



Vertiefung Numerische Mathematik für den Masterstudiengang UTRM

R. Verfürth

Fakultät für Mathematik
Ruhr-Universität Bochum

www.ruhr-uni-bochum.de/num1

Vorlesung / Bochum / Sommersemester 2010



Inhalt

Randwertprobleme für gewöhnliche
Differentialgleichungen

Partielle Differentialgleichungen
Differenzenverfahren für partielle
Differentialgleichungen

Finite-Element-Methoden für elliptische
Differentialgleichungen

A posteriori Fehlerschätzung und Adaptivität

Effiziente Löser

Parabolische Differentialgleichungen

Finite-Volumen-Methoden

Literatur



Randwertprobleme für gewöhnliche Differentialgleichungen

- ▶ Anfangswertprobleme
- ▶ Numerische Verfahren für Anfangswertprobleme
- ▶ Randwertprobleme
- ▶ Schießverfahren
- ▶ Mehrzielmethode
- ▶ Differenzenverfahren
- ▶ Variationsmethoden



Anfangswertprobleme für gewöhnliche Differentialgleichungen erster Ordnung

- ▶ Gegeben:
 - ▶ I Intervall
 - ▶ $D \subset \mathbb{R}^d$ Menge
 - ▶ $f(t, y) : I \times D \rightarrow \mathbb{R}^d$ Funktion
 - ▶ $t_0 \in I$ Anfangszeit
 - ▶ $y_0 \in D$ Anfangswert
- ▶ Gesucht:
Differenzierbare Funktion $y(t) : I \rightarrow D$ mit
 $y'(t) = f(t, y(t))$ für alle $t \in I$ (Differentialgleichung)
und $y(t_0) = y_0$ (Anfangsbedingung)



Beispiel: Konstante Wachstums- oder Sterberate

- ▶ $y'(t) = \lambda y(t), y(0) = c$
- ▶ Entspricht:
 - ▶ $I = \mathbb{R},$
 - ▶ $D = \mathbb{R},$
 - ▶ $f(t, y) = \lambda y,$
 - ▶ $t_0 = 0,$
 - ▶ $y_0 = c$
- ▶ Lösung:

$$y(t) = ce^{\lambda t}$$

5 / 264



Beispiel: Gedämpfte Schwingung

- ▶ $y'(t) = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix} y(t), y(0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$
- ▶ Entspricht:
 - ▶ $I = \mathbb{R},$
 - ▶ $D = \mathbb{R}^2,$
 - ▶ $f(t, y) = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix} y,$
 - ▶ $t_0 = 0,$
 - ▶ $y_0 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$
- ▶ Lösung:

$$y(t) = e^{\lambda t} \begin{pmatrix} c_1 \cos(\omega t) - c_2 \sin(\omega t) \\ c_1 \sin(\omega t) + c_2 \cos(\omega t) \end{pmatrix}$$

6 / 264



Beispiel: Explodierende Lösung

- ▶ $y'(t) = y(t)^2, y(0) = 1$
- ▶ Entspricht:
 - ▶ $I = \mathbb{R},$
 - ▶ $D = \mathbb{R},$
 - ▶ $f(t, y) = y^2,$
 - ▶ $t_0 = 0,$
 - ▶ $y_0 = 1$
- ▶ Lösung:

$$y(t) = \frac{1}{1-t} \text{ explodiert für } t \rightarrow 1-$$

7 / 264



Beispiel: Viele Lösungen

- ▶ $y'(t) = \sqrt{|y(t)|}, y(0) = 0$
- ▶ Entspricht:
 - ▶ $I = \mathbb{R},$
 - ▶ $D = \mathbb{R},$
 - ▶ $f(t, y) = \sqrt{|y|},$
 - ▶ $t_0 = 0,$
 - ▶ $y_0 = 0$
- ▶ Lösungen:

$$y(t) = \begin{cases} 0 & \text{für } t < 0 \\ \frac{1}{4}t^2 & \text{für } t \geq 0 \end{cases}$$

und unendlich viele weitere
Lösungen



8 / 264



Differentialgleichungen höherer Ordnung

- ▶ Differentialgleichungen höherer Ordnung können durch Einführen neuer Unbekannter in Differentialgleichungen erster Ordnung transformiert werden.
- ▶ **Beispiel: mechanisches System**
 - ▶ $Mx''(t) + Rx'(t) + Kx(t) = F(t)$, $x(0) = x_0$, $x'(0) = v_0$
 - ▶ Einführen von $v(t) = x'(t)$ führt auf

$$\begin{aligned} x'(t) &= v(t), \\ v'(t) &= M^{-1}F(t) - M^{-1}Rv(t) - M^{-1}Kx(t), \\ x(0) &= x_0, v(0) = v_0 \end{aligned}$$
 - ▶ Dies entspricht

$$y(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix},$$

$$f(t, y) = \begin{pmatrix} M^{-1}F(t) \\ -M^{-1}Kx - M^{-1}Rv \end{pmatrix} y$$



Eindeutige Lösbarkeit

- ▶ Falls f stetig differenzierbar ist bzgl. der Variablen y , gibt es ein Intervall $J = (t_-, t_+) \subset I$ mit $t_0 \in J$ und eine eindeutige auf J stetig differenzierbare Funktion y , die das Anfangswertproblem $y'(t) = f(t, y(t))$, $y(t_0) = y_0$ löst.
- ▶ Es ist $J = I$ oder $y(t)$ strebt für $t \rightarrow t_{\pm}$ gegen den Rand von D .
- ▶ Ist die Ableitung von f bzgl. der Variablen y auf $I \times D$ beschränkt, so ist $J = I$.



Abhängigkeit von den Anfangswerten

- ▶ Falls f zweimal stetig differenzierbar ist bzgl. der Variablen y , hängt die Lösung y des Anfangswertproblems $y'(t) = f(t, y(t))$, $y(t_0) = y_0$ differenzierbar vom Anfangswert y_0 ab, d.h. $y(t) = y(t; y_0)$.
- ▶ Die Ableitung $Z(t)$ der Funktion $y_0 \mapsto y(t; y_0)$ löst das Anfangswertproblem

$$Z'(t) = D_y f(t, y(t; y_0)) Z(t), Z(t_0) = I.$$
 Dabei ist $D_y f(t, y)$ die **Jacobi-Matrix** von f bzgl. der Variablen y und I die **Einheitsmatrix**.



Beispiel: Gedämpfte Schwingung

- ▶ $y'(t) = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix} y(t)$, $y(0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$
- ▶ $y(t) = e^{\lambda t} \begin{pmatrix} c_1 \cos(\omega t) - c_2 \sin(\omega t) \\ c_1 \sin(\omega t) + c_2 \cos(\omega t) \end{pmatrix}$
- ▶ $D_y f(t, y) = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix}$
- ▶ $Z(t) = \begin{pmatrix} z_{1,1}(t) & z_{1,2}(t) \\ z_{2,1}(t) & z_{2,2}(t) \end{pmatrix}$
- ▶ $Z'(t) = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix} Z(t)$, $Z(0) = I$
- ▶ $z'_{1,1}(t) = \lambda z_{1,1}(t) - \omega z_{2,1}(t)$, $z_{1,1}(0) = 1$
- ▶ $z'_{1,2}(t) = \lambda z_{1,2}(t) - \omega z_{2,2}(t)$, $z_{1,2}(0) = 0$
- ▶ $z'_{2,1}(t) = \omega z_{1,1}(t) + \lambda z_{2,1}(t)$, $z_{2,1}(0) = 0$
- ▶ $z'_{2,2}(t) = \omega z_{1,2}(t) + \lambda z_{2,2}(t)$, $z_{2,2}(0) = 1$



Grundidee

- ▶ Approximiere die Lösung y des Anfangswertproblems zu Zeiten $t_0 < t_1 < t_2 < \dots$
- ▶ Bezeichne mit $h_i = t_{i+1} - t_i$ die i -te Schrittweite.
- ▶ Im einfachsten Fall ist $h_i = h$ für alle i , d.h. $t_i = t_0 + ih$.
- ▶ Bezeichne mit η_i die Approximation für $y(t_i)$.
- ▶ Berechne η_{i+1} aus f und η_i (**Einschrittverfahren**) oder aus f und $\eta_i, \dots, \eta_{i-m}$ (**Mehrschrittverfahren**).
- ▶ Viele Verfahren, insbesondere **Runge-Kutta-Verfahren**, erhält man durch Anwenden einer **Quadraturformel** auf das Integral in der Identität

$$\eta_{i+1} - \eta_i \approx y(t_{i+1}) - y(t_i) = \int_{t_i}^{t_{i+1}} f(s, y(s)) ds.$$

13 / 264



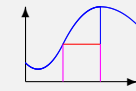
Die einfachsten Verfahren

- ▶ **Explizites Euler-Verfahren:**

$$\eta_0 = y_0,$$

$$\eta_{i+1} = \eta_i + h_i f(t_i, \eta_i),$$

$$t_{i+1} = t_i + h_i$$

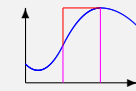


- ▶ **Implizites Euler-Verfahren:**

$$\eta_0 = y_0,$$

$$\eta_{i+1} = \eta_i + h_i f(t_{i+1}, \eta_{i+1}),$$

$$t_{i+1} = t_i + h_i$$

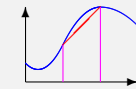


- ▶ **Trapezregel oder Verfahren von Crank-Nicolson:**

$$\eta_0 = y_0,$$

$$\eta_{i+1} = \eta_i + \frac{h_i}{2} [f(t_i, \eta_i) + f(t_{i+1}, \eta_{i+1})],$$

$$t_{i+1} = t_i + h_i$$



14 / 264



Runge-Kutta-Verfahren

- ▶ $\eta_0 = y_0$

$$\eta_{i,j} = \eta_i + h_i \sum_{k=1}^r a_{jk} f(t_i + c_k h, \eta_{i,k}) \quad \text{für } j = 1, \dots, r$$

$$\eta_{i+1} = \eta_i + h_i \sum_{k=1}^r b_k f(t_i + c_k h, \eta_{i,k})$$

$$t_{i+1} = t_i + h_i$$

- ▶ $0 \leq c_1 \leq \dots \leq c_r \leq 1$
- ▶ r heißt **Stufe** des Runge-Kutta-Verfahrens.
- ▶ Das Verfahren heißt **explizit**, wenn $a_{jk} = 0$ ist für alle $k \geq j$.
- ▶ Das Verfahren heißt **implizit**, wenn $a_{j,k} \neq 0$ ist für mindestens ein $k \geq j$.

15 / 264



Ordnung

- ▶ Ein Einschrittverfahren hat die **Ordnung** $p > 0$, wenn gilt $|y(t_1) - \eta_1| = O(h_1^{p+1})$.
- ▶ Die Ordnung ist ein Maß für den Fehler eines **einzelnen** Schrittes des Verfahrens.
- ▶ Hat das Einschrittverfahren die Ordnung p und ist f bzgl. der Variablen y stetig differenzierbar mit beschränkter Ableitung, gilt $|y(t_i) - \eta_i| = O\left(\left(\max_{1 \leq j \leq i} h_j\right)^p\right)$ für alle i .
- ▶ Die beiden Euler-Verfahren haben die Ordnung 1.
- ▶ Das Verfahren von Crank-Nicolson hat die Ordnung 2.

16 / 264



Stabilität

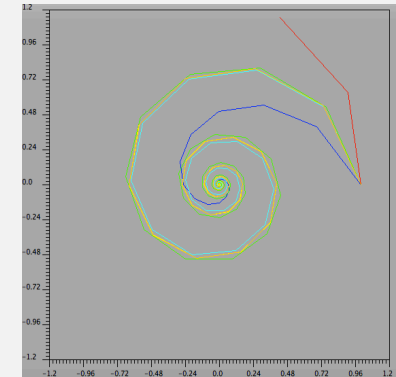
- ▶ Das numerische Verfahren sollte für einen möglichst großen Bereich von Schrittweiten eine Näherungslösung liefern, die das gleiche **qualitative** Verhalten hat wie die exakte Lösung des Anfangswertproblems.
- ▶ Für das Anfangswertproblem $y'(t) = -100y(t)$, $y(0) = 1$ mit exakter Lösung $y(t) = e^{-100t}$ gilt:
 - ▶ Das explizite Euler-Verfahren liefert nur dann eine abklingende Lösung, wenn $h_i \leq \frac{1}{50}$ ist für alle i .
 - ▶ Das implizite Euler-Verfahren und das Verfahren von Crank-Nicolson liefern **für jede Schrittweite** eine abklingende Lösung.
- ▶ Explizite Verfahren können nicht stabil sein.
- ▶ Es gibt stabile implizite Runge-Kutta-Verfahren beliebig hoher Ordnung.

17/ 264



Beispiel: Gedämpfte Schwingung

- ▶ $y'(t) = \begin{pmatrix} -0.9 & -6.3 \\ 6.3 & -0.9 \end{pmatrix} y(t)$
 $y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
- ▶ Lösung:
 $y(t) = e^{-0.9t} \begin{pmatrix} \cos(6.3t) \\ \sin(6.3t) \end{pmatrix}$
- ▶ 100 Schritte
expliziter Euler,
impliziter Euler,
Crank-Nicolson,
klassischer Runge-Kutta,
SDIRK Ordnung 3,
SDIRK Ordnung 4
mit $h_i = 0.1$ für alle i

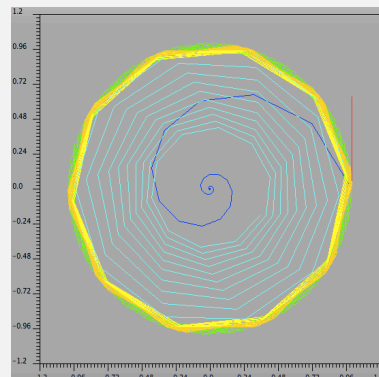


18/ 264



Beispiel: Ungedämpfte Schwingung

- ▶ $y'(t) = \begin{pmatrix} 0 & -6.3 \\ 6.3 & 0 \end{pmatrix} y(t)$
 $y(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
- ▶ Lösung:
 $y(t) = \begin{pmatrix} \cos(6.3t) \\ \sin(6.3t) \end{pmatrix}$
- ▶ 100 Schritte
expliziter Euler,
impliziter Euler,
Crank-Nicolson,
klassischer Runge-Kutta,
SDIRK Ordnung 3,
SDIRK Ordnung 4
mit $h_i = 0.1$ für alle i



19/ 264



Randwertprobleme für gewöhnliche Differentialgleichungen erster Ordnung

- ▶ Gegeben:
 - ▶ I Intervall
 - ▶ $a, b \in I$ zwei verschiedene Punkte
 - ▶ $D \subset \mathbb{R}^d$ Menge
 - ▶ $f(t, y) : I \times D \rightarrow \mathbb{R}^d$ Funktion
 - ▶ $r(u, v) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ Funktion
- ▶ Gesucht:
Differenzierbare Funktion $y(t) : I \rightarrow D$ mit
 $y'(t) = f(t, y(t))$ für alle $t \in I$ (Differentialgleichung)
und $r(y(a), y(b)) = 0$ (Randbedingung)

20/ 264



Beispiel: Gedämpfte Schwingung

▶ $y'(t) = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix} y(t)$, $y_1(0) = 1$, $y_1(\frac{\pi}{2\omega}) = 0$

▶ Entspricht:

▶ $I = \mathbb{R}$,

▶ $D = \mathbb{R}^2$,

▶ $f(t, y) = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix} y$,

▶ $a = 0$,

▶ $b = \frac{\pi}{2\omega}$,

▶ $r(u, v) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} u + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} v - \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

▶ Lösung:

$$y(t) = e^{\lambda t} \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \end{pmatrix}$$



Beispiel: Mechanisches System

▶ $Mx''(t) + Rx'(t) + Kx(t) = F(t)$, $x(0) = x_0$, $x(L) = x_L$

▶ Einführen von $v(t) = x'(t)$ führt auf

$$x'(t) = v(t),$$

$$v'(t) = M^{-1}F(t) - M^{-1}Rv(t) - M^{-1}Kx(t),$$

$$x(0) = x_0, x(L) = x_L$$

▶ Dies entspricht

$$y(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix},$$

$$f(t, y) = \begin{pmatrix} M^{-1}F(t) \\ -M^{-1}Kx - M^{-1}Rv \end{pmatrix} y,$$

$$r(u, v) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} u + \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix} v - \begin{pmatrix} x_0 \\ x_L \end{pmatrix}$$



Beispiel: Eigenwertproblem

▶ Gesucht sind $u : [a, b] \rightarrow \mathbb{R}$ und $\lambda \in \mathbb{R}$ mit

$$u'(t) = g(t, u(t)), \rho(u(a), u(b), \lambda) = 0$$

▶ Entspricht:

▶ $y(t) = \begin{pmatrix} u(t) \\ \lambda \end{pmatrix}$,

▶ $f(t, y) = \begin{pmatrix} g(t, y_1) \\ 0 \end{pmatrix}$,

▶ $r(u, v) = \rho(u_1, v_1, v_2)$



Beispiel: Freies Randwertproblem

▶ Gesucht sind $\beta > 0$ und $u : [0, \beta] \rightarrow \mathbb{R}$ mit

$$u'(s) = g(s, u(s)), \rho(u(0), u(\beta)) = 0$$

▶ Entspricht:

▶ $y(t) = \begin{pmatrix} u(t\beta) \\ \beta \end{pmatrix}$,

▶ $t = \frac{s}{\beta}$,

▶ $f(t, y) = \begin{pmatrix} y_2 g(ty_2, y_1) \\ 0 \end{pmatrix}$,

▶ $r(u, v) = \rho(u_1, v_1)$



Eindeutige Lösbarkeit

- ▶ Für Randwertprobleme gibt es keine allgemeine Existenz- und Eindeutigkeitsaussage wie für Anfangswertprobleme.
- ▶ Die Lösbarkeit und die Zahl allfälliger Lösungen hängt von dem konkreten Beispiel und dem Zusammenspiel von Differentialgleichung und Randbedingung ab.
- ▶ **Beispiel: Schwingung**
 - ▶ $y'(t) = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} y(t)$, $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} y(0) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} y(L) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$
 - ▶ Allgemeine Lösung der Differentialgleichung:

$$y(t) = \begin{pmatrix} c_1 \cos(\omega t) - c_2 \sin(\omega t) \\ c_1 \sin(\omega t) + c_2 \cos(\omega t) \end{pmatrix}$$
 - ▶ $L = \frac{2\pi}{\omega}$, $\alpha = 0$, $\beta = 1$ führt auf die **widersprüchlichen** Bedingungen $c_1 = 0$ und $c_1 = 1$.
 - ▶ $L = \frac{2\pi}{\omega}$, $\alpha = 0$, $\beta = 0$ führt auf die **einzig**e Bedingung $c_1 = 0$, so dass c_2 **beliebig** ist.

