

# Zusammenfassung Numerik von PDEs

© Tim Baumann, <http://timbaumann.info/uni-spicker>

**Def.** Sei  $\Omega \subseteq \mathbb{R}^n$  offen. Eine DGL der Form

$$F(x, u, Du, \dots, D^k u) = 0$$

heißt **partielle DGL/PDE** der Ordnung  $k \geq 1$ , wobei

$$F : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \dots \times \mathbb{R}^{n^k} \rightarrow \mathbb{R}$$

eine gegebene Funktion und  $u : \Omega \rightarrow \mathbb{R}$  gesucht ist.

**Def (Klassifikation von PDEs).**

- Die PDE heißt **linear**, wenn sie die Form

$$\sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha u = f(x)$$

mit Funktionen  $a_\alpha, f : \Omega \rightarrow \mathbb{R}$  besitzt.

- Die PDE heißt **semilinear**, wenn sie die Form

$$\sum_{|\alpha| = k} a_\alpha(x) D^\alpha u + a_0(x, u, Du, \dots, D^{k-1} u) = 0$$

besitzt, wobei  $a_\alpha : \Omega \rightarrow \mathbb{R}$  und  $a_0 : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \dots \times \mathbb{R}^{n^k} \rightarrow \mathbb{R}$  gegeben sind.

- Die PDE heißt **quasilinear**, wenn sie die Form

$$\sum_{|\alpha| = k} a_\alpha(x, u, Du, \dots, D^{k-1} u) D^\alpha u + a_0(x, u, Du, \dots, D^{k-1} u) = 0$$

hat, wobei  $a_\alpha, a_0 : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \dots \times \mathbb{R}^{n^k}$  gegeben sind.

- Die PDE heißt **nichtlinear**, falls die Ableitungen der höchsten Ordnung nicht linear vorkommen.

**Def.** Sei  $\Omega \subseteq \mathbb{R}^n$  offen und  $F : \Omega \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  eine gegebene Funktion. Eine PDE der Form

$$F(x, u, \partial_{x_1} u, \dots, \partial_{x_n} u, \partial_{x_1} \partial_{x_1} u, \dots, \partial_{x_1} \partial_{x_n} u, \dots, \partial_{x_n} \partial_{x_n} u) = 0$$

heißt **PDE zweiter Ordnung**.

**Notation.**  $p_i := \partial_{x_i} u$ ,  $p_{ij} := \partial_{x_i x_j}^2 u$

$$M(x) := \begin{pmatrix} \frac{\partial F}{\partial p_{11}} & \cdots & \frac{\partial F}{\partial p_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial p_{n1}} & \cdots & \frac{\partial F}{\partial p_{nn}} \end{pmatrix} = M(x)^T.$$

**Def (Typeneinteilung für PDEs der 2. Ordnung).**

Obige PDE zweiter Ordnung heißt

- elliptisch** in  $x$ , falls die Matrix  $M(x)$  positiv o. definit ist.
- parabolisch** in  $x$ , falls genau ein EW von  $M(x)$  gleich null ist und alle anderen dasselbe Vorzeichen haben.
- hyperbolisch** in  $x$ , falls genau ein EW ein anderes Vorzeichen als die anderen EWE hat.

## Lösungstheorie elliptischer PDEs

**Def.** Sei  $\Omega \subset \mathbb{R}^n$  offen, zusammenhängend und beschränkt.

- $\mathcal{C}(\overline{\Omega}, \mathbb{R}^m) := \{u : \overline{\Omega} \rightarrow \mathbb{R}^m \mid u \text{ stetig}\}$ ,  $\mathcal{C}(\overline{\Omega}) := \mathcal{C}(\overline{\Omega}, \mathbb{R})$ , mit Norm

$$\|u\|_{\mathcal{C}(\overline{\Omega}, \mathbb{R}^m)} = \sup_{x \in \overline{\Omega}} \|u(x)\|. \quad (\text{Supremumsnorm})$$

- $\mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m)$ ,  $k \in \mathbb{N}$  ist der Raum aller auf  $\Omega$   $k$ -mal stetig diff'baren Funktionen  $u : \Omega \rightarrow \mathbb{R}^m$ , die zusammen mit ihren Ableitungen bis zur Ordnung  $k$  stetig auf  $\overline{\Omega}$  fortgesetzt werden können.

$$\|u\|_{\mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m)} = \sum_{|\alpha| \leq k} \|D^\alpha u\|_{\mathcal{C}(\overline{\Omega}, \mathbb{R}^m)}$$

- Sei  $\alpha \in [0, 1)$ .  $\mathcal{C}^{0, \alpha}(\overline{\Omega}, \mathbb{R}^m) = \{u \in \mathcal{C}(\overline{\Omega}, \mathbb{R}^m) \mid H_\alpha(u, \overline{\Omega}) < \infty\}$  mit

$$H_\alpha(u, \overline{\Omega}) := \sup_{x, y \in \overline{\Omega}, x \neq y} \frac{\|u(x) - u(y)\|}{\|x - y\|^\alpha} \quad (\text{Hölder-Koeffizient})$$

heißt **Raum der glm. Hölder-stetigen Fktn** zum Exponent  $\alpha$ . Der Hölder-Koeffizient ist dabei eine Seminorm auf  $\mathcal{C}^{0, \alpha}(\overline{\Omega}, \mathbb{R}^m)$ .

