten Gleichung eliminiert werden. Damit ergibt sich als Lösung für u_0, \dots, u_N :

$$\begin{aligned} u_0 &= 0 \\ \frac{2u_i - u_{i+1} - u_{i-1}}{h^2} &\stackrel{!}{=} f(x_i) \quad \text{für } i = 1, \dots, N-1 \\ \frac{2(u_N - u_{N-1})}{h^2} &= f(x_N) + \frac{2g}{h} \end{aligned}$$

6.3.2 Nicht gleichmäßige Gitter in 1D

Bei nichtgleichmäßigen Gittern ist h_i , $i = 1, \dots, N$, nicht uniform. Eine Approximation an $-u''(x_i)$ ist dann z.B folgendermaßen gegeben:

$$-u''(x_i) \approx \frac{2}{h_i + h_{i+1}} \left(\frac{u_i - u_{i+1}}{h_{i+1}} - \frac{u_i - u_{i-1}}{h_i} \right), \quad i = 1, \dots, N-1$$

Diese Approximation ist allerdings nur konsistent von der Ordnung 1.

6.3.3 Krummlinige Ränder

Betrachte das Problem

$$-\Delta u = f \text{ auf } \Omega, \quad u = 0 \text{ auf } \partial\Omega$$

Die Diskretisierung in *randfernen* Punkten verursacht keine Schwierigkeiten und erfolgt mittels einer „Standard“-Diskretisierung, z.B. durch den 5-Punkt-Stern aus Abbildung 6.2.

Bei der Diskretisierung in *randnahen* Punkten tritt das Problem auf, dass das Gitter viereckig ist, aber Ω nicht, vergleiche dazu auch Abb. 6.5. Legt man das Gitter über Ω , und betrachtet nun einen randnahen Punkt dann ist es im Allgemeinen so, dass es Gitterpunkte gibt, die nicht in Ω liegen, sondern außerhalb.

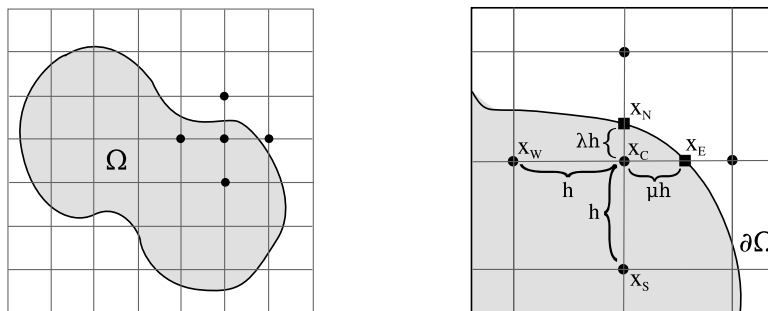


Abbildung 6.5: Ein gleichmäßiges Gitter überdeckt das Gebiet Ω mit krummlinigen Rand im Bild links. Rechts ist ein Ausschnitt mit $\partial\Omega$ zu sehen. Die runden Punkte liegen an den Schnittpunkten des gleichmäßigen Gitters. Die zwei eckigen Punkte x_N und x_E sind am $\partial\Omega$ liegenden Punkte, die anstatt der außerhalb von $\bar{\Omega}$ liegenden rund eingezeichneten Punkte zur Diskretisierung verwendet werden.

Eine Möglichkeit der Diskretisierung liefert die Verwendung von sogenannten „Shortley-Wellen“. Diese orientiert sich am 5-Punkt-Stern. Dabei werden jene Punkte, die bei Verwendung eines äquidistanten Gitters verwendet werden würden, aber außerhalb von Ω liegen, entlang der Gitterlinien an den Rand von Ω verschoben. Die Gitterpunkte werden dabei mit x_C , x_N , x_E , x_S und x_W bezeichnet. Dabei steht C für Center und N , E , S und W für North, East, South und West. Seien jetzt oBdA x_N und x_E jene Knoten, die auf $\partial\Omega$ liegen, wobei der Abstand von x_C zu x_N die Länge λh und jener von x_C zu x_E die Länge μh hat. Die Bedingungen für den Knoten x_C lauten dann

$$(1) \quad (-\Delta u)(x_C) \approx (L_h u^h)(x_C) := \frac{2}{h(1+\mu)} \left(\frac{u_C - u_E}{\mu h} - \frac{u_C - u_W}{h} \right) + \frac{2}{h(1+\lambda)} \left(\frac{u_C - u_N}{\lambda h} - \frac{u_C - u_S}{h} \right)$$

(2) setze die Randbedingung $u_N = u_E = 0$ explizit ein und

$$(3) \quad \text{fordere } (L_h u^h)(x_C) \stackrel{!}{=} f(x_C)$$

Bemerkung 6.20 Obwohl diese Diskretisierung nur Konsistenzordnung 1 hat, gilt unter geeigneten Voraussetzungen $\| [u]_h - u^h \|_\infty \leq Ch^2$.