25 / 264



Idee

- ▶ Bezeichne mit $y(t; s)$ die Lösung des **Anfangswertproblems** $y'(t) = f(t, y(t))$, $y(a; s) = s$.
- ▶ Dann löst $y(t; s)$ das **Randwertproblem** $y'(t) = f(t, y(t))$, $r(y(a), y(b)) = 0$ genau dann, wenn gilt $r(s, y(b; s)) = 0$.
- ▶ Bestimme mit dem **Newton-Verfahren** eine **Nullstelle** der Funktion $F(s) = r(s, y(b; s))$.
- ▶ Die Ableitung $DF(s)$ von F an der Stelle s ist gegeben durch $DF(s) = D_u r(s, y(b; s)) + D_v r(s, y(b; s))Z(b; s)$, wobei Z das **Anfangswertproblem** $Z'(t; s) = D_y f(t, y(t; s))Z(t; s)$, $Z(a; s) = I$ löst.
- ▶ Löse die beiden Anfangswertprobleme für $y(t; s)$ und $Z(t; s)$ näherungsweise mit einem numerischen Verfahren für Anfangswertprobleme, wobei beide Male die gleichen Gitterpunkte t_i benutzt werden.

26 / 264



Schießverfahren

0. Gegeben sei ein Startwert $s^{(0)} \in \mathbb{R}^d$. Setze $i = 0$.
1. Berechne mit einem numerischen Verfahren für Anfangswertprobleme eine Näherung $\eta^{(i)}(t)$ für die Lösung $y^{(i)}$ des **Anfangswertproblems** $y^{(i)'}(t) = f(t, y^{(i)}(t))$, $y^{(i)}(a) = s^{(i)}$. Setze $F^{(i)} = r(s^{(i)}, \eta^{(i)}(b))$.
2. Berechne mit dem gleichen Verfahren wie in Schritt 1 und den gleichen Gitterpunkten eine Näherung $\zeta^{(i)}(t)$ für die Lösung $Z^{(i)}$ des **Anfangswertproblems** $Z^{(i)'}(t) = D_y f(t, \eta^{(i)}(t))Z^{(i)}(t)$, $Z^{(i)}(a) = I$. Setze $D^{(i)} = D_u r(s^{(i)}, \eta^{(i)}(b)) + D_v r(s^{(i)}, \eta^{(i)}(b))\zeta^{(i)}(b)$.
3. Löse das **lineare Gleichungssystem** $D^{(i)}\Delta s^{(i)} = -F^{(i)}$. Setze $s^{(i+1)} = s^{(i)} + \Delta s^{(i)}$, erhöhe i um 1 und gehe zu Schritt 1 zurück.

27 / 264



Eigenschaften

- ▶ Die Anfangswertprobleme in Schritt 1 haben d Unbekannte.
- ▶ Die Anfangswertprobleme in Schritt 2 haben d^2 Unbekannte.
- ▶ Die Anfangswertprobleme in Schritt 2 sind **linear**.
- ▶ Die linearen Gleichungssysteme in Schritt 2 haben d Gleichungen und Unbekannte.
- ▶ Das Newton-Verfahren sollte gedämpft werden.
- ▶ Falls das Newton-Verfahren konvergiert, ist die Konvergenz quadratisch.

28 / 264



Ein warnendes Beispiel

- ▶ Randwertproblem:

$$y'(t) = \begin{pmatrix} 0 & 1 \\ 110 & 1 \end{pmatrix} y(t), \quad y_1(0) = 1, \quad y_1(10) = 1$$

- ▶ Lösung:

$$y(t) = c_1 e^{-10t} \begin{pmatrix} 1 \\ -10 \end{pmatrix} + c_2 e^{11t} \begin{pmatrix} 1 \\ 11 \end{pmatrix}$$

$$\text{mit } c_1 = \frac{e^{110} - 1}{e^{110} - e^{-100}}, \quad c_2 = \frac{1 - e^{-100}}{e^{110} - e^{-100}}$$

- ▶ Lösung des Anfangswertproblems zum Anfangswert s :

$$y(t; s) = \frac{11s_1 - s_2}{21} e^{-10t} \begin{pmatrix} 1 \\ -10 \end{pmatrix} + \frac{10s_1 + s_2}{21} e^{11t} \begin{pmatrix} 1 \\ 11 \end{pmatrix}$$

- ▶ Exakter Anfangswert:

$$s^* = \begin{pmatrix} 1 \\ -10 + 21 \cdot \frac{1 - e^{-100}}{e^{110} - e^{-100}} \end{pmatrix}$$

- ▶ Der falsche Anfangswert $\tilde{s} = \begin{pmatrix} 1 \\ -10 + 10^{-9} \end{pmatrix}$ mit einem **relativen Fehler** von 10^{-10} liefert $y_1(10; \tilde{s}) \approx 10^{37}$.

29/ 264



Beobachtung

- ▶ Das Schießverfahren versagt, weil Lösungen zu verschiedenen Anfangswerten exponentiell auseinander laufen können.
- ▶ Man kann diesen Effekt vermeiden, indem man Anfangswertprobleme nur auf kleinen Intervallen löst.

30/ 264



Idee

- ▶ Unterteile das Intervall $[a, b]$ durch Zwischenpunkte $a = \tau_1 < \tau_2 < \dots < \tau_m = b$.
- ▶ Für $s_1, \dots, s_m \in \mathbb{R}^d$ bezeichne mit $y(t; \tau_k, s_k)$ die Lösung des **Anfangswertproblems** $y'(t) = f(t, y(t))$, $y(\tau_k; s_k) = s_k$.
- ▶ Definiere die **stückweise** Funktion \tilde{y} durch $\tilde{y}(t) = y(t; \tau_k, s_k)$ für $\tau_k \leq t < \tau_{k+1}$ und $1 \leq k \leq m-1$ und $\tilde{y}(\tau_m) = s_m$.
- ▶ Dann löst \tilde{y} das **Randwertproblem** $y'(t) = f(t, y(t))$, $r(y(a), y(b)) = 0$ genau dann, wenn gilt $y(\tau_{k+1}; \tau_k, s_k) = s_{k+1}$ für $1 \leq k \leq m-1$ und $r(s_1, s_m) = 0$.
- ▶ Dies definiert ein **Gleichungssystem** $F(s_1, \dots, s_m) = 0$, das mit dem Newton-Verfahren gelöst werden kann.
- ▶ Die Berechnung der Ableitung von F erfordert das Lösen von Anfangswertproblemen auf den Intervallen $[\tau_k, \tau_{k+1}]$.

31/ 264



Struktur von DF

- ▶ DF hat die Struktur
$$\begin{pmatrix} G_1 & -I & & & 0 \\ & G_2 & -I & & \\ 0 & & \ddots & \ddots & \\ & & & G_{m-1} & -I \\ A & 0 & & 0 & B \end{pmatrix}$$
- ▶ Daher ist in jedem Newton-Schritt ein System zu lösen, das die Form hat:

$$G_1 \Delta s_1 - \Delta s_2 = -F_1, \quad \dots, \quad G_{m-1} \Delta s_{m-1} - \Delta s_m = -F_{m-1}, \\ A \Delta s_1 + B \Delta s_m = -F_m$$

- ▶ Sukzessive Elimination von $\Delta s_2, \dots, \Delta s_m$ führt auf

$$(A + B G_{m-1} \dots G_1) \Delta s_1 = -F_m - B \sum_{j=1}^{m-1} \left(\prod_{i=j+1}^{m-1} G_i \right) F_j$$

32/ 264



Mehrzielmethode I

0. Gegeben seien m Punkte $a = \tau_1 < \dots < \tau_m = b$ und m Vektoren $s_1^{(0)}, \dots, s_m^{(0)} \in \mathbb{R}^d$. Setze $i = 0$.
1. Bestimme mit einem numerischen Verfahren für Anfangswertprobleme Näherungen $\eta^{(i,j)}(t)$, $1 \leq j \leq m-1$, für die Lösungen $y^{(i,j)}$ der **Anfangswertprobleme** $y^{(i,j)'}(t) = f(t, y^{(i,j)}(t))$, $y^{(i,j)}(\tau_j) = s_j^{(i)}$ für $1 \leq j \leq m-1$.
Setze $F_j^{(i)} = \eta^{(i,j)}(\tau_{j+1}) - s_{j+1}^{(i)}$ für $1 \leq j \leq m-1$ und $F_m^{(i)} = r(s_1^{(i)}, s_m^{(i)})$.

33 / 264



Mehrzielmethode II

2. Bestimme mit dem gleichen Verfahren wie in Schritt 1 und den gleichen Gitterpunkten Näherungen $\zeta^{(i,j)}(t)$ für die Lösungen $Z^{(i,j)}$ der Anfangswertprobleme $Z^{(i,j)'}(t) = D_y f(t, \eta^{(i,j)}(t)) Z^{(i,j)}(t)$, $Z^{(i,j)}(\tau_j) = I$ für $1 \leq j \leq m-1$.
Setze $G_j^{(i)} = \zeta^{(i,j)}(\tau_{j+1})$ für $1 \leq j \leq m-1$ und $A^{(i)} = D_u r(s_1^{(i)}, s_m^{(i)})$,
 $B^{(i)} = D_v r(s_1^{(i)}, s_m^{(i)})$.

34 / 264



Mehrzielmethode III

3. Berechne die Matrix $H^{(i)} = A^{(i)} + B^{(i)} G_{m-1}^{(i)} \cdot \dots \cdot G_1^{(i)}$ und den Vektor $\varphi^{(i)} = -F_m^{(i)} - B^{(i)} \sum_{j=1}^{m-1} \left(\prod_{l=j+1}^{m-1} G_l^{(i)} \right) F_j^{(i)}$.
Löse das **lineare Gleichungssystem** $H^{(i)} \Delta s_1^{(i)} = \varphi^{(i)}$ und berechne rekursiv die Vektoren $\Delta s_{k+1}^{(i)} = G_k^{(i)} \Delta s_k^{(i)} + F_k^{(i)}$ für $1 \leq k \leq m-1$.
Setze $s_k^{(i+1)} = s_k^{(i)} + \Delta s_k^{(i)}$ für $1 \leq k \leq m$, erhöhe i um 1 und gehe zu Schritt 1 zurück.

35 / 264



Eigenschaften

- ▶ Bei gleicher Zahl von Gitterpunkten auf dem gesamten Intervall $[a, b]$ erfordern die Anfangswertprobleme beim Schießverfahren und bei der Mehrzielmethode den gleichen Aufwand.
- ▶ Die Anfangswertprobleme auf den Teilintervallen können **parallel** gelöst werden.
- ▶ Wenn keine zusätzlichen Informationen bekannt sind, können die Punkte τ_1, \dots, τ_m **äquidistant** gewählt werden.

36 / 264



Sturm-Liouville-Problem

► Gegeben:

- $p: [0, 1] \rightarrow \mathbb{R}$ stetig differenzierbare Funktion mit

$$\underline{p} = \min_{0 \leq x \leq 1} p(x) > 0$$

- $q: [0, 1] \rightarrow \mathbb{R}$ stetige Funktion mit

$$\underline{q} = \min_{0 \leq x \leq 1} q(x) > 0$$

► Gesucht:

Zweimal stetig differenzierbare Funktion $u: [0, 1] \rightarrow \mathbb{R}$ mit
 $-(pu')' + qu = f$ in $(0, 1)$ (Differentialgleichung)
 und $u(0) = 0, u(1) = 0$ (Randbedingung)



Verallgemeinerung

- Jedes Sturm-Liouville-Problem der Form

$$-(pu')' + qu = f \text{ in } (a, b), u(a) = \alpha, u(b) = \beta$$

kann transformiert werden in ein äquivalentes Problem mit:

$$a = 0, b = 1, \alpha = 0, \beta = 0.$$

- Suche dazu u in der Form

$$u(x) = \alpha + \frac{\beta - \alpha}{b - a}(x - a) + v(x)$$

mit $v(0) = 0, v(1) = 0$

und führe eine neue Variable ein durch

$$t = \frac{x - a}{b - a}.$$



Symmetrischer Differenzenquotient

- Der **symmetrische Differenzenquotient** ist definiert durch

$$\partial_h \varphi(x) = \frac{1}{h} [\varphi(x + \frac{h}{2}) - \varphi(x - \frac{h}{2})].$$

- **Taylor-Entwicklung** liefert für jede dreimal stetig differenzierbare Funktion:

$$\partial_h \varphi(x) = \varphi'(x) + \frac{h^2}{24} \varphi'''(x + \theta h)$$

mit einem geeigneten $\theta \in (-\frac{1}{2}, \frac{1}{2})$.



Idee

- Ersetze Ableitungen durch Differenzenquotienten ∂_h

$$\begin{aligned} & (-pu')'(x) \\ & \approx (-\partial_h(pu'))(x) \\ & = \frac{1}{h} [p(x - \frac{h}{2})u'(x - \frac{h}{2}) - p(x + \frac{h}{2})u'(x + \frac{h}{2})] \\ & \approx \frac{1}{h} [p(x - \frac{h}{2})\partial_h u(x - \frac{h}{2}) - p(x + \frac{h}{2})\partial_h u(x + \frac{h}{2})] \\ & = \frac{1}{h^2} [p(x - \frac{h}{2})(u(x) - u(x - h)) - p(x + \frac{h}{2})(u(x + h) - u(x))] \end{aligned}$$

- Fordere die resultierenden Gleichungen nur in Gitterpunkten ih mit $h = \frac{1}{n+1}$ und $1 \leq i \leq n$.



Differenzdiskretisierung

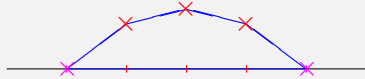
- ▶ Wähle eine **Gitterweite** $h = \frac{1}{n+1}$.
- ▶ Für $1 \leq i \leq n$ setze $f_i = f(ih)$, $q_i = q(ih)$, $p_{i \pm \frac{1}{2}} = p(ih \pm \frac{h}{2})$
- ▶ Bestimme u_0, \dots, u_{n+1} so dass

$$u_0 = 0, u_{n+1} = 0$$

und für $1 \leq i \leq n$

$$f_i = -\frac{1}{h^2} p_{i-\frac{1}{2}} u_{i-1} + \left(\frac{1}{h^2} [p_{i-\frac{1}{2}} + p_{i+\frac{1}{2}}] + q_i \right) u_i - \frac{1}{h^2} p_{i+\frac{1}{2}} u_{i+1}$$

- ▶ Bezeichne mit u_h die stetige, stückweise lineare Funktion, die an den Stellen ih mit u_i übereinstimmt.



Eigenschaften

- ▶ Die Differenzdiskretisierung führt auf ein lineares Gleichungssystem mit n Gleichungen für die n Unbekannten u_1, \dots, u_n .
- ▶ Die Matrix des Gleichungssystems ist symmetrisch, positiv definit und **tridiagonal** mit positiven Diagonalelementen und negativen Nebendiagonalelementen.
- ▶ Das Gleichungssystem hat eine eindeutige Lösung.
- ▶ Die Lösung des Gleichungssystems mit dem **Gaußschen Eliminationsverfahren** oder der **Cholesky-Zerlegung** erfordert $O(n)$ Operationen.



Fehlerabschätzung

- ▶ Es gelte:
 - ▶ p ist dreimal stetig differenzierbar.
 - ▶ Die Lösung u des Sturm-Liouville-Problems ist viermal stetig differenzierbar.

- ▶ Dann gilt die **Fehlerabschätzung**

$$\max_{0 \leq x \leq 1} |u(x) - u_h(x)| \leq ch^2.$$

- ▶ Die Konstante c hängt von der unteren Schranke \underline{q} für q , den Ableitungen bis zur Ordnung 3 von p und den Ableitungen bis zur Ordnung 4 von u ab.