- $\mathcal{C}^{k, \alpha}(\overline{\Omega}, \mathbb{R}^m) := \{u \in \mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m) \mid \forall |\gamma| = k : D^\gamma u \in \mathcal{C}^{0, \alpha}(\overline{\Omega}, \mathbb{R}^m)\}$  heißt **Hölder-Raum**. Eine Norm ist gegeben durch

$$\|u\|_{\mathcal{C}^{k, \alpha}(\overline{\Omega}, \mathbb{R}^m)} := \|u\|_{\mathcal{C}^k(\overline{\Omega}, \mathbb{R}^m)} + \sum_{|\gamma| = k} H_\alpha(D^\gamma u, \overline{\Omega}).$$

**Bem.** • Jede Hölder-stetige Funktion ist gleichmäßig stetig.

- $\mathcal{C}^{0, 1}(\overline{\Omega}, \mathbb{R}^m)$  heißt **Raum der Lipschitz-stetigen Funktionen**.
- $\mathcal{C}$ ,  $\mathcal{C}^k$  und  $\mathcal{C}^{k, \alpha}$  sind Banach-Räume mit den jeweiligen Normen.

**Def.** Sei  $\Omega \subset \mathbb{R}^n$  offen, zusammenhängend und beschränkt. Das Gebiet  $\Omega$  gehört zur **Klasse  $\mathcal{C}^{k, \alpha}$** , wenn in jedem Punkt  $x \in \partial\Omega$  eine Umgebung in  $\partial\Omega$  existiert, die sich in einem geeigneten Koordinatensystem als ein Graph einer Funktion aus  $\mathcal{C}^{k, \alpha}$  darstellen lässt und  $\Omega$  lokal immer auf einer Seite von  $\partial\Omega$  liegt.

**Satz (Gauß'scher Integralsatz).** Sei  $\Omega \subset \mathbb{R}^n$  ein Lipschitz-Gebiet und  $u \in \mathcal{C}(\overline{\Omega}, \mathbb{R}^n) \cap \mathcal{C}^1(\Omega, \mathbb{R}^n)$ . Dann gilt

$$\int_\Omega \operatorname{div} u \, dx = \int_\Omega \sum_{i=1}^n \frac{\partial u_i}{\partial x_i} \, dx = \int_{\partial\Omega} \sum_{i=1}^n u_i \nu_i \, d\rho(x) = \int_{\partial\Omega} u \cdot \nu \, d\rho(x),$$

wobei  $\nu$  der äußere Normalenvektor an den Rand von  $\Omega$  ist.

**Problem.** Wir betrachten das Randwertproblem

$$(\text{RWP}) \begin{cases} \mathcal{L}u &= f & \text{in } \Omega & (\text{PDE}) \\ \mathcal{R}u &= g & \text{auf } \partial\Omega & (\text{Randbedingung}) \end{cases}$$

wobei  $\mathcal{L}$  der lineare Differentialoperator

$$\mathcal{L}u = - \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(x) \frac{\partial u}{\partial x_i} + c(x)u$$

mit Fktn  $a_{ij}, b_i, c, f : \Omega \rightarrow \mathbb{R}$ ,  $g : \partial\Omega \rightarrow \mathbb{R}$  ist, sodass  $A(x) := (a_{ij}(x))$  symmetrisch ist. Als Randbedingung (RB) verlangen wir:

$$\begin{array}{lll} \text{Dirichlet-RB:} & u &= g \quad \text{auf } \partial\Omega, \\ \text{Neumann-RB:} & (A(x)\nabla u) \cdot \nu &= g \quad \text{auf } \partial\Omega \text{ oder} \\ \text{Robin-RB:} & (A(x)\nabla u) \cdot \nu + \delta u &= g \quad \text{auf } \partial\Omega. \end{array}$$

**Bem.** Man kann auch auf verschiedenen Teilstücken des Randes verschiedene Bedingungen stellen.

**Bem.** Falls die Funktionen  $a_{ij}$  differenzierbar sind, so kann  $\mathcal{L}$  in **Divergenzform** geschrieben werden:

$$\begin{aligned} \mathcal{L}u &= - \sum_{i,j=1}^n \frac{\partial}{\partial x_j} \left( a_{ij}(x) \frac{\partial u}{\partial x_i} \right) + \sum_{i=1}^n \underbrace{\left( \left( \sum_{j=1}^n \frac{\partial}{\partial x_j} a_{ij}(x) \right) + b_i(x) \right)}_{\tilde{b}(x)} \frac{\partial u}{\partial x_i} + c(x)u \\ &= - \operatorname{div}(A(x)\nabla u) + \tilde{b}(x) \cdot \nabla u + c(x)u \end{aligned}$$

**Voraussetzung.** Wir nehmen im Folgenden an:

- $\mathcal{L}$  ist **gleichmäßig elliptisch**, d. h.

$$\exists \lambda_0 > 0 : \forall \xi \in \mathbb{R}^n : \forall x \in \Omega : \xi^T A(x) \xi \geq \lambda_0 \|\xi\|^2$$

Dabei heißt  $\lambda_0$  **Elliptizitätskonstante**.

- $a_{ij}, b_i, c, f \in \mathcal{C}(\overline{\Omega})$ ,  $g \in \mathcal{C}(\partial\Omega)$

**Bem.**  $\mathcal{L} = f$  ist elliptisch auf  $\Omega \iff A(x) > 0$  (spd) für alle  $x \in \Omega$

**Def.** Eine Fkt  $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$  heißt **klassische Lsg** vom (RWP) mit  $\mathcal{R}u := u$ , wenn die beiden Gleichungen in (RWP) in jedem Punkt von  $\Omega$  bzw. des Randes  $\partial\Omega$  erfüllt sind.

**Satz (Maximumsprinzip).** Sei  $\Omega \subset \mathbb{R}^n$  offen, zshgd u. beschränkt. Sei  $u \in \mathcal{C}^2(\omega) \cap \mathcal{C}(\overline{\Omega})$  eine Lösung vom (RWP),  $f \leq 0$  in  $\Omega$  und  $c \equiv 0$ . Dann nimmt  $u$  sein Maximum auf dem Rand  $\partial\Omega$  an, d. h.

$$\sup_{x \in \overline{\Omega}} u(x) = \sup_{x \in \partial\Omega} u(x) = \sup_{x \in \partial\Omega} g(x)$$

**Kor.** Sei  $c \geq 0$  und  $f \leq 0$ . Dann gilt  $\sup_{x \in \overline{\Omega}} u(x) \leq \max\{\sup_{x \in \partial\Omega} u(x), 0\}$ .

**Kor (Vergleichsprinzip).** Für  $u_1, u_2 \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$  und  $c \geq 0$  gelte  $\mathcal{L}u_1 \leq \mathcal{L}u_2$  in  $\Omega$  und  $u_1 \leq u_2$  auf  $\partial\Omega$ . Dann gilt  $u_1 \leq u_2$  auf  $\overline{\Omega}$ .

**Kor (Eindeutigkeit).** Sei  $c \geq 0$ . Dann hat (RWP) höchstens eine Lösung  $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ .

**Satz.** Sei  $\Omega$  ein beschr. Lipschitz-Gebiet,  $a_{ij}, b_i, c, f \in \mathcal{C}(\overline{\Omega})$ ,  $c \geq 0$ ,  $g \in \mathcal{C}(\partial\Omega)$ . Dann besitzt (RWP) genau eine Lsg  $u \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ .

**Achtung.** Es muss aber nicht  $u \in \mathcal{C}^2(\overline{\Omega})$  gelten!

# Differenzenverfahren

**Verfahren (DV).** Am Beispiel des Poisson-Problems

$$(RWP_1) \quad \begin{cases} -\Delta u = f & \text{in } \Omega = (0, 1) \\ u(0) = g_0, u(1) = g_1 & \text{auf } \partial\Omega \end{cases}$$

Wir führen folgende Schritte durch:

1. Diskretisierung: Wähle  $n \in \mathbb{N}$ , setze  $h := \frac{1}{n}$  und

$$\begin{aligned} \Omega_h &:= \{x_i := ih \mid i = 1, \dots, n-1\} & (\text{innere Gitterpunkte}) \\ \partial\Omega_h &:= \{x_0 = 0, x_n = 1\} & (\text{Randpunkte}) \end{aligned}$$

2. Approx. der Ableitungen durch Differenzenquotienten (DQ)

$$\begin{aligned} u'(x_i) &\approx \frac{1}{h} (u(x_i + h) - u(x_i)) & (\text{Vorwärts-DQ}) \\ u'(x_i) &\approx \frac{1}{h} (u(x_i) - u(x_i - h)) & (\text{Rückwärts-DQ}) \\ u'(x_i) &\approx \frac{1}{2h} (u(x_i + h) - u(x_i - h)) & (\text{zentraler DQ}) \end{aligned}$$

Für die zweite Ableitung ergibt sich

$$\begin{aligned} u''(x_i) &= (u'(x_i))' \approx \frac{1}{h} (u'(x_i + h) - u'(x_i)) \approx \\ &\approx \frac{1}{h} \cdot \left( \frac{1}{h} (u(x_i + h) - u(x_i)) - \frac{1}{h} (u(x_i) - u(x_i - h)) \right) \\ &= \frac{1}{h^2} (u(x_i + h) - 2 \cdot u(x_i) + u(x_i - h)) =: \Delta_h u \end{aligned}$$

Dabei heißt  $\Delta_h$  der diskrete eindim. Laplace-Operator. Das diskretisierte Randwertproblem ist nun

$$(RWP_1)_h \quad \begin{cases} -\Delta_h u_h = f & \text{in } \Omega_h, \\ u_h(0) = g_0, u_h(1) = g_1 & \text{auf } \partial\Omega_h. \end{cases}$$