Idee der Variationsformulierung

- ▶ Multipliziere die **Differentialgleichung** mit einer stetig differenzierbaren Funktion v mit $v(0) = 0$, $v(1) = 0$
 $-(pu')'(x)v(x) + q(x)u(x)v(x) = f(x)v(x)$ für $0 \leq x \leq 1$.
- ▶ Integriere das Ergebnis von 0 bis 1

$$\int_0^1 [-(pu')'(x)v(x) + q(x)u(x)v(x)] dx = \int_0^1 f(x)v(x) dx.$$

- ▶ Integriere den **Ableitungsterm** partiell

$$\begin{aligned} & - \int_0^1 (pu')'(x)v(x) dx \\ &= p(0)u'(0)v(0) - p(1)u'(1)v(1) + \int_0^1 p(x)u'(x)v'(x) dx \\ &= \int_0^1 p(x)u'(x)v'(x) dx \end{aligned}$$



Probleme

- ▶ Für eine sinnvolle Variationsformulierung müssen die Eigenschaften der Funktionen u und v präziser gefasst werden.
- ▶ Klassische Eigenschaften wie stetige Differenzierbarkeit sind zu restriktiv.
- ▶ Der Begriff der Ableitung muss daher geeignet verallgemeinert werden.
- ▶ In Hinblick auf die Diskretisierung sollten insbesondere stückweise differenzierbare Funktionen im erweiterten Sinn differenzierbar sein.

45 / 264



Schwache Ableitung

- ▶ Partielle Integration liefert für stetig differenzierbare Funktionen u und v mit $v(0) = 0$, $v(1) = 0$:

$$\int_0^1 u'(x)v(x)dx = u(1)v(1) - u(0)v(0) - \int_0^1 u(x)v'(x)dx$$

$$= - \int_0^1 u(x)v'(x)dx$$

- ▶ Die Funktion u heißt **schwach differenzierbar** mit **schwacher Ableitung** w , wenn für jede stetig differenzierbare Funktion v mit $v(0) = 0$, $v(1) = 0$ gilt


$$\int_0^1 w(x)v(x)dx = - \int_0^1 u(x)v'(x)dx$$

46 / 264



Beispiele

- ▶ Jede stetig differenzierbare Funktion ist schwach differenzierbar und die schwache Ableitung stimmt mit der klassischen Ableitung überein.
- ▶ Jede stetige, stückweise stetig differenzierbare Funktion ist schwach differenzierbar und die schwache Ableitung stimmt mit der stückweisen klassischen Ableitung überein.
- ▶ $u(x) = 1 - |2x - 1|$ ist schwach differenzierbar mit schwacher Ableitung $w(x) = \begin{cases} 2 & \text{für } 0 < x < \frac{1}{2} \\ -2 & \text{für } \frac{1}{2} < x < 1 \end{cases}$.

(Beachte: Der Wert $w(\frac{1}{2})$ ist beliebig.) 

47 / 264




Sobolev-Räume

- ▶ $\|v\| = \left\{ \int_0^1 |v(x)|^2 dx \right\}^{\frac{1}{2}}$ ist die L^2 -Norm.
- ▶ $L^2(0, 1)$ ist der **Lebesgue-Raum** aller Funktionen v mit endlicher L^2 -Norm $\|v\|$.
- ▶ $H^1(0, 1)$ ist der **Sobolev-Raum** aller Funktionen v in $L^2(0, 1)$, deren schwache Ableitung existiert und ebenfalls in $L^2(0, 1)$ ist.
- ▶ $H_0^1(0, 1)$ ist der **Sobolev-Raum** aller Funktionen v in $H^1(0, 1)$ mit $v(0) = 0$ und $v(1) = 0$.

48 / 264



Beispiele

- ▶ Jede beschränkte Funktion ist in $L^2(0, 1)$.
- ▶ $v(x) = \frac{1}{\sqrt{x}}$ ist nicht in $L^2(0, 1)$, da das Integral von $\frac{1}{x} = v(x)^2$ nicht endlich ist.
- ▶ Jede stetig differenzierbare Funktion ist in $H^1(0, 1)$.
- ▶ Jede stetige, stückweise stetig differenzierbare Funktion ist in $H^1(0, 1)$.
- ▶ $v(x) = 1 - |2x - 1|$ ist in $H_0^1(0, 1)$. 
- ▶ $v(x) = 2\sqrt{x}$ ist nicht in $H^1(0, 1)$, da das Integral von $\frac{1}{x} = (v'(x))^2$ nicht endlich ist.
- ▶ Im Gegensatz zu mehrdimensionalen Problemen sind Funktionen in $H^1(0, 1)$ immer stetig.



Variationsproblem

Finde $u \in H_0^1(0, 1)$ so, dass für alle $v \in H_0^1(0, 1)$ gilt

$$\int_0^1 [p(x)u'(x)v'(x) + q(x)u(x)v(x)] dx = \int_0^1 f(x)v(x) dx$$




Eigenschaften des Variationsproblems

- ▶ Das Variationsproblem hat eine eindeutige Lösung.
- ▶ Die Lösung des Variationsproblems ist das eindeutige Minimum in $H_0^1(0, 1)$ der Energiefunktion

$$\frac{1}{2} \int_0^1 [p(x)u'(x)^2 + q(x)u(x)^2] dx - \int_0^1 f(x)u(x) dx.$$



Finite-Element-Räume

- ▶ \mathcal{T} bezeichnet eine beliebige Unterteilung des Intervalls $(0, 1)$ in nicht überlappende Teilintervalle. 
- ▶ $k \geq 1$ bezeichnet einen beliebigen Polynomgrad.
- ▶ $S^{k,0}(\mathcal{T})$ ist der Finite-Element-Raum aller stetigen Funktionen, die stückweise auf den Intervallen von \mathcal{T} Polynome vom Grad höchstens k sind.
- ▶ $S_0^{k,0}(\mathcal{T})$ ist der Finite-Element-Raum aller Funktionen v in $S^{k,0}(\mathcal{T})$ mit $v(0) = 0$ und $v(1) = 0$.



Finite-Element-Problem

Finde $u_{\mathcal{T}} \in S_0^{k,0}(\mathcal{T})$ (**Ansatzfunktion**) so, dass für alle $v_{\mathcal{T}} \in S_0^{k,0}(\mathcal{T})$ (**Testfunktion**) gilt

$$\int_0^1 [p(x)u'_{\mathcal{T}}(x)v'_{\mathcal{T}}(x) + q(x)u_{\mathcal{T}}(x)v_{\mathcal{T}}(x)] dx = \int_0^1 f(x)v_{\mathcal{T}}(x) dx$$



Eigenschaften des Finite-Element-Problems

- ▶ Das Finite-Element-Problem hat eine eindeutige Lösung.
- ▶ Die Lösung des Finite-Element-Problems ist das eindeutige **Minimum** in $S_0^{k,0}(\mathcal{T})$ der **Energiefunktion**.
- ▶ Nach Wahl einer Basis für $S_0^{k,0}(\mathcal{T})$ führt das Finite-Element-Problem auf ein lineares Gleichungssystem mit $k \cdot \#\mathcal{T} - 1$ Unbekannten und einer symmetrischen, positiv definiten, tridiagonalen Matrix (**Steifigkeitsmatrix**).
- ▶ Meistens werden die Integrale durch **Quadraturformeln** angenähert.
- ▶ Meistens wird k gleich 1 (**lineare Elemente**) oder 2 (**quadratische Elemente**) gewählt.
- ▶ In der Regel wird eine **nodale** Basis für $S_0^{k,0}(\mathcal{T})$ benutzt.

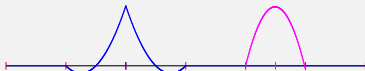


Nodale Basisfunktionen

- ▶ **Lineare Elemente**: Die Funktionen, die in genau einem **Intervall-Endpunkt** den Wert 1 annehmen und an allen anderen Intervall-Endpunkten den Wert 0 haben.



- ▶ **Quadratische Elemente**: Die Funktionen, die in genau einem **Intervall-End-** oder **-Mittelpunkt** den Wert 1 annehmen und an allen anderen Intervall-End- oder Mittelpunkten den Wert 0 haben.



A priori Fehlerabschätzung

- ▶ Bezeichne mit $h_{\mathcal{T}}$ die **maximale Länge** der Intervalle in \mathcal{T} .
- ▶ Dann gelten folgende **a priori Fehlerabschätzungen** für die Lösungen u des Variationsproblems und $u_{\mathcal{T}}$ des Finite-Element-Problems:

$$\|u' - u'_{\mathcal{T}}\| \leq c_1 h_{\mathcal{T}}$$

$$\|u - u_{\mathcal{T}}\| \leq c_2 h_{\mathcal{T}}^2$$

- ▶ Die Konstanten c_1 und c_2 hängen nur von der unteren Schranke \underline{p} für p , Ableitungen bis zur Ordnung 1 von p , dem Maximum von q und Ableitungen bis zur Ordnung 2 von u ab.



Grenzen der a priori Fehlerabschätzung

- ▶ Sie trifft nur eine Aussage über das **asymptotische** Verhalten des Fehlers für $h_{\mathcal{T}} \rightarrow 0$ d.h. für immer feiner werdende Unterteilungen.
- ▶ Sie gibt keine Information über die tatsächliche Größe des Fehlers.
- ▶ Sie erlaubt keine Information über die räumliche Verteilung des Fehlers, die benötigt wird, um zusätzliche Gitterpunkte dort zu platzieren, wo der Fehler groß ist.



A posteriori Fehlerabschätzung

- ▶ Für jedes Intervall K in \mathcal{T} bezeichne mit h_K seine **Länge** und definiere den **Fehlerindikator** η_K durch

$$\eta_K = h_K \left\{ \int_K |f + (pu'_{\mathcal{T}})' - qu_{\mathcal{T}}|^2 dx \right\}^{\frac{1}{2}}.$$

- ▶ Dann gelten die **a posteriori Fehlerabschätzungen**:

$$\|u' - u'_{\mathcal{T}}\| \leq (p)^{-1} \left\{ \sum_{K \in \mathcal{T}} \eta_K^2 \right\}^{\frac{1}{2}}$$

$$\eta_K \leq c \left\{ \int_K |u' - u'_{\mathcal{T}}|^2 dx \right\}^{\frac{1}{2}} \text{ für jedes } K \text{ in } \mathcal{T}$$



Eigenschaften

- ▶ $f + (pu'_{\mathcal{T}})' - qu_{\mathcal{T}}$ ist das intervallweise Residuum der Finite-Element-Lösung bzgl. der Differentialgleichung.
- ▶ Der Aufwand zur Berechnung der Fehlerindikatoren ist vernachlässigbar.
- ▶ Die a posteriori Fehlerabschätzung gibt zuverlässige Informationen über die tatsächliche Größe des Fehlers und seine räumliche Verteilung.
- ▶ Die a posteriori Fehlerabschätzung kann für eine **adaptive Gitterverfeinerung** genutzt werden, so dass jede vorgegebene Genauigkeit mit (nahezu) minimaler Anzahl an Unbekannten erreicht werden kann.



Adaptive Gitterverfeinerung

0. Gegeben: eine Toleranz ε .
Gesucht: Eine Finite-Element-Lösung mit Fehler $\leq \varepsilon$.
1. Wähle eine grobe Unterteilung \mathcal{T}_0 und setze $k = 0$.
2. Löse das Finite-Element-Problem zu \mathcal{T}_k .
3. Für jedes Intervall K in \mathcal{T}_k bestimme den Fehlerindikator η_K und den Maximalwert $\eta_k = \max_{K \in \mathcal{T}_k} \eta_K$.
4. Falls $\eta_k \leq \varepsilon$ ist, **stopp**.
5. Für jedes K in \mathcal{T}_k prüfe, ob $\eta_K \geq \frac{1}{2} \eta_k$ ist. Falls dies der Fall ist, halbiere K , sonst lasse K unverändert. Dies bestimmt die neue Unterteilung \mathcal{T}_{k+1} . Erhöhe k um 1 und gehe zu Schritt **2** zurück.



Partielle Differentialgleichungen

- ▶ Beispiele
- ▶ Typen
- ▶ Lösungseigenschaften
- ▶ Überblick über Diskretisierungsmethoden



Poisson-Gleichung I

- ▶ Die vertikale **Auslenkung** $u : \Omega \rightarrow \mathbb{R}$ einer elastischen, nicht dehnbaren **Membran** mit **Querschnittsfläche** Ω in der (x, y) -Ebene unter Einfluss einer vertikalen **Last** $f : \Omega \rightarrow \mathbb{R}$ wird durch die **Membran-** oder **Poisson-Gleichung** beschrieben:

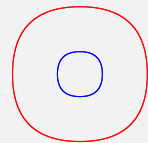
$$-\Delta u = -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f \text{ in } \Omega.$$

- ▶ Auf dem **Rand** Γ von Ω muss die Auslenkung eine der folgenden **Randbedingungen** erfüllen:
 - ▶ **Dirichlet-Randbedingung** $u = 0$ für eine eingespannte Membran,
 - ▶ **Neumann-Randbedingung** $\mathbf{n} \cdot \nabla u = \frac{\partial u}{\partial n} = 0$ für eine frei bewegliche Membran.



Poisson-Gleichung II

- ▶ Die Randbedingungen können in dem Sinne gemischt werden, dass auf einem Teil des Randes die **Dirichlet-** und auf einem anderen, dazu disjunkten Teil die **Neumann-Bedingung** gelten soll.
- ▶ In jedem Punkt des Randes muss genau eine der Bedingungen gefordert werden.
- ▶ Die rechten Seiten 0 der Randbedingungen können durch beliebige gegebene Funktionen ersetzt werden.
- ▶ Die Auslenkung u minimiert die **Energiefunktion** $\frac{1}{2} \int_{\Omega} |\nabla u|^2 dx - \int_{\Omega} f u dx$ in einer geeigneten Menge zulässiger Auslenkungen.



Biharmonische Gleichung I

- ▶ Die vertikale **Auslenkung** $u : \Omega \rightarrow \mathbb{R}$ der **Mittelebene** Ω einer dünnen, starren **Platte** unter Einfluss einer vertikalen **Last** $f : \Omega \rightarrow \mathbb{R}$ wird durch die **Platten-** oder **biharmonische Gleichung** beschrieben:

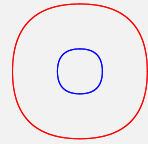
$$\Delta^2 u = \frac{\partial^4 u}{\partial x^4} + 2 \frac{\partial^4 u}{\partial x^2 \partial y^2} + \frac{\partial^4 u}{\partial y^4} = f \text{ in } \Omega.$$

- ▶ Auf dem **Rand** Γ von Ω muss die Auslenkung eine der folgenden **Randbedingungen** erfüllen:
 - ▶ $u = 0$ und $\frac{\partial u}{\partial n} = 0$ für eine eingespannte Platte,
 - ▶ $u = 0$ und $\Delta u = 0$ für eine frei bewegliche Platte.



Biharmonische Gleichung II

- ▶ Die Randbedingungen können in dem Sinne gemischt werden, dass auf einem Teil des Randes die Platte **fest** und auf einem anderen, dazu disjunkten Teil **frei** sein soll.
- ▶ In jedem Punkt des Randes muss genau eine der Bedingungen gefordert werden.
- ▶ Die rechten Seiten 0 der Randbedingungen können durch beliebige gegebene Funktionen ersetzt werden.
- ▶ Die Auslenkung u minimiert die **Energiefunktion** $\frac{1}{2} \int_{\Omega} |\Delta u|^2 dx - \int_{\Omega} f u dx$ in einer geeigneten Menge zulässiger Auslenkungen.



Lineare Elastizitätstheorie I

- ▶ Die **Verformung** $\mathbf{u} : \Omega \rightarrow \mathbb{R}^3$ eines **Körpers** $\Omega \subset \mathbb{R}^3$ unter dem Einfluss äußerer **Kräfte** $\mathbf{f} : \Omega \rightarrow \mathbb{R}^3$ wird durch die Gleichungen der **linearen Elastizitätstheorie** beschrieben:
 - $\text{div } \sigma = \mathbf{f}$ in Ω .

- ▶ Dabei ist:

$$\text{div } \sigma = \begin{pmatrix} \frac{\partial \sigma_{11}}{\partial x} + \frac{\partial \sigma_{21}}{\partial y} + \frac{\partial \sigma_{31}}{\partial z} \\ \frac{\partial \sigma_{12}}{\partial x} + \frac{\partial \sigma_{22}}{\partial y} + \frac{\partial \sigma_{32}}{\partial z} \\ \frac{\partial \sigma_{13}}{\partial x} + \frac{\partial \sigma_{23}}{\partial y} + \frac{\partial \sigma_{33}}{\partial z} \end{pmatrix}$$

$$\sigma = C \varepsilon$$

$$\varepsilon = \frac{1}{2} (\nabla \mathbf{u} + (\nabla \mathbf{u})^t) = \begin{pmatrix} \frac{\partial u_1}{\partial x} & \frac{1}{2} \frac{\partial u_1}{\partial y} + \frac{1}{2} \frac{\partial u_2}{\partial x} & \frac{1}{2} \frac{\partial u_1}{\partial z} + \frac{1}{2} \frac{\partial u_3}{\partial x} \\ \frac{1}{2} \frac{\partial u_1}{\partial y} + \frac{1}{2} \frac{\partial u_2}{\partial x} & \frac{\partial u_2}{\partial y} & \frac{1}{2} \frac{\partial u_2}{\partial z} + \frac{1}{2} \frac{\partial u_3}{\partial y} \\ \frac{1}{2} \frac{\partial u_1}{\partial z} + \frac{1}{2} \frac{\partial u_3}{\partial x} & \frac{1}{2} \frac{\partial u_2}{\partial z} + \frac{1}{2} \frac{\partial u_3}{\partial y} & \frac{\partial u_3}{\partial z} \end{pmatrix}$$



Lineare Elastizitätstheorie II

- ▶ ε heißt **Verzerrung** (engl. **strain**).
- ▶ σ heißt **Spannung** (engl. **stress**).
- ▶ C beschreibt das **Materialgesetz**.
- ▶ Ein wichtiger Spezialfall ist $\sigma = 2\mu\varepsilon + \lambda(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33})\mathbf{1}$ mit den sog. **Lamé-Parametern** λ und μ .
- ▶ Auf dem Rand Γ von Ω muss eine der **Randbedingungen** $\mathbf{u} = 0$ oder $\mathbf{n} \cdot \sigma = 0$ gefordert werden.
- ▶ Die Poisson- und biharmonische Gleichung entstehen aus den Gleichungen der Elastizitätstheorie durch zusätzliche vereinfachende **Modellannahmen**.
- ▶ Die Verschiebung \mathbf{u} minimiert die **Energiefunktion** $\frac{1}{2} \int_{\Omega} \varepsilon : \sigma dx - \int_{\Omega} \mathbf{f} \cdot \mathbf{u} dx$ in einer geeigneten Menge zulässiger Verformungen.