3. Aufstellen des linearen Gleichungssystems

$$\begin{aligned} \frac{1}{h^2} (2u_h(x_1) - u_h(x_2)) &= f(x_1) + \frac{g_0}{h^2} & (i=1) \\ \frac{1}{h^2} (-u_h(x_{i-1}) + u_h(x_i) - u_h(x_{i+1})) &= f(x_i) & (i=2, \dots, n-2) \\ \frac{1}{h^2} (-u_h(x_{n-2}) + 2u_h(x_{n-1})) &= f(x_{n-1}) + \frac{g_1}{h^2} & (i=n-1) \end{aligned}$$

Als lineares Gleichungssystem:  $-\tilde{\Delta}_h \tilde{u}_h = \tilde{f}_h$  mit

$$-\tilde{\Delta}_h = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ 0 & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{(n-1) \times (n-1)},$$

$$\tilde{u}_h = \begin{pmatrix} u_h(x_1) \\ \vdots \\ u_h(x_{n-1}) \end{pmatrix}, \quad \tilde{f}_h = \begin{pmatrix} f(x_1) + \frac{g_0}{h^2} \\ f(x_2) \\ \vdots \\ f(x_{n-2}) \\ f(x_{n-1}) + \frac{g_1}{h^2} \end{pmatrix}$$

**Ziel.** Herausfinden, was die Lösung  $u_h$  von  $(RWP)_h$  (die man durch Lösen von (LGS) erhält) mit der Lösung  $u$  zum ursprünglichen Problem (RWP) zu tun hat. Ist etwa  $u_h$  die Einschränkung von  $u$ , oder zumindest annäherungsweise? Wenn ja, wie klein muss man  $h$  wählen, damit die Approximation gut wird?

$$\begin{aligned} (RWP) \quad & \begin{cases} -\mathcal{L}u = f & \text{in } \Omega, \\ u = g & \text{auf } \partial\Omega \end{cases} \\ (RWP)_h \quad & \begin{cases} -\mathcal{L}_h u = f_h & \text{in } \Omega_h, \\ u_h = g_h & \text{auf } \partial\Omega_h \end{cases} \\ (LGS) \quad & \tilde{\mathcal{L}}_h \tilde{u}_h = \tilde{f}_h \end{aligned}$$

**Notation.**  $U_h := \{\Omega_h \rightarrow \mathbb{R}\}$ ,  $R_h : \mathcal{C}(\bar{\Omega}) \rightarrow U_h$ ,  $u \mapsto u|_{\Omega_h}$

**Def.** Das Differenzenverfahren  $(RWP)_h$  heißt

- **konvergent** von der Ordnung  $p$ , falls  $C > 0$ ,  $h_0 > 0$  existieren, sodass für die Lösung  $u$  von (RWP) und die Lösung  $u_h$  von  $(RWP)_h$  gilt:
$$\|u_h - R_h u\|_h \leq Ch^p \quad \text{für alle } 0 < h \leq h_0,$$
wobei  $\|\cdot\|_h$  eine Norm zu  $U_h$  ist, wie z. B.  $\|u_h\|_h := \max_{x \in \Omega_h} |u_h(x)|$ .
- **konsistent** von der Ordnung  $p$ , falls
$$\|\mathcal{L}_h R_h u - R_h \mathcal{L} u\|_h \leq ch^p \|u\|_{\mathcal{C}^{p+2}(\bar{\Omega})} \quad \forall u \in \mathcal{C}^{p+2}(\bar{\Omega}).$$
- **stabil**, falls  $\tilde{L}_h$  invertierbar ist und ein  $h_0 > 0$  existiert mit

$$\sup_{0 < h \leq h_0} \|\tilde{\mathcal{L}}_h^{-1}\|_h < \infty, \quad \text{wobei } \|\tilde{\mathcal{L}}_h^{-1}\|_h := \sup_{f \neq 0} \frac{\|\tilde{\mathcal{L}}_h^{-1} f\|_h}{\|f\|_h}.$$

**Bem.** Die ind. Matrixnorm ist  $\|\tilde{\mathcal{L}}_h^{-1}\|_h = \|\tilde{\mathcal{L}}_h^{-1}\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |l_{ij}|$ .

**Satz.** Ist das DV  $(RWP)_h$  konsistent und stabil, so auch konvergent. Genauer gilt: Ist  $(RWP)_h$  stabil und konsistent von der Ordnung  $p$  und  $u \in \mathcal{C}^{p+2}(\bar{\Omega})$ , dann ist  $(RWP)_h$  konvergent von der Ordnung  $p$ .