Minimalflächen

- ▶ Zu einer gegebenen Menge Ω in der (x, y) -Ebene und einer gegebenen Funktion $u_0 : \Gamma \rightarrow \mathbb{R}$ auf dem Rand Γ von Ω wird eine Fläche der Form $S = \{(x, y, u(x, y)) : (x, y) \in \Omega, u(x, y) = u_0(x, y) \text{ auf } \Gamma\}$ mit **minimalem Flächeninhalt** gesucht.
- ▶ Dann ist $u : \Omega \rightarrow \mathbb{R}$ Lösung der **Minimalflächengleichung**
 - $\text{div}(\{1 + |\nabla u|^2\}^{-\frac{1}{2}} \nabla u) = 0$ in Ω
 - $u = u_0$ auf Γ .



Gasgleichung

- ▶ Betrachte die **rotationsfreie Strömung** eines **idealen, kompressiblen Gases**.
- ▶ Dann gibt es ein skalares **Potential** u , so dass für die **Geschwindigkeit** \mathbf{v} des Gases gilt $\mathbf{v} = \nabla u$.
- ▶ Aus der **Massenerhaltung** folgt $\operatorname{div}(\rho \mathbf{v}) = 0$, wobei $\rho = \rho(\mathbf{v})$ die **Dichte** des Gases ist.
- ▶ Da das Gas ideal ist, gilt die **Zustandsgleichung**
 $\rho(\mathbf{v}) = \left[1 - \frac{\gamma - 1}{2} |\mathbf{v}|^2\right]^{\frac{1}{\gamma - 1}}$, wobei $\gamma > 1$ der **spezifische Wärmekoeffizient** ist.

- ▶ Daher erfüllt das Potential u die **Gasgleichung**

$$-\operatorname{div}\left(\left[1 - \frac{\gamma - 1}{2} |\nabla u|^2\right]^{\frac{1}{\gamma - 1}} \nabla u\right) = 0 \text{ in } \Omega$$

$$u = u_0 \text{ auf } \Gamma.$$



Wärmeleitungsgleichung I

- ▶ Die **Temperatur** $u(x, t)$ im Punkt x eines **Körpers** $\Omega \subset \mathbb{R}^3$ zur **Zeit** $t > 0$ wird unter dem Einfluss einer äußeren **Wärmequelle** $f(x, t)$ durch die **Wärmeleitungsgleichung** beschrieben:

$$\frac{\partial u}{\partial t} - \Delta u = f \text{ in } \Omega \times (0, \infty).$$

- ▶ Die anfängliche Temperaturverteilung wird beschrieben durch die **Anfangsbedingung**

$$u(x, 0) = u_0(x) \text{ für alle } x \in \Omega.$$

- ▶ Zusätzlich ist auf dem Rand Γ des Körpers eine der folgenden **Randbedingungen** zu erfüllen:

- ▶ $u(x, t) = 0$ für alle $x \in \Gamma$ und $t > 0$ (**feste Temperatur**),
- ▶ $\frac{\partial u}{\partial n}(x, t) = 0$ für alle $x \in \Gamma$ und $t > 0$ (**Isolation**).



Wärmeleitungsgleichung II

- ▶ Die Randbedingungen können wie üblich gemischt werden.
- ▶ Die rechten Seiten 0 in den Randbedingungen können durch gegebene Funktionen ersetzt werden.
- ▶ Die **Energie** $\frac{1}{2} \int_{\Omega} u(x, t)^2 dx$ ist eine monoton fallende Funktion der Zeit.



Grundwasserströmung I

- ▶ Die räumliche und zeitliche Verteilung $u(x, t)$ einer Flüssigkeit wie Grundwasser in einem Medium wie Erde wird durch die **Transport-Diffusions-Gleichung** beschrieben:

$$\frac{\partial u}{\partial t} - \operatorname{div}(D(x, u) \nabla u) + \mathbf{k}(x, u) \cdot \nabla u = f \text{ in } \Omega \times (0, \infty).$$

- ▶ Dabei beschreibt der **Quellterm** f die Zufuhr (Quelle) bzw. Entnahme (Brunnen) von Flüssigkeit.
- ▶ Die **Diffusivität** $D(x, u) \in \mathbb{R}^{3 \times 3}$ und die **Konduktivität** $\mathbf{k}(x, u) \in \mathbb{R}^3$ beschreiben spezifische Eigenschaften des Mediums (Ton, Lehm, Sand usw.).
- ▶ Wie bei der Wärmeleitungsgleichung ist die Transport-Diffusions-Gleichung durch **Anfangs-** und **Randbedingungen** zu ergänzen.



Grundwasserströmung II

- ▶ Falls der Quellterm ab einem gewissen Zeitpunkt zeitlich konstant ist, strebt die Lösung u der Transport-Diffusions-Gleichung für $t \rightarrow \infty$ gegen einen zeitlich konstanten stationären Zustand v .
- ▶ Der stationäre Zustand v wird durch die Konvektions-Diffusions-Gleichung beschrieben:
$$-\operatorname{div}(D(x, v)\nabla v) + \mathbf{k}(x, v) \cdot \nabla v = f \text{ in } \Omega.$$
- ▶ Diese ist wie die Poisson-Gleichung durch Randbedingungen zu ergänzen.

73 / 264



Wellengleichung

- ▶ Die zeitlich veränderliche vertikale Auslenkung u einer elastischen, nicht dehnbaren Membran unter Einfluss einer vertikalen, zeitlich veränderlichen Last f wird durch die Wellen-Gleichung beschrieben:

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = f \text{ in } \Omega \times (0, \infty).$$

- ▶ Der Anfangszustand wird beschrieben durch die zwei Anfangsbedingungen:

$$u(x, 0) = u_0(x) \text{ und } \frac{\partial}{\partial t} u(x, 0) = u_1(x) \text{ für alle } x \text{ in } \Omega.$$

- ▶ Die Wellengleichung ist durch Randbedingungen zu ergänzen.
- ▶ Die Energie $\frac{1}{2} \int_{\Omega} \left\{ \left(\frac{\partial u}{\partial t} \right)^2 + |\nabla u|^2 \right\} dx$ ist zeitlich konstant.

74 / 264



Ordnung

- ▶ Die Ordnung einer partiellen Differentialgleichung ist die höchste Differentiationsstufe der in der Gleichung auftretenden partiellen Ableitungen.
- ▶ Die biharmonische Gleichung hat die Ordnung 4.
- ▶ Alle anderen Beispiele haben die Ordnung 2.
- ▶ Bei einer Differentialgleichung der Ordnung $2k$ sind typischerweise k Randbedingungen zu fordern.

75 / 264



Lineare und nichtlineare Gleichungen

- ▶ Eine partielle Differentialgleichung heißt linear, wenn eine Überlagerung oder Skalierung der Last zu einer entsprechenden Überlagerung bzw. Skalierung der Lösung führt, d.h. die Zuordnung Last \rightarrow Lösung ist eine lineare Funktion.
- ▶ Die Poisson-Gleichung, die biharmonische Gleichung, die Gleichungen der linearen Elastizitätstheorie, die Wärmeleitungsgleichung und die Wellengleichung sind linear, alle anderen Beispiele sind nichtlinear.
- ▶ Die Transport-Diffusions- und die Konvektions-Diffusions-Gleichung sind linear, wenn die Diffusivität und die Konduktivität nicht von der Lösung u abhängen.

76 / 264



Elliptisch, parabolisch, hyperbolisch I

- ▶ Die allgemeine Form einer linearen partiellen Differentialgleichung zweiter Ordnung ist
 $A(x) : D^2u + \mathbf{a}(x) \cdot \nabla u + \alpha(x)u = f.$
- ▶ Wegen der Symmetrie der **Hesse-Matrix** D^2u kann die Matrix $A(x)$ als symmetrisch vorausgesetzt werden.
- ▶ Die Differentialgleichung heißt
 - ▶ **elliptisch**, wenn für alle x alle Eigenwerte von $A(x)$ ungleich Null sind und gleiches Vorzeichen haben,
 - ▶ **parabolisch**, wenn für alle x genau ein Eigenwert von $A(x)$ gleich Null ist und alle anderen Eigenwerte gleiches Vorzeichen haben,
 - ▶ **hyperbolisch**, wenn für alle x alle Eigenwerte von $A(x)$ ungleich Null sind und genau ein Eigenwert anderes Vorzeichen hat als die restlichen Eigenwerte.

77 / 264



Elliptisch, parabolisch, hyperbolisch II

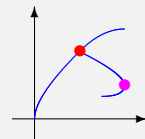
- ▶ Die Wellengleichung ist hyperbolisch.
- ▶ Die Wärmeleitungsgleichung und die Transport-Diffusions-Gleichung sind parabolisch.
- ▶ Die Gasgleichung ist elliptisch, falls der Gradient des Potentials hinreichend klein ist (**Unterschallströmung**).
- ▶ Alle anderen Beispiele sind elliptisch.
- ▶ Elliptische Gleichungen beschreiben häufig ein **Variations- oder Minimierungsproblem**.
- ▶ Parabolische Gleichungen beschreiben häufig ein **Dissipationsphänomen**, bei dem eine Energie monoton fällt.
- ▶ Hyperbolische Gleichungen beschreiben häufig einen **Erhaltungssatz**.

78 / 264



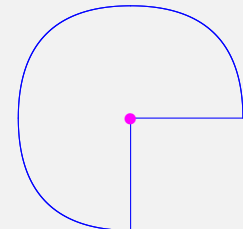
Existenz-, Eindeutigkeit und Regularität

- ▶ Lineare Differentialgleichungen haben in der Regel eine eindeutige Lösung.
- ▶ Nichtlineare Differentialgleichungen können mehrere Lösungen zulassen, insbesondere können **Verzweigungen** und **Umkehrpunkte** auftreten.
- ▶ **Anders als bei gewöhnlichen Differentialgleichungen hängt die Regularität (Differenzierbarkeit) der Lösung einer partiellen Differentialgleichung von Eigenschaften des Randes Γ ab. Einspringende Ecken führen zu einem Regularitätsverlust.**



Beispiel

- ▶ Betrachte die Poisson-Gleichung $-\Delta u = 0$ auf dem Kreissegment $\Omega = \{(r \cos \varphi, r \sin \varphi) : 0 \leq r < 1, 0 \leq \varphi \leq \frac{3\pi}{2}\}$ mit Randbedingung $u = \sin(\frac{2}{3}\varphi)$ auf dem Kreisbogen und $u = 0$ auf den geraden Randstücken.
- ▶ Die Lösung ist $u = r^{\frac{2}{3}} \sin(\frac{2}{3}\varphi)$.
- ▶ Die Lösung hat in der Nähe der einspringenden Ecke $(0, 0)$ keine beschränkte Ableitung.



79 / 264

80 / 264



Diskretisierungsmethoden

- ▶ **Differenzenverfahren** ersetzen Ableitungen durch **Differenzenquotienten** und fordern die resultierenden Gleichungen nur in den Punkten eines (regelmäßigen) **Gitters**.
- ▶ **Finite-Element-Methoden** basieren auf einer **Variationsformulierung** der Differentialgleichung und approximieren die dabei auftretenden Funktionen durch **stetige, stückweise Polynome** auf einer **Unterteilung** des Gebietes Ω in einfache Teilgebiete wie Dreiecke oder Vierecke.
- ▶ **Finite-Volumen-Methoden** basieren auf einer **Erhaltungsgleichung** und erfüllen diese für **stückweise konstante Funktionen** auf einfachen Teilgebieten von Ω , den sog. **Kontrollvolumina**.



Differenzenverfahren für partielle Differentialgleichungen

- ▶ Elliptische Differentialgleichungen
- ▶ Parabolische Differentialgleichungen
- ▶ Hyperbolische Differentialgleichungen



Reaktions-Diffusions-Gleichung

$$\begin{aligned} -\operatorname{div}(A\nabla u) + \alpha u &= f && \text{in } \Omega \\ u &= 0 && \text{auf } \Gamma \end{aligned}$$

- ▶ $\Omega \subset \mathbb{R}^d$ mit $d = 2$ oder $d = 3$
- ▶ $A(x)$ eine symmetrische, positiv definite, $d \times d$ Matrix für jedes x in Ω
- ▶ $\alpha(x)$ eine nicht-negative Zahl für jedes x in Ω

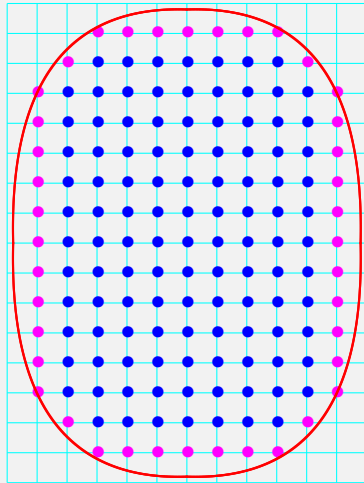


Idee

- ▶ Ersetze alle partiellen Ableitungen durch **Differenzenquotienten**.
- ▶ Fordere die resultierenden Gleichungen nur für die Punkte eines (regelmäßigen) **Gitters**.



Gitter



- ▶ $h > 0$ Gitterweite
- ▶ $G_h = \{ih : i \in \mathbb{Z}^d\}$
- ▶ $\bar{\Omega}_h = G_h \cap \bar{\Omega}$
- ▶ $\Gamma_h = \{x \in \bar{\Omega}_h : \min_{y \in \Gamma} |x - y| < h\}$
diskreter Rand
- ▶ $\Omega_h = \bar{\Omega}_h \setminus \Gamma_h$ diskretes Gebiet

85 / 264



Nummerierung

- ▶ $N_h = \#\Omega_h$ Zahl der Punkte im diskreten Gebiet
- ▶ $\bar{N}_h = \#\bar{\Omega}_h$ Zahl aller Gitterpunkte
- ▶ $\bar{N}_h - N_h$ Zahl der diskreten Randpunkte
- ▶ Nummeriere zuerst die Punkte in Ω_h **lexikographisch**, d.h. zeilenweise von links nach rechts beginnend mit der obersten Zeile.
- ▶ Nummeriere dann fortlaufend die Punkte in Γ_h .
- ▶ Es ist $N_h \approx h^{-d}$ und $\bar{N}_h - N_h \approx h^{-(d-1)} \approx N_h^{\frac{d-1}{d}}$.

86 / 264



Differenzenquotienten

- ▶ e_i i -te Einheitsvektor
(i -te Komponente gleich 1, alle anderen Komponenten 0)
- ▶ $\partial_{h,i}^+ u(x) = \frac{1}{h} [u(x + he_i) - u(x)]$ i -te vorwärts
Differenzenquotient
- ▶ $\partial_{h,i}^- u(x) = \frac{1}{h} [u(x) - u(x - he_i)]$ i -te rückwärts
Differenzenquotient
- ▶ Taylor-Entwicklung liefert mit geeignetem $\theta \in (0, 1)$:

$$\partial_{h,i}^\pm u(x) = \frac{\partial}{\partial x_i} u(x) \pm \frac{1}{2} h \frac{\partial^2}{\partial x_i^2} u(x \pm \theta h e_i)$$

$$\begin{aligned} \partial_{h,i}^+ (\partial_{h,i}^- u)(x) &= \partial_{h,i}^- (\partial_{h,i}^+ u)(x) \\ &= \frac{\partial^2}{\partial x_i^2} u(x) + \frac{1}{12} h^2 \frac{\partial^4}{\partial x_i^4} u(x \pm \theta h e_i) \end{aligned}$$

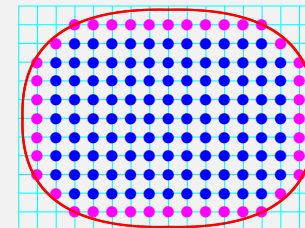
87 / 264



Differenzdiskretisierung

Bestimme den Vektor $u_h = (u_h(x))_{x \in \bar{\Omega}_h}$ so, dass gilt:

- ▶ $u_h(x) = 0$
für alle Punkte x auf dem **diskreten Rand** Γ_h
- ▶ $-\sum_{i=1}^d \sum_{j=1}^d \partial_{h,i}^- (A_{i,j} \partial_{h,j}^+ u)(x) + \alpha(x) u(x) = f(x)$
für alle Punkte x im diskreten Gebiet Ω_h



88 / 264



Eigenschaften

- ▶ Die Differenzdiskretisierung führt auf ein **lineares Gleichungssystem** mit N_h Gleichungen für die N_h Unbekannten $u_h(x)$, $x \in \Omega_h$.
- ▶ Die Matrix L_h ist symmetrisch, positiv definit.
- ▶ Die Matrix ist **dünn besetzt**, pro Zeile sind höchstens 3^d Elemente von Null verschieden.
- ▶ Die Diagonalelemente sind positiv.
- ▶ Die Elemente außerhalb der Diagonalen verschwinden oder sind negativ.
- ▶ Die Matrix hat **Bandstruktur**, die **Bandbreite** ist $\approx N_h^{1-\frac{1}{d}}$.
- ▶ Die Lösung mit **Gauß-Elimination** oder **Cholesky-Zerlegung** erfordert $\approx N_h^{3-\frac{2}{d}}$ Operationen und $\approx N_h^{2-\frac{1}{d}}$ Speicherplätze.