**Beweis.** Setze  $w_h := u_h - R_h u$ . Für  $x \in \partial\Omega_h$  gilt dann  $w_h(x) = 0$  und für  $x \in \Omega_h$  gilt

$$\begin{aligned} \tilde{\mathcal{L}}_h w_h(x) &= \mathcal{L}_h w_h(x) = \mathcal{L}_h u_h(x) - \mathcal{L}_h R_h u(x) \\ &= f_h(x) - \mathcal{L}_h R_h u(x) = R_h f(x) - \mathcal{L}_h R_h u(x) \\ &= R_h \mathcal{L} u(x) - \mathcal{L}_h R_h u(x) \end{aligned}$$

Somit gilt  $w_h = \tilde{\mathcal{L}}_h^{-1} (R_h \mathcal{L} u - \mathcal{L}_h R_h u)$  in  $\Omega_h$ , also

$$\begin{aligned} \|w_h\|_h &= \|\tilde{\mathcal{L}}_h^{-1} (R_h \mathcal{L} u - \mathcal{L}_h R_h u)\| \leq \|\tilde{\mathcal{L}}_h^{-1}\|_h \cdot \|R_h \mathcal{L} u - \mathcal{L}_h R_h u\|_h \\ &\leq c_1 \cdot c_2 \cdot h^p \cdot \|u\|_{\mathcal{C}^{p+2}(\bar{\Omega})} \leq Ch^p \quad \text{für } 0 < h \leq h_0. \end{aligned} \quad \square$$

**Lem.** Das DV  $(RWP_1)_h$  ist konsistent von der Ordnung 2. Es gilt

$$\|\Delta_h R_h u - R_h \Delta u\|_h \leq \frac{1}{12} \|u\|_{\mathcal{C}^4(\bar{\Omega})} h^2 \quad \forall u \in \mathcal{C}^4(\bar{\Omega}).$$

**Bem.** Um zu zeigen, dass  $(RWP_1)_h$  konvergent ist, müssen wir noch zeigen, dass  $\tilde{L}_h = -\tilde{\Delta}_h$  invertierbar ist und  $\sup_{0 < h \leq h_0} \|\tilde{\Delta}_h\| < \infty$ .

**Def.** Eine Matrix  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  heißt **M-Matrix**, falls  
a)  $a_{ii} > 0$  für  $i = 1, \dots, n$ , b)  $a_{ij} \leq 0$  für  $i \neq j$ ,  $i, j = 1, \dots, n$ ,  
c)  $A$  invertierbar ist und d) für  $A^{-1} =: B = (b_{ij})$  gilt  $b_{ij} \geq 0$ .

**Lem.** Erfülle  $A \in \mathbb{R}^{n \times n}$  die Bedingungen a) und b). Zerlege  $A = D + L + R$  in eine Diagonalmatrix und strikte untere/obere Dreiecksmatrizen. Dann ist  $A$  genau dann eine M-Matrix wenn

$$\rho(D^{-1}(L + R)) < 1.$$

**Bem.** Es gilt folgende Monotonie-Eigenschaft für M-Matrizen:

$$x \leq y \implies A^{-1}x \leq A^{-1}y.$$

**Def.** Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt **reduzibel** (oder zerlegbar), wenn es eine Permutationsmatrix  $P \in \mathbb{R}^{n \times n}$  gibt, sodass

$$PAP^T = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \quad \text{mit } A_{11} \in \mathbb{R}^{k \times k}, 0 < k < n.$$

**Lem** (Gerschgorin). Alle EWe einer Matrix  $A = (a_{ij}) \in \mathbb{C}^{n \times n}$  liegen in der Menge

$$\bigcup_{i=1}^n \overline{B_{r_i}(a_{ii})} \quad \text{mit } r_i := \sum_{j=1}^n |a_{ij}|.$$

Falls  $A$  irreduzibel ist, so liegen sie sogar in

$$\left( \bigcup_{i=1}^n B_{r_i}(a_{ii}) \right) \cup \left( \bigcap_{i=1}^n \partial B_{r_i}(a_{ii}) \right)$$

**Def.** Sei  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  eine Matrix.

- $A$  heißt **schwach diagonaldominant**, falls

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \leq |a_{ii}| \quad \text{für } i = 1, \dots, n$$

und ein  $i_0$  existiert, sodass die Ungleichung strikt ist.

- $A$  heißt **diagonaldominant**, falls

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| < |a_{ii}| \quad \text{für } i = 1, \dots, n$$

- $A$  heißt **irreduzibel diagonaldominant**, falls  $A$  irreduzibel und schwach diagonaldominant ist.

**Lem.** Sei  $A = (a_{ij}) \in \mathbb{R}^{n \times n}$  eine Matrix mit  $a_{ii} > 0$ ,  $i = 1, \dots, n$  und  $a_{ij} \leq 0$ ,  $i, j = 1, \dots, n$ ,  $i \neq j$ , die diagonaldominant oder irreduzibel diagonaldominant ist. Dann ist  $A$  eine M-Matrix.

**Bem.**  $-\tilde{\Delta}_h$  ist irreduzibel diagonaldominant, also eine M-Matrix.

**Lem.** Sei  $A$  eine irreduzible M-Matrix. Dann gilt  $A^{-1} > 0$ .

**Lem.** Sei  $A \in \mathbb{R}^{n \times n}$  eine M-Matrix und es existiere ein Vektor  $v$ , sodass  $(Av)_j \geq 1$ ,  $j = 1, \dots, n$ . Dann gilt  $\|A^{-1}\|_\infty \leq \|v\|_\infty$ .