89 / 264



Beispiel

h	d = 2		d = 3	
	Speicher	Aufwand	Speicher	Aufwand
$\frac{1}{16}$	$3.3 \cdot 10^3$	$7.6 \cdot 10^5$	$7.6 \cdot 10^5$	$1.7 \cdot 10^8$
$\frac{1}{32}$	$2.9 \cdot 10^4$	$2.8 \cdot 10^7$	$2.8 \cdot 10^7$	$2.8 \cdot 10^{10}$
$\frac{1}{64}$	$2.5 \cdot 10^5$	$9.9 \cdot 10^8$	$9.9 \cdot 10^8$	$3.9 \cdot 10^{12}$
$\frac{1}{128}$	$2.0 \cdot 10^6$	$3.3 \cdot 10^{10}$	$3.3 \cdot 10^{10}$	$5.3 \cdot 10^{14}$

90 / 264



Fehlerabschätzung

- ▶ Für beliebige Diffusion A gilt

$$\max_{x \in \Omega_h} |u(x) - u_h(x)| \leq c_1 h.$$

Die Konstante c_1 hängt von Ableitungen bis zur Ordnung 2 von A und von Ableitungen von u bis zur Ordnung 3 ab.

- ▶ Ist die Diffusion A eine konstante Diagonalmatrix, so gilt

$$\max_{x \in \Omega_h} |u(x) - u_h(x)| \leq c_2 h^2.$$

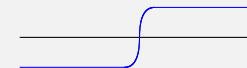
Die Konstante c_2 hängt von Ableitungen von u bis zur Ordnung 4 ab.

91 / 264



Grenzen

- ▶ Die Differenzierbarkeitsanforderungen an die Lösung u der Differentialgleichung sind unrealistisch.
- ▶ In der Praxis variiert die Lösung u in weiten Teilen des Gebietes kaum, während sie sich in relativ kleinen Teilen stark ändert.
 - ▶ Ein gleichmäßiges Gitter vergeudet die Unbekannten in den uninteressanten Teilen des Gebietes.
 - ▶ Die Gitterpunkte sollten in den Teilen des Gebietes konzentriert werden, in denen sich die Lösung stark ändert.
 - ▶ **Problem:** Die relevanten Bereiche des Gebietes sind Teil der Lösung und nicht von vornherein bekannt.

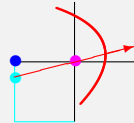


92 / 264



Neumann-Randbedingung

- ▶ Gegeben seien ein **diskreter Randpunkt** x und ein **Einheitsvektor** ν .
 - ▶ Gesucht ist eine Approximation $\partial_{h,\nu}u(x)$ für $\nu \cdot \nabla u(x)$.
 - ▶ Bestimme den Schnittpunkt y der **Geraden** durch x in Richtung ν mit dem Rand γ_x des Quadrates mit Mittelpunkt x und Kantenlänge $2h$.
 - ▶ Bestimme den am dichtesten an y liegenden Gitterpunkt z auf γ_x .
- $$\partial_{h,\nu}u(x) = \frac{1}{|x-z|} [u(x) - u(z)]$$



93 / 264



Konvektionsterm I

- ▶ Gegeben seien ein Punkt x im **diskreten Gebiet** und ein Vektor $\mathbf{a} \neq 0$.
- ▶ Gesucht ist eine Approximation für den **Konvektionsterm** $\mathbf{a} \cdot \nabla u(x) = a_1 \frac{\partial u}{\partial x_1}(x) + \dots + a_d \frac{\partial u}{\partial x_d}(x)$.
- ▶ Für die Diskretisierung von $a_i \frac{\partial u}{\partial x_i}(x)$ stehen zur Auswahl:
 - ▶ **vorwärts Differenzenquotient**
 $a_i \partial_{h,i}^+ u(x) = a_i \frac{1}{h} [u(x + he_i) - u(x)]$
 - ▶ **rückwärts Differenzenquotient**
 $a_i \partial_{h,i}^- u(x) = a_i \frac{1}{h} [u(x) - u(x - he_i)]$
 - ▶ **symmetrischer Differenzenquotient**
 $a_i \frac{1}{2} [\partial_{h,i}^+ u(x) + \partial_{h,i}^- u(x)] = a_i \frac{1}{2h} [u(x + he_i) - u(x - he_i)]$

94 / 264



Konvektionsterm II

- ▶ Falls die **Péclet-Zahl** $\frac{a_i h}{\nu}$ (ν kleinster Eigenwert der Diffusionsmatrix $A(x)$) größer ist als 1, kann jede Wahl zu unphysikalischen Oszillationen der numerischen Lösung führen.
- ▶ Die Oszillationen treten auf, wenn die Matrix der Differenzendiskretisierung negative Diagonalelemente und positive Nicht-Diagonalelemente aufweist.
- ▶ Daher ist je nach Vorzeichen von a_i der rückwärts- oder vorwärts Differenzenquotient zu wählen (**upwinding**):

$$a_i \partial_{h,i}^u u(x) = \begin{cases} a_i \frac{1}{h} [u(x) - u(x - he_i)] & \text{falls } a_i > 0 \\ a_i \frac{1}{h} [u(x + he_i) - u(x)] & \text{falls } a_i < 0 \end{cases}$$

95 / 264



Wärmeleitungsgleichung

$$\begin{aligned} \frac{\partial u}{\partial t} - \operatorname{div}(A \nabla u) + \alpha u &= f && \text{in } \Omega \times (0, \infty) \\ u &= 0 && \text{auf } \Gamma \times (0, \infty) \\ u(x, 0) &= u_0(x) && \text{in } \Omega \end{aligned}$$

- ▶ $\Omega \subset \mathbb{R}^d$ mit $d = 2$ oder $d = 3$
- ▶ $A(x)$ eine **zeitlich konstante**, symmetrische, positiv definite, $d \times d$ Matrix für jedes x in Ω
- ▶ $\alpha(x)$ eine **zeitlich konstante**, nicht-negative Zahl für jedes x in Ω

96 / 264



Idee

- ▶ Diskretisiere die **Ortsableitungen** wie bei der Reaktions-Diffusions-Gleichung.
- ▶ Dies führt auf ein System **gewöhnlicher Differentialgleichungen** der Form

$$\dot{u}_h(t) = f_h(t) - L_h u_h(t) \text{ für } t > 0,$$

$$u_h(0) = u_{0,h}.$$
- ▶ Ersetze die Zeitableitung durch einen rückwärtigen Differenzenquotienten und setze diesen gleich einer Konvexkombination der rechten Seite zu den entsprechenden Zeiten (θ -Schema).

97 / 264



Differenzendiskretisierung

- ▶ Wähle eine **Ortschrittweite** $h > 0$, eine **Zeitschrittweite** $\tau > 0$ und einen Parameter $\theta \in [0, 1]$.
- ▶ Bezeichne mit L_h die Matrix der Differenzendiskretisierung der **Reaktion-Diffusions-Gleichung** und setze $f_h^n = f(x, n\tau)$ für alle $x \in \Omega_h$.
- ▶ Setze $u_h^0 = u_0(x)$ für alle $x \in \bar{\Omega}_h$.
- ▶ Bestimme u_h^n für $n = 1, 2, \dots$ sukzessive durch

$$\frac{1}{\tau} (u_h^n - u_h^{n-1}) = \theta (f_h^n - L_h u_h^n) + (1 - \theta) (f_h^{n-1} - L_h u_h^{n-1}).$$

98 / 264



Eigenschaften

- ▶ $\theta = 0$ entspricht dem **expliziten Euler-Verfahren**, $\theta = 1$ dem **impliziten Euler-Verfahren** und $\theta = \frac{1}{2}$ dem **Verfahren von Crank-Nicolson**.
- ▶ Für $\theta = 0$ erfordert die Berechnung von u_h^n lediglich eine **Matrix-Vektor-Multiplikation**.
- ▶ Für $\theta > 0$ erfordert die Berechnung von u_h^n das **Lösen eines linearen Gleichungssystems** mit symmetrisch, positiv definiten Matrix $\frac{1}{\tau} I + L_h$.
- ▶ Für $\theta < \frac{1}{2}$ muss für Orts- und Zeitschritt die **CFL-Bedingung** $\tau \lesssim h^2$ erfüllt sein.
- ▶ Für $\theta \geq \frac{1}{2}$ können Orts- und Zeitschritt unabhängig voneinander gewählt werden.

99 / 264



Fehlerabschätzung

- ▶ Setze $\gamma = \begin{cases} 1 & \text{für } \theta \neq \frac{1}{2} \\ 2 & \text{für } \theta = \frac{1}{2} \end{cases}$.
- ▶ Für beliebige Diffusion A gilt

$$\max_{x \in \Omega_h, n \geq 0} |u(x, n\tau) - u_h^n(x)| \leq c_1 (\tau^\gamma + h).$$
 Die Konstante c_1 hängt von Ableitungen bis zur Ordnung 2 von A und von Ableitungen von u bis zur Ordnung 3 ab.
- ▶ Ist die Diffusion A eine konstante Diagonalmatrix, so gilt

$$\max_{x \in \Omega_h, n \geq 0} |u(x, n\tau) - u_h^n(x)| \leq c_2 (\tau^\gamma + h^2).$$
 Die Konstante c_2 hängt von Ableitungen von u bis zur Ordnung 4 ab.

100 / 264



Wellengleichung

$$\frac{\partial^2 u}{\partial t^2} - \operatorname{div}(A \nabla u) + \alpha u = f \quad \text{in } \Omega \times (0, \infty)$$

$$u = 0 \quad \text{auf } \Gamma \times (0, \infty)$$

$$u(x, 0) = u_0(x) \quad \text{in } \Omega$$

$$\frac{\partial u}{\partial t}(x, 0) = u_1(x) \quad \text{in } \Omega$$

- ▶ $\Omega \subset \mathbb{R}^d$ mit $d = 2$ oder $d = 3$
- ▶ $A(x)$ eine zeitlich konstante, symmetrische, positiv definite, $d \times d$ Matrix für jedes x in Ω
- ▶ $\alpha(x)$ eine zeitlich konstante, nicht-negative Zahl für jedes x in Ω

101 / 264



Idee

- ▶ Diskretisiere die Ortsableitungen wie bei der Reaktions-Diffusions-Gleichung.
- ▶ Dies führt auf ein System gewöhnlicher Differentialgleichungen der Form

$$\begin{aligned} \dot{u}_h(t) &= f_h(t) - L_h u_h(t) \quad \text{für } t > 0, \\ u_h(0) &= u_{0,h}, \\ \dot{u}_h(0) &= u_{1,h}. \end{aligned}$$
- ▶ Ersetze die Zeitableitung durch einen symmetrischen Differenzenquotienten und setze diesen gleich der rechten Seite zur mittleren Zeit.

102 / 264



Differenzdiskretisierung

- ▶ Wähle eine Ortsschrittweite $h > 0$ und eine Zeitschrittweite $\tau > 0$.
- ▶ Bezeichne mit L_h die Matrix der Differenzdiskretisierung der Reaktions-Diffusions-Gleichung und setze $f_h^n = f(x, n\tau)$ für alle $x \in \Omega_h$.
- ▶ Setze $u_h^0 = u_0(x)$ für alle $x \in \bar{\Omega}_h$.
- ▶ Setze $u_h^1 = u_h^0(x) + \tau u_1(x)$ für alle $x \in \bar{\Omega}_h$.
- ▶ Bestimme u_h^n für $n = 2, 3, \dots$ sukzessive durch

$$\frac{1}{\tau^2} (u_h^n - 2u_h^{n-1} + u_h^{n-2}) = (f_h^{n-1} - L_h u_h^{n-1}).$$

103 / 264



Eigenschaften

- ▶ Die Berechnung von u_h^n erfordert lediglich eine Matrix-Vektor-Multiplikation.
- ▶ Für Orts- und Zeitschritt muss die CFL-Bedingung $\tau \lesssim h$ erfüllt sein.

104 / 264



Fehlerabschätzung

- ▶ Für beliebige Diffusion A gilt

$$\max_{x \in \Omega_h, n \geq 0} |u(x, n\tau) - u_h^n(x)| \leq c_1(\tau + h).$$

Die Konstante c_1 hängt von Ableitungen bis zur Ordnung 2 von A und von Ableitungen von u bis zur Ordnung 3 ab.

- ▶ Ist die Diffusion A eine konstante Diagonalmatrix, so gilt

$$\max_{x \in \Omega_h, n \geq 0} |u(x, n\tau) - u_h^n(x)| \leq c_2(\tau^2 + h^2).$$

Die Konstante c_2 hängt von Ableitungen von u bis zur Ordnung 4 ab.



Ergänzungen

- ▶ Zeitabhängige Diffusions- und Reaktionskoeffizienten werden bei parabolischen und hyperbolischen Gleichungen zu den gleichen diskreten Zeiten ausgewertet wie die entsprechenden u_h -Terme.
- ▶ Konvektionsterme und Neumann-Randbedingungen werden bei parabolischen und hyperbolischen Gleichungen diskretisiert wie bei elliptischen Gleichungen; die Matrix L_h ändert sich entsprechend.



Finite-Element-Methoden für elliptische Differentialgleichungen

- ▶ Variationsformulierung
- ▶ Finite-Element-Diskretisierung
- ▶ Praktische Aspekte
- ▶ Upwind und Petrov-Galerkin-Verfahren
- ▶ Gemischte Finite-Element-Methoden



Reaktions-Diffusions-Gleichung

$$\begin{aligned} -\operatorname{div}(A\nabla u) + \alpha u &= f && \text{in } \Omega \\ u &= 0 && \text{auf } \Gamma \end{aligned}$$

- ▶ Ω ein **Polyeder** in \mathbb{R}^d mit $d = 2$ oder $d = 3$
- ▶ $A(x)$ eine symmetrische, positiv definite, $d \times d$ Matrix für jedes x in Ω
- ▶ $\alpha(x)$ eine nicht-negative Zahl für jedes x in Ω



Satz von Gauß

- ▶ Divergenz:

$$\operatorname{div} \mathbf{w} = \sum_{i=1}^d \frac{\partial w_i}{\partial x_i}$$

- ▶ Satz von Gauß:

$$\int_{\Omega} \operatorname{div} \mathbf{w} dx = \int_{\Gamma} \mathbf{w} \cdot \mathbf{n} dS$$

109 / 264



Mehrdimensionale partielle Integration I

- ▶ Satz von Gauß angewandt auf $\mathbf{w} = v(A\nabla u)$ liefert

$$\begin{aligned} & \int_{\Omega} v \operatorname{div}(A\nabla u) dx + \int_{\Omega} \nabla v \cdot A\nabla u dx \\ &= \int_{\Omega} \operatorname{div}(vA\nabla u) dx = \int_{\Omega} \operatorname{div} \mathbf{w} dx = \int_{\Gamma} \mathbf{w} \cdot \mathbf{n} dS \\ &= \int_{\Gamma} v \mathbf{n} \cdot A\nabla u dS \end{aligned}$$

- ▶ Falls $v = 0$ ist auf Γ , folgt

$$\int_{\Omega} \nabla v \cdot A\nabla u dx = - \int_{\Omega} v \operatorname{div}(A\nabla u) dx$$

110 / 264



Idee der Variationsformulierung

- ▶ Multipliziere die Differentialgleichung mit einer stetig differenzierbaren Funktion v mit $v = 0$ auf Γ .
 $-\operatorname{div}(A\nabla u)(x)v(x) + \alpha(x)u(x)v(x) = f(x)v(x)$ für $x \in \Omega$.

- ▶ Integriere das Ergebnis über Ω

$$\int_{\Omega} [-\operatorname{div}(A\nabla u)v + \alpha uv] dx = \int_{\Omega} f v dx.$$

- ▶ Integriere den Ableitungsterm partiell

$$-\int_{\Omega} \operatorname{div}(A\nabla u)v dx = \int_{\Omega} \nabla v \cdot A\nabla u dx$$

111 / 264



Probleme

- ▶ Für eine sinnvolle Variationsformulierung müssen die Eigenschaften der Funktionen u und v präziser gefasst werden.
- ▶ Klassische Eigenschaften wie stetige Differenzierbarkeit sind zu restriktiv.
- ▶ Der Begriff der Ableitung muss daher geeignet verallgemeinert werden.
- ▶ In Hinblick auf die Diskretisierung sollten insbesondere stückweise differenzierbare Funktionen im erweiterten Sinn differenzierbar sein.