**Lem.**  $\|\tilde{\Delta}_h^{-1}\|_\infty \leq \frac{1}{8}$

**Satz.** Das DV  $(RWP_1)_h$  ist konvergent von der Ordnung 2, falls die Lösung von  $(RWP_1)$  zu  $\mathcal{C}^4([0, 1])$  gehört. Es gilt die Abschätzung

$$\|u_h - R_h u\|_\infty \leq \frac{h^2}{96} \|u\|_{\mathcal{C}^4([0, 1])}.$$

**Problem.** Wir betrachten nun

$$(RWP_2) \quad \begin{cases} -\Delta u &= f & \text{in } \Omega = (0, 1) \times (0, 1) \\ u &= g & \text{auf } \partial\Omega \end{cases}$$

1. Diskretisierung: Setze  $h := \frac{1}{n}$ ,  $n \in \mathbb{N}$  und

$$\Omega_h := \{(x, y) \in \Omega \mid x = ih, y = jh, i, j = 1, \dots, n-1\}$$

$$\partial\Omega_h := \{(x, y) \in \partial\Omega \mid x = ih, y = jh, i, j = 1, \dots, n-1\}$$

2. Approximation der Ableitungen

$$-\Delta u(x, y) = -\frac{\partial^2 u}{\partial x^2}(x, y) - \frac{\partial^2 u}{\partial y^2}(x, y)$$

$$\approx -\frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2} - \frac{u(x, y+h) - 2u(x, y) + u(x, y-h)}{h^2}$$

$$= -\frac{u(x+h, y) + u(x-h, y) - 4u(x, y) + u(x, y+h) + u(x, y-h)}{h^2} =: -\Delta_h u$$

Dabei hat der diskrete Laplace-Operator  $\Delta_h$  die Form eines Differenzensterns. Gesucht ist die Lsg  $u_h : \Omega_h \cup \partial\Omega_h \rightarrow \mathbb{R}$  von

$$(\text{RWP}_2)_h \begin{cases} -\Delta_h u_h = f_h & \text{in } \Omega_h \\ u_h = g & \text{auf } \partial\Omega_h. \end{cases}$$

3. Aufstellen des linearen Gleichungssystems  $-\tilde{\Delta}_h \tilde{u}_h = f_h$ :

$$\tilde{u}_h = \begin{pmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{n-1, n-2} \\ u_{n-1, n-1} \end{pmatrix} \in \mathbb{R}^{(n-1)^2},$$

$$-\tilde{\Delta}_h = \frac{1}{h^2} \begin{pmatrix} A & -I & & 0 \\ -I & A & -I & \\ & \ddots & \ddots & \ddots \\ 0 & & -I & A \end{pmatrix} \in \mathbb{R}^{(n-1)^2 \times (n-1)^2},$$

$$A = \begin{pmatrix} 4 & -1 & & 0 \\ -1 & 4 & -1 & \\ & \ddots & \ddots & \ddots \\ 0 & & -1 & 4 \end{pmatrix} \in \mathbb{R}^{n-1 \times n-1}$$

**Lem.** Das DV  $(\text{RWP}_2)_h$  ist konsistent von der Ordnung 2. Es gilt

$$\|\Delta_h R_h u - R_h \Delta u\|_h \leq \frac{1}{6} \|u\|_{C^2(\bar{\Omega})} h^2.$$

**Lem.** Das DV  $(\text{RWP}_2)_h$  ist stabil. Es gilt  $\|\tilde{\Delta}_h^{-1}\|_\infty \leq 1/8$ .

**Satz.** Das DV  $(\text{RWP}_2)_h$  ist konvergent von der Ordnung 2, falls die Lösung von  $(\text{RWP}_2)$  zu  $C^4(\bar{\Omega})$  gehört. Es gilt

$$\|u_h - R_h u\|_h \leq 1/48 \|u\|$$

*Bem.* Durch die Einbeziehung weiterer Gitterpunkte zur Approximation des Differentialoperators lässt sich die Konvergenzordnung erhöhen:

$$-\Delta_h^{(9)} u(x, y) = \frac{1}{12h^2} (u(x-2h, y) - 16u(x-h, y) + 30u(x, y) - 16u(x+h, y) + u(x+2h, y) + u(x, y-2h) - 16u(x, y-h) + 30u(x, y) - 16u(x, y+h) + u(x, y+2h)) \approx -\Delta u(x, y)$$

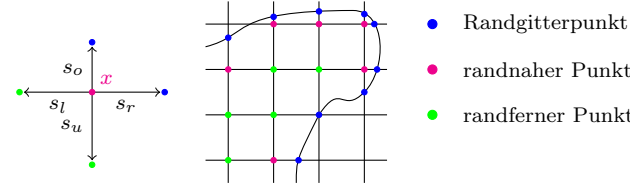
Damit erreicht man die Konsistenzordnung 4.

**Situation.** Sei  $\Omega \subset \mathbb{R}^2$  beschränkt.

**Def.** •  $\Omega_h := \{x, y \in \Omega \mid x/h, y/h \in \mathbb{Z}\}$  heißen **innere Gitterpunkte**.

- Ein Punkt  $z_R \in \partial\Omega$  heißt **Randgitterpunkt** (notiert  $z_R \in \partial\Omega_h$ ), falls es einen inneren Gitterpunkt  $z \in \Omega_h$  gibt, sodass  $z_R = r + \alpha h e_1$  oder  $z_R = z + \alpha h e_2$  mit  $|\alpha| \leq 1$ . Die Nachbarn  $N(x, y)$  eines Punktes  $(x, y)$  sind  $(x + s_r h, y)$ ,  $(x - s_l h, y)$ ,  $(x, y + s_o h)$ ,  $(x, y - s_u h)$ , falls  $s_r, s_l, s_o, s_u \in (0, 1]$  und die Verbindungsstrecken zu  $(x, y)$  in  $\Omega$  liegen.
- Ein Punkt  $(x, y) \in \Omega_h$  heißt **randnah**, falls  $(x, y)$  die Nachbarn  $(x - s_l h, y)$ ,  $(x + s_r h, y)$ ,  $(x, y - s_u h)$ ,  $(x, y + s_o h)$  hat mit mindestens einem  $s_i < 1$ . Ansonsten heißt  $(x, y)$  **randfern**.

**Notation.** Wir haben eine Einteilung  $\Omega_h = \Omega_h^{\text{rn}} \sqcup \Omega_h^{\text{rf}}$  der Gitterpunkte in randnahe und randferne Punkte.



**Lem (Dividierte Differenzen von Newton).**

Für  $u \in C^3([x_l, x_r])$ ,  $x \in (x_l, x_r)$  gilt

$$u''(x) = \frac{2}{x_r - x_l} \left( \frac{u(x_r) - u(x)}{x_r - x} - \frac{u(x) - u(x_l)}{x - x_l} \right) + \mathcal{O}(x_r - x_l)$$

$$= \frac{2}{x_r - x_l} \left( \frac{1}{x_r - x} u(x_r) + \frac{1}{x - x_l} u(x_l) \right) - \frac{2}{(x_r - x)(x - x_l)} u(x)$$

**Verfahren (Shortley-Weller-Diskretisierung).**

Dadurch inspiriert approximieren wir den Laplace-Operator durch

$$\mathcal{D}_h u(x, y) = \frac{1}{h^2} \left( \frac{2u(x - s_l h, y)}{s_l(s_r + s_l)} + \frac{2u(x + s_r h, y)}{s_r(s_r + s_l)} + \frac{2u(x, y - s_u h)}{s_u(s_o + s_u)} + \frac{2u(x, y + s_o h)}{s_o(s_o + s_u)} - \left( \frac{2}{s_l s_r} + \frac{2}{s_o s_u} \right) u(x, y) \right)$$

wobei  $x_r - x = s_r h$ ,  $x - x_l = s_l h$ ,  $y_o - y = s_o h$ ,  $y - y_u = s_u h$ . Wir betrachten nun

$$(\text{RWP}_2)'_h \begin{cases} -\mathcal{D}_h u_h = f_h & \text{in } \Omega_h \\ u_h = g & \text{auf } \partial\Omega_h \end{cases}$$

$$(\text{LGS}_2)' \begin{cases} -\tilde{\mathcal{D}}_h \tilde{u}_h = \tilde{f}_h \\ \tilde{f}_h = f_h + g_h \end{cases}$$

$$\text{mit } g_h(x, y) = \frac{1}{h^2} \sum_{(x_N, y_N) \in N(x, y) \cap \partial\Omega_h} S_{x_N, y_N} g(x_N, y_N)$$

wobei

$$S_{x_N, y_N} := \begin{cases} 2/s_l(s_l + s_r) & \text{falls } (x_N, y_N) = (x - s_l h, y), \\ 2/s_r(s_l + s_r) & \text{falls } (x_N, y_N) = (x + s_r h, y), \\ 2/s_o(s_o + s_u) & \text{falls } (x_N, y_N) = (x, y + s_o h), \\ 2/s_u(s_o + s_u) & \text{falls } (x_N, y_N) = (x, y - s_u h), \end{cases}$$

$$-\tilde{\mathcal{D}}_h = (d_{ij}) \quad \text{mit} \quad d_{ii} = 1/h^2 \left( \frac{2}{s_{il}s_{ir}} + \frac{2}{s_{iu}s_{io}} \right) \quad \text{und}$$

$$d_{ij} = 1/h^2 \begin{cases} -2/s_{il}(s_{il} + s_{ir}) & \text{falls } j \text{ der linke Nachbar von } i \text{ ist,} \\ -2/s_{ir}(s_{il} + s_{ir}) & \text{falls } j \text{ der rechte Nachbar von } i \text{ ist,} \\ -2/s_{iu}(s_{iu} + s_{io}) & \text{falls } j \text{ der untere Nachbar von } i \text{ ist,} \\ -2/s_{io}(s_{iu} + s_{io}) & \text{falls } j \text{ der obere Nachbar von } i \text{ ist.} \end{cases}$$

**Lem.** • Die Matrix  $-\tilde{\mathcal{D}}_h$  ist eine M-Matrix.

- Sei  $\Omega \subset \mathbb{R}^2$  beschränkt und gehöre zu dem Streifen  $(x_0, x_0 + d) \times \mathbb{R}$  oder  $\mathbb{R} \times (y_0, y_0 + d)$ . Dann gilt  $\|\tilde{\mathcal{D}}_h^{-1}\| \leq d^2/s$ .