112 / 264



Mehrdimensionale partielle Integration II

- ▶ Satz von Gauß angewandt auf $\mathbf{w} = uv\mathbf{e}_i$ (\mathbf{e}_i i -te Einheitsvektor mit i -ter Komponente 1 und restlichen Komponenten 0) liefert

$$\begin{aligned} & \int_{\Omega} \frac{\partial u}{\partial x_i} v dx + \int_{\Omega} u \frac{\partial v}{\partial x_i} dx \\ &= \int_{\Omega} \frac{\partial(uv)}{\partial x_i} dx = \int_{\Omega} \operatorname{div} \mathbf{w} dx = \int_{\Gamma} \mathbf{w} \cdot \mathbf{n} dS \\ &= \int_{\Gamma} uv \mathbf{n}_i dS \end{aligned}$$

- ▶ Falls $u = 0$ oder $v = 0$ auf Γ ist, folgt

$$\int_{\Omega} \frac{\partial u}{\partial x_i} v dx = - \int_{\Omega} u \frac{\partial v}{\partial x_i} dx$$

113 / 264



Schwache Ableitung

- ▶ Die Funktion u heißt **schwach differenzierbar** (bzgl. x_i) mit **schwacher Ableitung** w_i , wenn für jede stetig differenzierbare Funktion v mit $v = 0$ auf Γ gilt

$$\int_{\Omega} w_i v dx = - \int_{\Omega} u \frac{\partial v}{\partial x_i} dx$$

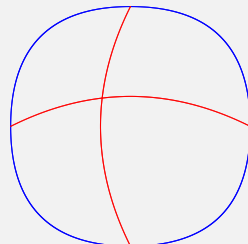
- ▶ Ist u bzgl. jeder der Variablen x_1, \dots, x_d schwach differenzierbar, so nennt man u **schwach differenzierbar** und schreibt ∇u für den Vektor (w_1, \dots, w_d) der schwachen Ableitungen.

114 / 264



Beispiele

- ▶ Jede stetig differenzierbare Funktion ist schwach differenzierbar und die schwache Ableitung ∇u stimmt mit dem klassischen Gradienten überein.
- ▶ Eine stückweise stetig differenzierbare Funktion ist genau dann schwach differenzierbar, wenn sie global stetig ist; dann stimmt die schwache Ableitung ∇u mit dem stückweise definierten klassischen Gradienten überein.



115 / 264



Sobolev-Räume

- ▶ $\|v\| = \left\{ \int_{\Omega} |v|^2 dx \right\}^{\frac{1}{2}}$ ist die L^2 -Norm.
- ▶ $L^2(\Omega)$ ist der **Lebesgue-Raum** aller Funktionen v mit endlicher L^2 -Norm $\|v\|$.
- ▶ $H^1(\Omega)$ ist der **Sobolev-Raum** aller Funktionen v in $L^2(\Omega)$, die schwach differenzierbar sind und für die $|\nabla v|$, die euklidische Norm von ∇v , in $L^2(\Omega)$ ist.
- ▶ $H_0^1(\Omega)$ ist der **Sobolev-Raum** aller Funktionen v in $H^1(\Omega)$ mit $v = 0$ auf Γ .

116 / 264



Beispiele

- ▶ Jede beschränkte Funktion ist in $L^2(\Omega)$.
- ▶ $v(x) = \frac{1}{\sqrt{x^2+y^2}}$ ist nicht in $L^2(B(0,1))$ ($B(0,1)$ Kreis mit Radius 1 und Mittelpunkt im Ursprung), da
$$\int_{B(0,1)} |v(x)|^2 dx = 2\pi \int_0^1 \frac{1}{r} dr$$
 nicht endlich ist.
- ▶ Jede stetig differenzierbare Funktion ist in $H^1(\Omega)$.
- ▶ Eine stückweise stetig differenzierbare Funktion ist genau dann in $H^1(\Omega)$, wenn sie global stetig ist.
- ▶ **Punktwerte sind für Funktionen in $H^1(\Omega)$ nicht definiert.**
 $v(x) = \ln(|\ln(\sqrt{x^2+y^2})|)$ ist in $H^1(B(0,1))$, besitzt aber keinen endlichen Wert im Ursprung.

117 / 264



Variationsproblem

Finde $u \in H_0^1(\Omega)$ so, dass für alle $v \in H_0^1(\Omega)$ gilt

$$\int_{\Omega} [\nabla v \cdot A \nabla u + \alpha uv] dx = \int_{\Omega} f v dx$$

118 / 264



Eigenschaften des Variationsproblems

- ▶ Das Variationsproblem hat eine eindeutige Lösung.
- ▶ Die Lösung des Variationsproblems ist das eindeutige **Minimum** in $H_0^1(\Omega)$ der **Energiefunktion**
$$\frac{1}{2} \int_{\Omega} [\nabla u \cdot A \nabla u + \alpha u^2] dx - \int_{\Omega} f u dx.$$

119 / 264



Idee

- ▶ Zerlege Ω in nicht überlappende, einfache Teilgebiete, sog. **Elemente**, wie Dreiecke, Parallelogramme, Tetraeder, Parallelepipede, ... (**Unterteilung**).
- ▶ Ersetze in der Variationsformulierung den Raum $H_0^1(\Omega)$ durch einen **endlich dimensionalen Unterraum** bestehend aus **stetigen** Funktionen, die stückweise auf den Elementen Polynome sind (**Finite-Element-Raum**).
- ▶ Dies führt auf ein **lineares Gleichungssystem** für die Approximation $u_{\mathcal{T}}$ an die Lösung u der partiellen Differentialgleichung.

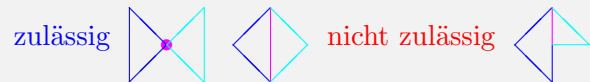
120 / 264



Unterteilung

$\mathcal{T} = \{K_i : 1 \leq i \leq N_{\mathcal{T}}\}$ bezeichnet eine **Unterteilung** von Ω mit folgenden Eigenschaften:

- ▶ Ω ist die Vereinigung aller Elemente K in \mathcal{T} .
- ▶ **Zulässigkeit:** Je zwei Elemente K und K' in \mathcal{T} sind entweder disjunkt oder haben einen Eckpunkt oder eine ganze Kante oder, falls $d = 3$ ist, eine ganze Seitenfläche **gemeinsam**.



- ▶ **Affine Äquivalenz:** Jedes Element K ist ein Dreieck oder Parallelogramm, falls $d = 2$ ist, oder ein Tetraeder oder Parallelepiped, falls $d = 3$ ist.



Bemerkungen

- ▶ Wenn Ω kein **Polyeder** ist, kann man es nicht durch geradlinige Elemente wie Dreiecke überdecken, so dass der **Rand** von Ω zusätzlich **approximiert** werden muss.
- ▶ Die Zulässigkeit wird benötigt, damit die Finite-Element-Räume in $H_0^1(\Omega)$ enthalten sind.
- ▶ Bei fehlender Zulässigkeit muss die Inklusion der Finite-Element-Räume explizit erzwungen werden, was die Implementierung erschwert.
- ▶ Es können auch allgemeine Vierecke und Quader betrachtet werden, was die Implementierung erschwert.



Finite-Element-Räume

- ▶ $R_k(K) = \begin{cases} \text{span}\{x_1^{\alpha_1} \cdot \dots \cdot x_d^{\alpha_d} : \alpha_1 + \dots + \alpha_d \leq k\} \\ K \text{ Dreieck oder Tetraeder} \\ \text{span}\{x_1^{\alpha_1} \cdot \dots \cdot x_d^{\alpha_d} : \max\{\alpha_1, \dots, \alpha_d\} \leq k\} \\ K \text{ Parallelogram oder Parallelepiped} \end{cases}$
- ▶ $S^{k,-1}(\mathcal{T}) = \{v : \Omega \rightarrow \mathbb{R} : v|_K \in R_k(K) \text{ für alle } K \in \mathcal{T}\}$
- ▶ $S^{k,0}(\mathcal{T}) = S^{k,-1}(\mathcal{T}) \cap C(\bar{\Omega})$
- ▶ $S_0^{k,0}(\mathcal{T}) = S^{k,0}(\mathcal{T}) \cap H_0^1(\Omega)$
 $= \{v \in S^{k,0}(\mathcal{T}) : v = 0 \text{ auf } \Gamma\}$



Bemerkungen

- ▶ Wegen der globalen Stetigkeit ist $S^{k,0}(\mathcal{T}) \subset H^1(\Omega)$.
- ▶ Der Polynomgrad k kann auch von Element zu Element variieren, was auf einen höheren Aufwand für die Implementierung führt.



Diskretes Problem

Finde $u_{\mathcal{T}} \in S_0^{k,0}(\mathcal{T})$ (Ansatzfunktion) so, dass für alle $v_{\mathcal{T}} \in S_0^{k,0}(\mathcal{T})$ (Testfunktion) gilt

$$\int_{\Omega} [\nabla v_{\mathcal{T}} \cdot A \nabla u_{\mathcal{T}} + \alpha u_{\mathcal{T}} v_{\mathcal{T}}] dx = \int_{\Omega} f v_{\mathcal{T}} dx$$

125 / 264



Eigenschaften des diskreten Problems

- ▶ Das diskrete Problem hat eine eindeutige Lösung.
- ▶ Die Lösung des diskreten Problems ist das eindeutige Minimum in $S_0^{k,0}(\mathcal{T})$ der Energiefunktion $\frac{1}{2} \int_{\Omega} [\nabla u \cdot A \nabla u + \alpha u^2] dx - \int_{\Omega} f u dx$.
- ▶ Das diskrete Problem führt nach Wahl einer Basis für $S_0^{k,0}(\mathcal{T})$ auf ein lineares Gleichungssystem.
- ▶ Die Matrix, genannt Systemsteifigkeits- oder Steifigkeitsmatrix, ist symmetrisch, positiv definit.
- ▶ Der Vektor der rechten Seite wird häufig als Lastvektor bezeichnet.
- ▶ Die Zahl der Gleichungen und Unbekannten ist $\approx k^d N_{\mathcal{T}}$ ($N_{\mathcal{T}} = \#\mathcal{T}$).

126 / 264



A priori Fehlerabschätzungen

- ▶ Bezeichne mit $h_{\mathcal{T}}$ den maximalen Durchmesser aller Elemente in \mathcal{T} .
- ▶ Dann gelten folgende a priori Fehlerabschätzungen für die Lösung u des Variationsproblems und die Lösung $u_{\mathcal{T}}$ des Finite-Element-Problems:

$$\|\nabla u - \nabla u_{\mathcal{T}}\| \leq c_1 h_{\mathcal{T}}^k$$

$$\|u - u_{\mathcal{T}}\| \leq c_2 h_{\mathcal{T}}^{k+1} \text{ (Falls } \Omega \text{ konvex ist.)}$$

- ▶ Die Konstanten c_1 und c_2 hängen von Ω , der Diffusion A , der Reaktion α und den L^2 -Normen der Ableitungen bis zur Ordnung $k - 1$ von f ab.

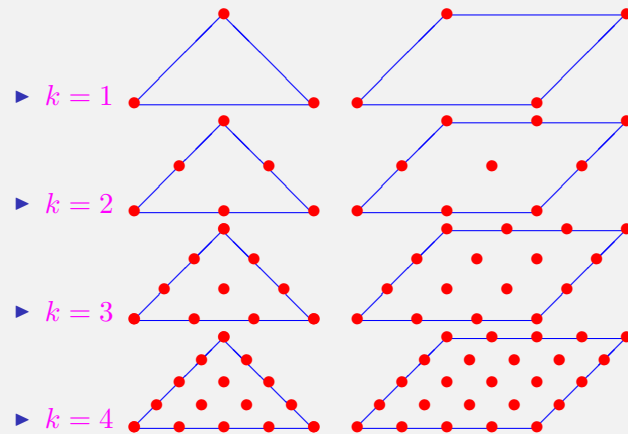
127 / 264



Praktische Aspekte

- ▶ Wahl einer Basis für $S_0^{k,0}(\mathcal{T})$
- ▶ Aufstellen der Steifigkeitsmatrix und des Lastvektors
- ▶ Berechnung der dabei auftretenden Integrale
- ▶ Behandlung gekrümmter Ränder
- ▶ Behandlung von Neumann-Randbedingungen
- ▶ Behandlung von Konvektionstermen (\rightarrow Petrov-Galerkin Verfahren)
- ▶ Bestimmung einer optimalen Unterteilung (\rightarrow Adaptivität)
- ▶ Lösung der diskreten Probleme (\rightarrow Effiziente Löser)

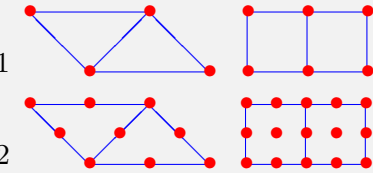
128 / 264

Elementfreiheitsgrade $\mathcal{N}_{K,k}$ 

129 / 264

Globale Freiheitsgrade $\mathcal{N}_{\mathcal{T},k}$

$$\mathcal{N}_{\mathcal{T},k} = \bigcup_{K \in \mathcal{T}} \mathcal{N}_{K,k}$$

 $k = 1$ $k = 2$ 

- ▶ Wegen der Zulässigkeit von \mathcal{T} sind die Funktionen in $S^{k,0}(\mathcal{T})$ eindeutig definiert durch ihre Werte in $\mathcal{N}_{\mathcal{T},k}$.

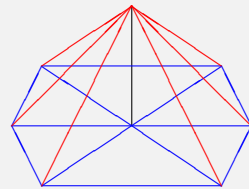
130 / 264



Nodale Basisfunktionen

Für $z \in \mathcal{N}_{\mathcal{T},k}$ ist die **nodale Basisfunktion** $\lambda_{z,k}$ eindeutig definiert durch die Bedingungen

- ▶ $\lambda_{z,k} \in S^{k,0}(\mathcal{T})$,
- ▶ $\lambda_{z,k}(z) = 1$,
- ▶ $\lambda_{z,k}(y) = 0$ für alle $y \in \mathcal{N}_{\mathcal{T},k} \setminus \{z\}$.



131 / 264




Eigenschaften

- ▶ $\{\lambda_{z,k} : z \in \mathcal{N}_{\mathcal{T},k}\}$ ist eine **Basis** für $S^{k,0}(\mathcal{T})$.
- ▶ $\{\lambda_{z,k} : z \in \mathcal{N}_{\mathcal{T},k} \setminus \Gamma\}$ ist eine **Basis** für $S_0^{k,0}(\mathcal{T})$.
(Freiheitsgrade auf dem Rand Γ werden unterdrückt.)
- ▶ $\lambda_{z,k}$ verschwindet außerhalb der Vereinigung aller der Elemente, die den Punkt z enthalten.
- ▶ Die Steifigkeitsmatrix ist **dünn besetzt**.

132 / 264





Berechnung durch Transformation auf ein Referenzelement

- ▶ Referenzelemente \hat{K} 
- ▶ Bestimme die nodalen Basisfunktionen $\hat{\lambda}_{z,k}$ zum Referenzelement \hat{K}
- ▶ Bestimme eine affine Transformation des Referenzelementes \hat{K} auf das aktuelle Element K
 $\hat{K} \ni \hat{x} \mapsto x = b_K + B_K \hat{x} \in K$
- ▶ Berechne $\lambda_{z,k}$ aus $\hat{\lambda}_{z,k}$ mit Hilfe der affinen Transformation
 $\lambda_{z,k}(x) = \hat{\lambda}_{z,k}(\hat{x})$

133 / 264



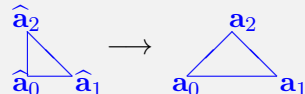
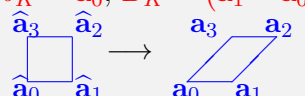
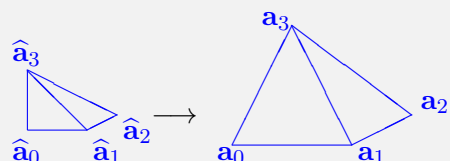
Beispiele für $\hat{\lambda}_{z,k}$

- ▶ Referenzdreieck 
 - ▶ $k = 1$
Ecken $1 - x - y, x, y$
 - ▶ $k = 2$
Ecken $(1 - x - y)(1 - 2x - 2y), x(2x - 1), y(2y - 1)$
Kantenmitten $4x(1 - x - y), 4xy, 4y(1 - x - y)$
- ▶ Referenzquadrat 
 - ▶ $k = 1$
Ecken $(1 - x)(1 - y), x(1 - y), xy, (1 - x)y$
 - ▶ $k = 2$
Ecken $(1 - 2x)(1 - x)(1 - 2y)(1 - y), x(2x - 1)(1 - 2y)(1 - y), x(2x - 1)y(2y - 1), (1 - 2x)(1 - x)y(2y - 1)$
Kantenmitten $4x(1 - x)(1 - y)(1 - 2y), 4x(2x - 1)y(1 - y), 4x(1 - x)y(2y - 1), 4y(1 - y)(1 - 2x)(1 - x)$
Schwerpunkt $16x(1 - x)y(1 - y)$

134 / 264



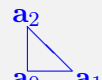
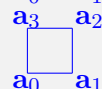
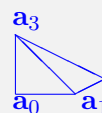
Beispiele für affine Transformationen

- ▶ 
 $b_K = \mathbf{a}_0, B_K = (\mathbf{a}_1 - \mathbf{a}_0, \mathbf{a}_2 - \mathbf{a}_0)$
- ▶ 
 $b_K = \mathbf{a}_0, B_K = (\mathbf{a}_1 - \mathbf{a}_0, \mathbf{a}_3 - \mathbf{a}_0)$
- ▶ 
 $b_K = \mathbf{a}_0, B_K = (\mathbf{a}_1 - \mathbf{a}_0, \mathbf{a}_2 - \mathbf{a}_0, \mathbf{a}_3 - \mathbf{a}_0)$
- ▶ Analoge Formeln gelten für Parallelepipede.

135 / 264



Berechnung aus Element-Geometrie ($k = 1$)

- ▶  $\lambda_{a_i,1}(x) = \frac{\det(x - \mathbf{a}_{i+1}, \mathbf{a}_{i+2} - \mathbf{a}_{i+1})}{\det(\mathbf{a}_i - \mathbf{a}_{i+1}, \mathbf{a}_{i+2} - \mathbf{a}_{i+1})}$
- ▶  $\lambda_{a_i,1}(x) = \frac{\det(x - \mathbf{a}_{i+2}, \mathbf{a}_{i+3} - \mathbf{a}_{i+2})}{\det(\mathbf{a}_i - \mathbf{a}_{i+2}, \mathbf{a}_{i+3} - \mathbf{a}_{i+2})} \cdot \frac{\det(x - \mathbf{a}_{i+2}, \mathbf{a}_{i+1} - \mathbf{a}_{i+2})}{\det(\mathbf{a}_i - \mathbf{a}_{i+2}, \mathbf{a}_{i+1} - \mathbf{a}_{i+2})}$
- ▶  $\lambda_{a_i,1}(x) = \frac{\det(x - \mathbf{a}_{i+1}, \mathbf{a}_{i+2} - \mathbf{a}_{i+1}, \mathbf{a}_{i+3} - \mathbf{a}_{i+1})}{\det(\mathbf{a}_i - \mathbf{a}_{i+1}, \mathbf{a}_{i+2} - \mathbf{a}_{i+1}, \mathbf{a}_{i+3} - \mathbf{a}_{i+1})}$
- ▶ Parallelepipede analog mit 3 Faktoren entsprechend 3 Tetraedern
- ▶ Alle Indizes sind modulo der Eckenzahl des Elementes zu nehmen.

136 / 264



Berechnung aus Element-Geometrie ($k \geq 2$)

- ▶ Die $\lambda_{z,k}$ können dargestellt werden als Produkte der Basisfunktionen erster Ordnung $\lambda_{\mathbf{a}_i,1}$ zu den Ecken.
- ▶ Beispiel: Dreieck, $k = 2$
 - ▶ Eckpunkt \mathbf{a}_i
 $\lambda_{\mathbf{a}_i,2} = \lambda_{\mathbf{a}_i} [\lambda_{\mathbf{a}_i} - \lambda_{\mathbf{a}_{i+1}} - \lambda_{\mathbf{a}_{i+2}}]$
 - ▶ Mittelpunkt z der Kante mit Endpunkten \mathbf{a}_i und \mathbf{a}_{i+1}
 $\lambda_{z,2} = 4\lambda_{\mathbf{a}_i} \lambda_{\mathbf{a}_{i+1}}$
- ▶ Beispiel: Parallelogramm, $k = 2$
 - ▶ Eckpunkt \mathbf{a}_i
 $\lambda_{\mathbf{a}_i,2} = \lambda_{\mathbf{a}_i} [\lambda_{\mathbf{a}_i} - \lambda_{\mathbf{a}_{i+1}} + \lambda_{\mathbf{a}_{i+2}} - \lambda_{\mathbf{a}_{i+3}}]$
 - ▶ Mittelpunkt z der Kante mit Endpunkten \mathbf{a}_i und \mathbf{a}_{i+1}
 $\lambda_{z,2} = 4\lambda_{\mathbf{a}_i} [\lambda_{\mathbf{a}_{i+1}} - \lambda_{\mathbf{a}_{i+2}}]$
 - ▶ Schwerpunkt z
 $\lambda_{z,2} = 16\lambda_{\mathbf{a}_0} \lambda_{\mathbf{a}_2}$

137 / 264



Aufstellen des Lastvektors

- ▶ Durchlaufe alle Elemente K :
 - ▶ Durchlaufe alle Elementfreiheitsgrade z von K :
 - ▶ Berechne $\int_K f \lambda_{z,k} dx$.
 - ▶ Addiere das Ergebnis zu dem entsprechenden Eintrag des Lastvektors.

138 / 264



Aufstellen der Steifigkeitsmatrix

- ▶ Durchlaufe alle Elemente K :
 - ▶ Durchlaufe alle Paare z, z' von Elementfreiheitsgraden von K :
 - ▶ Berechne $\int_{\Omega} [\nabla \lambda_{z',k} \cdot A \nabla \lambda_{z,k} + \alpha \lambda_{z',k} \lambda_{z,k}] dx$.
 - ▶ Addiere das Ergebnis zu dem entsprechenden Eintrag der Steifigkeitsmatrix.

139 / 264



Berechnung von Integralen

- ▶ Die exakte Berechnung der beim Aufstellen der Steifigkeitsmatrix und des Lastvektors auftretenden Integrale ist häufig nicht möglich oder zu aufwändig.
- ▶ Die Integrale werden daher durch **Quadraturformeln** näherungsweise berechnet:

$$\int_K \varphi dx \approx Q_k(\varphi) = \sum_{q \in Q_K} c_q \varphi(q).$$
- ▶ Damit der dadurch verursachte Fehler nicht den Fehler der Finite-Element-Diskretisierung dominiert, muss die Quadraturformel mindestens die **Ordnung $2k - 2$** haben:

$$\int_K \varphi dx = Q_K(\varphi) \text{ für alle } \varphi \in R_{2k-2}(K).$$
- ▶ Für lineare Elemente $S^{1,0}(\mathcal{T})$ reicht also die Ordnung 0; für quadratische Elemente $S^{2,0}(\mathcal{T})$ die Ordnung 2.

140 / 264



Beispiele für Quadraturformeln

- ▶ Dreieck:
 - ▶ Ordnung 1:
 - ▶ Q_K Schwerpunkt von K ,
 - ▶ $c_q = |K|$
 - ▶ Ordnung 2:
 - ▶ Q_K Kantenmittelpunkte von K ,
 - ▶ $c_q = \frac{1}{3}|K|$ für alle q
- ▶ Parallelogramm:
 - ▶ Ordnung 1:
 - ▶ Q_K Schwerpunkt von K ,
 - ▶ $c_q = |K|$
 - ▶ Ordnung 3:
 - ▶ Q_K Ecken, Kantenmitten und Schwerpunkt von K ,
 - ▶ $c_q = \begin{cases} \frac{1}{36}|K| & \text{falls } q \text{ Ecke} \\ \frac{4}{36}|K| & \text{falls } q \text{ Kantenmitte} \\ \frac{16}{36}|K| & \text{falls } q \text{ Schwerpunkt} \end{cases}$

141 / 264



Behandlung gekrümmter Ränder

- ▶ Gebiete mit gekrümmten Rändern können nicht durch geradlinige Elemente wie Dreiecke, Parallelogramme usw. exakt überdeckt werden.
- ▶ Mögliche Auswege:
 - ▶ Benutze in Randnähe krummlinige (z.B. **isoparametrische**) Elemente.
 - ▶ Dies macht die Implementierung erheblich aufwändiger.
 - ▶ Nur spezielle krummlinige Ränder können exakt dargestellt werden.
 - ▶ Approximiere das Gebiet Ω durch ein Polyeder $\Omega_{\mathcal{T}}$, das exakt überdeckt wird.
 - ▶ Der durch die Randapproximation erzeugte Fehler ist $\approx h_{\mathcal{T}}^2$.
 - ▶ Falls Ω **nicht konvex** ist, ist $\Omega_{\mathcal{T}} \cap \Omega \neq \emptyset$. Dann müssen die Daten f, A, α der Differentialgleichung ggf. auf das größere Gebiet **fortgesetzt** werden.

142 / 264



Neumann-Randbedingungen

- ▶ Die **Neumann-Randbedingung** $\mathbf{n} \cdot A \nabla u = g$ auf $\Gamma_N \subset \Gamma$ führt zu
 - ▶ einem zusätzlichen Term $\int_{\Gamma_N} g v dS$ auf der rechten Seite des Variationsproblems,
 - ▶ einem zusätzlichen Term $\int_{\Gamma_N} g v_{\mathcal{T}} dS$ auf der rechten Seite des diskreten Problems.
- ▶ Die zusätzlichen Beiträge zum Lastvektor werden beim Durchlaufen der Elemente berücksichtigt.
- ▶ Die **Elementfreiheitsgrade**, die auf dem Neumann-Rand Γ_N liegen, sind zusätzliche **Unbekannte**.

143 / 264



Konvektions-Diffusions-Gleichung

$$\begin{aligned} -\operatorname{div}(A \nabla u) + \mathbf{a} \cdot \nabla u + \alpha u &= f & \text{in } \Omega \\ u &= 0 & \text{auf } \Gamma \end{aligned}$$

- ▶ Ω ein **Polyeder** in \mathbb{R}^d mit $d = 2$ oder $d = 3$
- ▶ $A(x)$ eine symmetrische, positiv definite, $d \times d$ Matrix für jedes x in Ω
- ▶ $\mathbf{a}(x)$ ein **Vektor** in \mathbb{R}^d für jedes x in Ω
- ▶ $\alpha(x)$ eine nicht-negative Zahl für jedes x in Ω
- ▶ $\alpha(x) - \frac{1}{2} \operatorname{div} \mathbf{a}(x) \geq 0$ für jedes x in Ω

144 / 264



Variationsproblem

- ▶ Der **Konvektionsterm** $\mathbf{a} \cdot \nabla u$ führt auf einen zusätzlichen Term $\int_{\Omega} v \mathbf{a} \cdot \nabla u dx$ im Variationsproblem.
- ▶ Das Variationsproblem hat nach wie vor eine eindeutige Lösung.
- ▶ Die Lösung des Variationsproblems ist nicht mehr Minimum einer Energiefunktion.

145 / 264



Diskretes Problem

- ▶ Der **Konvektionsterm** $\mathbf{a} \cdot \nabla u$ führt auf einen zusätzlichen Term $\int_{\Omega} v \tau \mathbf{a} \cdot \nabla u \tau dx$ im diskreten Problem.
- ▶ Das diskrete Problem hat nach wie vor eine eindeutige Lösung.
- ▶ Die Lösung des diskreten Problems ist nicht mehr Minimum einer Energiefunktion.
- ▶ Die **Steifigkeitsmatrix ist nicht mehr symmetrisch**.

146 / 264



Problem

- ▶ Falls die **Péclet-Zahl** $\frac{|\mathbf{a}|h\tau}{\nu}$ (ν kleinster Eigenwert der Diffusionsmatrix $A(x)$) größer ist als 1, führt die übliche Finite-Element-Diskretisierung zu unphysikalischen Oszillationen der numerischen Lösung.
- ▶ Dies liegt daran, dass diese Diskretisierung einer zentralen Differenzdiskretisierung des Konvektionstermes gleicht.
- ▶ Mögliche Auswege sind:
 - ▶ **Upwind-Verfahren**,
 - ▶ **Petrov-Galerkin-Verfahren**.

147 / 264



Upwind-Verfahren

- ▶ Vorgehensweise:
 - ▶ Approximiere beim Aufstellen der Steifigkeitsmatrix die **Konvektionsterme** $\int_K \lambda_{z',k} \mathbf{a} \cdot \nabla \lambda_{z,k} dx$ durch eine **Quadraturformel** $\sum_{q \in Q_K} c_q \lambda_{z',k}(q) \mathbf{a}(q) \cdot \nabla \lambda_{z,k}(q)$.
 - ▶ Ersetze $\mathbf{a}(q) \cdot \nabla \lambda_{z,k}(q)$ durch einen **upwind-Differenzenquotienten** wie bei Differenzenverfahren.
- ▶ Nachteile:
 - ▶ Der Ansatz führt häufig zu einem Genauigkeitsverlust.
 - ▶ Wegen der nötigen Interpolation zwischen Elementfreiheitsgraden treten häufig neue unphysikalischen Oszillationen auf.
 - ▶ Bei nichtlinearen Differentialgleichungen hängt das resultierende diskrete Problem nicht mehr differenzierbar von der Lösung ab.

148 / 264



Petrov-Galerkin-Verfahren

- ▶ Idee:
Teste die Differentialgleichung zusätzlich **elementweise** mit $\mathbf{a} \cdot \nabla v_{\mathcal{T}}$.
- ▶ Diskretes Problem:
Finde $u_{\mathcal{T}} \in S_0^{k,0}(\mathcal{T})$ (**Ansatzfunktion**) so, dass für alle $v_{\mathcal{T}} \in S_0^{k,0}(\mathcal{T})$ (**Testfunktion**) gilt

$$\int_{\Omega} [\nabla v_{\mathcal{T}} \cdot A \nabla u_{\mathcal{T}} + v_{\mathcal{T}} \mathbf{a} \cdot \nabla u_{\mathcal{T}} + \alpha u_{\mathcal{T}} v_{\mathcal{T}}] dx$$

$$+ \sum_{K \in \mathcal{T}} \delta_K h_K \int_K [-\operatorname{div}(A \nabla u_{\mathcal{T}}) + \mathbf{a} \cdot \nabla u_{\mathcal{T}} + \alpha u_{\mathcal{T}}] \mathbf{a} \cdot \nabla v_{\mathcal{T}} dx$$

$$= \int_{\Omega} f v_{\mathcal{T}} dx + \sum_{K \in \mathcal{T}} \delta_K h_K \int_K f \mathbf{a} \cdot \nabla v_{\mathcal{T}} dx$$
 (h_K Durchmesser von K , $\delta_K > 0$ Stabilitätsparameter)

149 / 264



Eigenschaften

- ▶ Das diskrete Problem besitzt eine eindeutige Lösung.
- ▶ Unphysikalische Oszillationen werden weitest gehend vermieden.
- ▶ Es gelten die gleichen Fehlerabschätzungen wie bei der üblichen Diskretisierung (**kein Genauigkeitsverlust**).
- ▶ **Zusätzlich** gilt eine analoge Fehlerabschätzung für den Fehler $\|\mathbf{a} \cdot \nabla u - \mathbf{a} \cdot \nabla u_{\mathcal{T}}\|$ der **Konvektionsableitung**.
- ▶ Bei nichtlinearen Differentialgleichungen hängt das diskrete Problem differenzierbar von der Lösung ab.
- ▶ Eine beliebige Wahl für die Stabilitätsparameter ist

$$\delta_K = \frac{|\mathbf{a}| h_K}{\sqrt{\nu^2 + |\mathbf{a}|^2 h_K^2}}$$

150 / 264



Motivation

- ▶ Die üblichen Finite-Element-Verfahren liefern eine Approximation für die Lösung u (**Verschiebung**) der Differentialgleichung.
- ▶ Häufig sind abgeleitete Größen wie ∇u (**Spannungen**) physikalisch interessanter.
- ▶ Diese müssen nachträglich durch numerische Differentiation bestimmt werden und werden dadurch weniger genau approximiert.
- ▶ Bei Elastizitätsproblemen führt der klassische Ansatz (**Verschiebungsmethode**) zu unphysikalischen Lösungen (**locking**).
- ▶ Gesucht sind Variationsformulierungen und Finite-Element-Diskretisierungen, die Verschiebungen und Spannungen gleichzeitig mit gleicher Genauigkeit approximieren.

151 / 264



Poisson-Gleichung als System erster Ordnung

- ▶ Poisson-Gleichung

$$\begin{aligned} -\Delta u &= f & \text{in } \Omega \\ u &= 0 & \text{auf } \Gamma \end{aligned}$$

- ▶ Führe ∇u als zusätzliche Variable ein.
- ▶ **Resultierendes System erster Ordnung**

$$\begin{aligned} \sigma - \nabla u &= 0 & \text{in } \Omega \\ -\operatorname{div} \sigma &= f & \text{in } \Omega \\ u &= 0 & \text{auf } \Gamma \end{aligned}$$

152 / 264



Idee der Variationsformulierung

- ▶ Multipliziere die erste Gleichung mit einem differenzierbaren **Vektorfeld** $\tau : \Omega \rightarrow \mathbb{R}^d$ und integriere das Ergebnis über Ω

$$\int_{\Omega} \sigma \cdot \tau dx - \int_{\Omega} \nabla u \cdot \tau dx = 0.$$

- ▶ Integriere den ∇u -Term partiell

$$\int_{\Omega} \nabla u \cdot \tau dx = - \int_{\Omega} u \operatorname{div} \tau dx.$$

- ▶ Multipliziere die zweite Gleichung mit einer Funktion $v : \Omega \rightarrow \mathbb{R}$ und integriere das Ergebnis über Ω

$$- \int_{\Omega} \operatorname{div} \sigma v dx = \int_{\Omega} f v dx.$$

153 / 264



Variationsproblem

- ▶ $H(\operatorname{div}, \Omega) = \{\sigma \in L^2(\Omega)^d : \operatorname{div} \sigma \in L^2(\Omega)\}$
- ▶ Finde $u \in L^2(\Omega)$ und $\sigma \in H(\operatorname{div}, \Omega)$, so dass für alle $v \in L^2(\Omega)$ und alle $\tau \in H(\operatorname{div}, \Omega)$ gilt

$$\begin{aligned} \int_{\Omega} \sigma \cdot \tau dx + \int_{\Omega} u \operatorname{div} \tau dx &= 0 \\ - \int_{\Omega} v \operatorname{div} \sigma dx &= \int_{\Omega} f v dx \end{aligned}$$

154 / 264



Eigenschaften

- ▶ Das Variationsproblem hat eine eindeutige Lösung (σ, u) .
- ▶ Die **primale Variable** σ ist das eindeutige Minimum der **Energiefunktion** $\frac{1}{2} \int_{\Omega} \sigma : \sigma dx$ unter der **Nebenbedingung** (**Gleichgewichtsbedingung**) $-\operatorname{div} \sigma = f$; die **duale Variable** u ist der zugehörige **Lagrange-Multiplikator**.
- ▶ (σ, u) ist der eindeutige **Sattelpunkt** der Funktion $\mathcal{L}(\tau, v) = \frac{1}{2} \int_{\Omega} \tau : \tau dx + \int_{\Omega} v \operatorname{div} \tau dx$; σ minimiert \mathcal{L} bei festem v , u maximiert \mathcal{L} bei festem τ .