*Bem.* Das DV  $(\text{RWP}_2)'_h$  hat in den randnahen Punkten nur die Konsistenzordnung 1. Dennoch gilt:

**Satz.** Sei  $\Omega \subset \mathbb{R}^3$  beschränkt und Teilmenge des Streifens  $(x_0, x_0 + d) \times \mathbb{R}$  oder  $\mathbb{R} \times (y_0, y_0 + d)$ . Dann ist das Verfahren  $(\text{RWP}_2)'_h$  konvergent von der Ordnung 2. Es gilt

$$\|u_h - R_h u\|_h \leq (1/3h^3 + d^2/48h^2) \|u\|_{C^4(\bar{\Omega})}.$$

**Idee.** Bestimme den Wert von  $u$  bei randnahen Punkten  $(x, y)$  durch lineare Interpolation:

- $u(x, y) \approx \frac{s_r}{s_r + s_l} u(x - s_l h, y) + \frac{s_l}{s_r + s_l} u(x + s_r h, y)$
- $u(x, y) \approx \frac{s_o}{s_u + s_o} u(x, y - s_u h) + \frac{s_u}{s_u + s_o} u(x, y + s_o h)$

$$(\text{RWP}_2)''_h \begin{cases} -\mathcal{D}_h u = f_h & \text{in } \Omega_h \\ u_h = g & \text{auf } \partial\Omega_h \end{cases}$$

$$(\text{LGS}_2)'' - \tilde{\mathcal{D}}_h \tilde{u}_h = \tilde{f}_h$$

**Lem.** Dieses Verfahren besitzt Konsistenzordnung (und somit Konvergenzordnung) 2.

$$(\text{RWP}_3) \begin{cases} -\mathcal{L}u = f & \text{in } \Omega = (0, 1) \times (0, 1) \\ u = g & \text{auf } \partial\Omega \end{cases}$$

mit

$$-\mathcal{L}u = -(a_{11}(x, y)u_{xx} + 2a_{12}(x, y)u_{xy} + a_{22}(x, y)u_{yy}) + b_1(x, y)u_x + b_2(x, y)u_y$$

wobei  $c(x, y) \leq 0$ ,  $\xi^T A(x, y) \xi \geq \lambda \|\xi\|^2$  und

$$A(x, y) = \begin{pmatrix} a_{11}(x, y) & a_{12}(x, y) \\ a_{21}(x, y) & a_{22}(x, y) \end{pmatrix}$$

1. Diskretisierung:  $h = 1/n$ ,  $\Omega_h$ ,  $\partial\Omega_h$  wie früher.

2. Approximation:  $u_{xx}(x, y) \approx \frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2}$ ,

$$u_{yy}(x, y) \approx \dots, u_x(x, y) \approx \frac{u(x+h, y) - u(x-h, y)}{2h},$$

$$u_y(x, y) \approx \frac{u(y+h) - u(y-h)}{2h},$$

Mehrere Möglichkeiten für Approx. von  $u_{xy}$ :

- $u_{xy}(x,y) \approx \frac{u(x+h,y+h) - u(x+h,y-h) - u(x-h,y+h) + u(x-h,y-h)}{4h^2}$
- $\frac{1}{2h^2} (-u(x-h,y) + 2u(x,y) - u(x+h,y) + u(x-h,y-h) - u(x,y+h) + u(x+h,y+h))$   
falls  $a_{12}(x,y) \geq 0$ ,  $u_{xy}(x,y) \approx \frac{1}{2h^2} (u(x-h,y) - 2u(x,y) + u(x+h,y) + u(x,y+h) - u(x-h,y+h) - u(x+h,y-h) + u(x,y-h) + u(x+h,y+h))$   
falls  $a_{12}(x,y) < 0$

Beziehungsweise als 9-Punkte-Stern

$$\frac{1}{2h^2} \begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

Insgesamt haben wir als Approximation

$$-\mathcal{L}u(x,y) \approx \frac{1}{h^2} ((-a_{11} + |a_{12}| + h/2b_1)u(x-h,y) + (-a_{11} + |a_{12}| - h/2b_1)u(x+h,y) + (-a_{22} + |a_{12}| + h/2b_2)u(x,y-h) + (-a_{22} + |a_{12}| - h/2b_2)u(x,y+h) + a_{12}(u(x-h,y-h) + u(x+h,y+h))$$

(RWP<sub>3</sub>)<sub>h</sub>

$$\begin{cases} -\mathcal{L}_h u_h &= f_h & \text{in } \Omega_h \\ u_h &= g & \text{auf } \partial\Omega_h \end{cases}$$

(LGS<sub>3</sub>)

$$-\tilde{\mathcal{L}}_h \tilde{u}_h = \tilde{f}_h$$