155 / 264



Gemischte Finite-Element-Diskretisierung

- ▶ Bestimme eine zulässige Unterteilung \mathcal{T} von Ω .
- ▶ Wähle endlich dimensionale Teilräume $X(\mathcal{T})$ von $H(\operatorname{div}, \Omega)$ und $Y(\mathcal{T})$ von $L^2(\Omega)$.
- ▶ Finde $u_{\mathcal{T}} \in Y(\mathcal{T})$ und $\sigma_{\mathcal{T}} \in X(\mathcal{T})$, so dass für alle $v_{\mathcal{T}} \in Y(\mathcal{T})$ und alle $\tau_{\mathcal{T}} \in X(\mathcal{T})$ gilt

$$\begin{aligned} \int_{\Omega} \sigma_{\mathcal{T}} \cdot \tau_{\mathcal{T}} dx + \int_{\Omega} u_{\mathcal{T}} \operatorname{div} \tau_{\mathcal{T}} dx &= 0 \\ - \int_{\Omega} v_{\mathcal{T}} \operatorname{div} \sigma_{\mathcal{T}} dx &= \int_{\Omega} f v_{\mathcal{T}} dx \end{aligned}$$

156 / 264



Eigenschaften

- ▶ Nach Wahl von Basen für $X(\mathcal{T})$ und $Y(\mathcal{T})$ führt die gemischte Finite-Element-Diskretisierung auf ein lineares Gleichungssystem, Unbekannte sind die Koeffizienten von $\sigma_{\mathcal{T}}$ und $u_{\mathcal{T}}$.
- ▶ Die Steifigkeitsmatrix hat die Form $\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}$ mit einer symmetrischen, positiv definiten Matrix A und einer rechteckigen Matrix B .
- ▶ Die Steifigkeitsmatrix ist indefinit und hat positive und negative Eigenwerte.
- ▶ Die Steifigkeitsmatrix ist genau dann invertierbar, wenn $B^T A^{-1} B$ invertierbar ist.

157 / 264



Schwierigkeiten

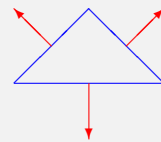
- ▶ Um optimale Fehlerabschätzungen zu erhalten, muss zusätzlich der kleinste Eigenwert der Matrix $B^T A^{-1} B$ unabhängig von \mathcal{T} von Null weg beschränkt sein (inf-sup-Bedingung).
- ▶ Nicht jede auf den ersten Blick plausible Wahl von $X(\mathcal{T})$ und $Y(\mathcal{T})$ erfüllt diese Bedingung.

158 / 264



Raviart-Thomas-Diskretisierung

- ▶ \mathcal{T} besteht aus Dreiecken oder Tetraedern.
- ▶ $RT(K) = \{\mathbf{a} + b\mathbf{x} : \mathbf{a} \in \mathbb{R}^d, b \in \mathbb{R}\}$ (Raviart-Thomas Element)
- ▶ Freiheitsgrade sind die Normalkomponenten in den Kantenmittelpunkten für $d = 2$ und in den Schwerpunkten der Seitenflächen für $d = 3$.
- ▶ $Y(\mathcal{T}) = S^{0,-1}(\mathcal{T})$ (stückweise konstante Verschiebungen)
- ▶ $X(\mathcal{T}) = \{\tau : \tau|_K \in RT(K) \text{ für alle } K \in \mathcal{T}, \tau \cdot \mathbf{n} \text{ ist stetig in den Kanten- bzw. Flächenmitten}\}$



159 / 264



Eigenschaften

- ▶ Die Raviart-Thomas Diskretisierung erfüllt die inf-sup Bedingung.
- ▶ Es gelten die Fehlerabschätzungen $\|\sigma - \sigma_{\mathcal{T}}\| + \|\text{div } \sigma - \text{div } \sigma_{\mathcal{T}}\| + \|u - u_{\mathcal{T}}\| \leq c h_{\mathcal{T}}$.
- ▶ Die Konstante c hängt von Ableitungen erster Ordnung von f , u und σ ab.
- ▶ Es gibt analoge Diskretisierungen höherer Ordnung und solche für Unterteilungen in Parallelogramme und Parallelepipede.

160 / 264



Lineare Elastizitätstheorie

- ▶ Bei Anwendung auf die Gleichungen der linearen Elastizitätstheorie entspricht u dem **Verschiebungsvektor** und σ dem **Spannungstensor**.
- ▶ Eine zusätzliche Schwierigkeit bereitet die **Symmetrie** des Spannungstensors.
 - ▶ Für das Variationsproblem ist sie automatisch erfüllt.
 - ▶ Wird sie für das diskrete Problem exakt gefordert, treten Stabilitätsprobleme auf.
 - ▶ Daher kann die Symmetrie für das diskrete Problem nur in abgeschwächter Form gefordert werden (**PEERS: plane elasticity elements with reduced symmetry**).

161 / 264



Strömungsmechanik

- ▶ In der Strömungsmechanik treten vorwiegend Variationsprobleme mit **Divergenz-Nebenbedingungen** auf, d.h. die Divergenz eines Vektorfeldes wie der Geschwindigkeit muss verschwinden (**Inkompressibilität**).
- ▶ Diese Nebenbedingung macht die Nutzung gemischter Finite-Element-Methoden zwingend erforderlich.
- ▶ Bei den resultierenden Variationsproblemen sind die Rollen von u und σ vertauscht; u ist die primale Variable, σ ist der Lagrange-Multiplikator zur Nebenbedingung.
- ▶ u entspricht der **Geschwindigkeit**, σ dem **Druck**.

162 / 264



A posteriori Fehlerschätzung und Adaptivität

- ▶ Motivation
- ▶ A posteriori Fehlerschätzer
- ▶ Gitteranpassung

163 / 264



Grenzen von a priori Fehlerabschätzungen

- ▶ Sie liefern nur eine **asymptotische** Aussage über den Fehler, d.h. über das Verhalten für immer feiner werdende Unterteilungen.
- ▶ Sie geben keine Auskunft über die tatsächliche Größe des Fehlers.
- ▶ Sie erlauben keine Rückschlüsse über die räumliche (und zeitliche) Verteilung des Fehlers.

164 / 264



Wunschliste

- ▶ Eine leicht berechenbare und zuverlässige Information über die tatsächliche Größe des Fehlers und seine räumliche (und zeitliche) Verteilung.
- ▶ Erhalte eine Näherungslösung für die gegebene Differentialgleichung mit vorgegebener Toleranz und (nahezu) minimalem Aufwand (z.B. gemessen in der Zahl der Freiheitsgrade).

165 / 264



Adaptiver Algorithmus

0. Gegeben: Daten einer partiellen Differentialgleichung und eine **Toleranz ε** .
Gesucht: Eine numerische Näherungslösung mit einem Fehler kleiner oder gleich der Toleranz.
1. Konstruiere eine erste grobe Unterteilung \mathcal{T}_0 , setze $k = 0$.
2. Stelle das diskrete Problem zu \mathcal{T}_k auf und löse es (näherungsweise).
3. Bestimme für jedes Element K in \mathcal{T}_k einen **Fehlerschätzer**.
4. Falls der geschätzte Gesamtfehler kleiner oder gleich der Toleranz ε ist, **stopp**.
5. Entscheide welche Elemente von \mathcal{T}_k unterteilt werden müssen und konstruiere eine entsprechende neue Unterteilung \mathcal{T}_{k+1} . Erhöhe k um 1 und gehe zu Schritt **2** zurück.

166 / 264



Wesentliche Ingredienzien

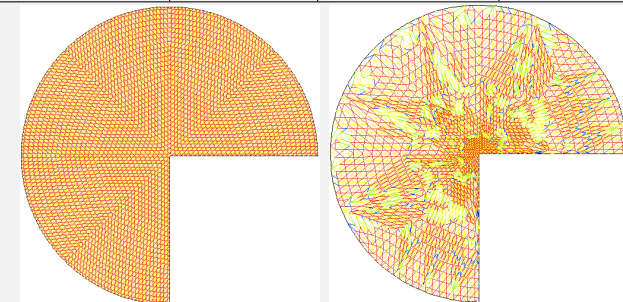
- ▶ Diskretisierungsmethode
- ▶ Lösungsverfahren für die diskreten Probleme (→ **Effiziente Löser**)
- ▶ Fehlerschätzer (→ **A posteriori Fehlerschätzer**)
- ▶ Gitteranpassungsstrategie (→ **Gitteranpassung**)

167 / 264



Beispiel: Poisson-Gleichung mit singulärer Lösung

Verfeinerung	Elemente	Unbekannte	rel. Fehler
gleichmäßig	24576	12033	0.5%
adaptiv	11242	5529	0.5%

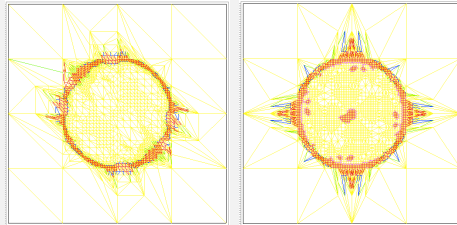


168 / 264



Beispiel: Reaktions-Diffusions-Gleichung mit innerer Grenzschicht

	Dreiecke		Vierecke	
	gleichm.	adaptiv	gleichm.	adaptiv
Unbekannte	16129	2923	16129	4722
Dreiecke	32768	5860	0	3830
Vierecke	0	0	16384	2814
rel. Fehler	3.8%	3.5%	6.1%	4.4%



169 / 264



Differentialgleichung und Diskretisierung

- ▶ Lineare elliptische Differentialgleichung zweiter Ordnung

$$-\operatorname{div}(A\nabla u) + \mathbf{a} \cdot \nabla u + \alpha u = f \quad \text{in } \Omega$$

$$u = 0 \quad \text{auf } \Gamma$$

- ▶ u Lösung des Variationsproblems
- ▶ Übliche Finite-Element-Diskretisierung zu Unterteilung \mathcal{T}
- ▶ $u_{\mathcal{T}}$ berechnete Lösung des diskreten Problems

170 / 264



Residueller Fehlerschätzer

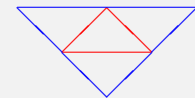
- ▶ $\eta_K = \left\{ h_K^2 \int_K |f + \operatorname{div}(A\nabla u_{\mathcal{T}}) - \mathbf{a} \cdot \nabla u_{\mathcal{T}} - \alpha u_{\mathcal{T}}|^2 dx + \sum_{E \in \mathcal{E}_K \setminus \Gamma} h_E \int_E |[\mathbf{n}_E \cdot A\nabla u_{\mathcal{T}}]_E|^2 dS \right\}^{\frac{1}{2}}$
- ▶ \mathcal{E}_K Kanten ($d = 2$) bzw. Seitenflächen ($d = 3$) von K
- ▶ h_K, h_E Durchmesser von K bzw. E
- ▶ \mathbf{n}_E ein Einheitsvektor senkrecht zu E
- ▶ $[\varphi]_E$ Sprung einer Funktion φ über E

171 / 264



A posteriori Fehlerabschätzungen

- ▶ $\left\{ \int_{\Omega} |\nabla u - \nabla u_{\mathcal{T}}|^2 dx \right\}^{\frac{1}{2}} \leq c^* \left\{ \sum_{K \in \mathcal{T}} \eta_K^2 \right\}^{\frac{1}{2}}$
- ▶ $\eta_K \leq c_* \left\{ \int_{\omega_K} |\nabla u - \nabla u_{\mathcal{T}}|^2 dx \right\}^{\frac{1}{2}}$
- ▶ ω_K Vereinigung aller Elemente, die mit K eine Kante ($d = 2$) bzw. Seitenfläche ($d = 3$) gemeinsam haben
- ▶ c^*, c_* hängen ab von
 - ▶ der relativen Größe von Diffusion A , Konvektion \mathbf{a} und Reaktion α zueinander
 - ▶ dem Polynomgrad von $u_{\mathcal{T}}$
 - ▶ dem Formparameter von \mathcal{T} (maximales Verhältnis von Elementdurchmesser zum Durchmesser des größten eingeschriebenen Balles)



172 / 264



Struktur des Fehlerschätzers

- ▶ $f + \operatorname{div}(A\nabla u_{\mathcal{T}}) - \mathbf{a} \cdot \nabla u_{\mathcal{T}} - \alpha u_{\mathcal{T}}$ (Elementresiduum) ist das Residuum auf K von $u_{\mathcal{T}}$ bzgl. der Differentialgleichung
- ▶ $[\mathbf{n}_E \cdot A\nabla u_{\mathcal{T}}]_E$ (Kantenresiduum) ist der Sprung über E des Operators, der starke und schwache Form der Differentialgleichung verbindet (Randterm bei partieller Integration)
- ▶ Für Elemente niedriger Ordnung ($k = 1$) ist das Kantenresiduum häufig entscheidend.

173 / 264



Struktur der a posteriori Fehlerabschätzungen

- ▶ Die obere Schranke ist **global**.
- ▶ Dies liegt daran, dass sie auf Eigenschaften des Lösungsoperators der Differentialgleichung beruht. (Lokale Last impliziert globale Verschiebung.)
- ▶ Die untere Schranke ist **lokal**.
- ▶ Dies liegt daran, dass sie auf Eigenschaften des Differentialoperators beruht. (Lokale Verschiebung impliziert lokale Last.)

174 / 264



Einige alternative Fehlerschätzer

- ▶ **Lokale Hilfsprobleme:** Löse lokale diskrete Hilfsprobleme höherer Ordnung mit Element- und Kantenresiduen als Lasttermen.
- ▶ **Hierarchische Schätzer:** Vergleiche die aktuelle Lösung mit einer Finite-Element-Lösung höherer Ordnung basierend auf einem **Lumping** der Steifigkeitsmatrix.
- ▶ **ZZ-Schätzer:** Vergleiche den Gradienten der Lösung mit einer Mittelung des Gradienten.
- ▶ Alle Schätzer sind **äquivalent** in dem Sinne, dass sie bis auf multiplikative Konstanten gegeneinander abgeschätzt werden können.

175 / 264



Überblick

- ▶ Die Gitterverfeinerung beruht auf zwei Säulen:
 - ▶ **Markierungsstrategien**, die festlegen, welche Elemente unterteilt werden sollen,
 - ▶ **Verfeinerungsregeln**, die festlegen, wie ein einzelnes Element unterteilt werden soll.
- ▶ Um die Zulässigkeit der Unterteilung zu gewährleisten und **hängende Knoten** zu vermeiden, erfolgt die Verfeinerung in zwei Etappen:
 - ▶ Zuerst werden alle Elemente unterteilt, die einen zu großen Wert des Fehlerschätzers η_K aufweisen (**reguläre Unterteilung**).
 - ▶ Danach werden zusätzliche Elemente unterteilt, um hängende Knoten zu beseitigen, die während der ersten Etappe erzeugt wurden (**irreguläre Unterteilung**).
- ▶ Die Gitterverfeinerung kann mit einer **Gittervergrößerung** und **Gitterglättung** kombiniert werden.

176 / 264



Markierung mit der Maximum-Strategie

0. Gegeben: Unterteilung \mathcal{T} , Fehlerschätzer η_K für die Elemente $K \in \mathcal{T}$, Schwellenwert $\theta \in (0, 1)$.
Gesucht: Teilmenge $\tilde{\mathcal{T}}$ von **markierten** Elementen, die unterteilt werden sollen.
1. Berechne $\eta_{\mathcal{T}, \max} = \max_{K \in \mathcal{T}} \eta_K$.
 2. Falls $\eta_K \geq \theta \eta_{\mathcal{T}, \max}$ ist, **markiere** K und füge es zu $\tilde{\mathcal{T}}$ hinzu.

177 / 264



Modifikation

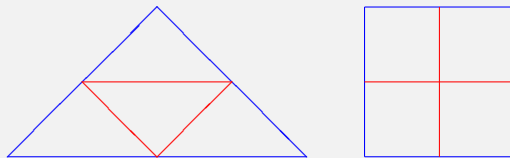
- ▶ Manchmal verteilen sich die Elemente auf drei Gruppen:
 - ▶ eine sehr kleine Zahl von Elementen mit einem sehr großen geschätzten Fehler,
 - ▶ eine sehr große Zahl von Elementen mit einem sehr kleinen geschätzten Fehler,
 - ▶ eine mittlere Zahl von Elementen mit einem geschätzten Fehler mittlerer Größe.
- ▶ Dann markiert die Maximum-Strategie nur die Elemente der ersten Gruppe.
- ▶ Dies verschlechtert die Effizienz des adaptiven Algorithmus.
- ▶ Dies kann durch folgende Modifikation vermieden werden:
Markiere zuerst einen kleinen Prozentsatz ε der Elemente mit dem größten geschätzten Fehler und wende danach die Maximum-Strategie auf die verbleibenden Elemente an.

178 / 264



Reguläre Unterteilung

- ▶ Verbinde die Kantenmittelpunkte der Elemente.



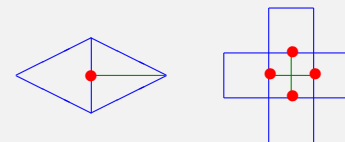
- ▶ Dies erhält den Formparameter (Verhältnis des Elementdurchmessers zum Durchmesser des größten eingeschriebenen Balles).

179 / 264



Hängende Knoten

- ▶ **Hängende Knoten** zerstören die Zulässigkeit der Unterteilung.



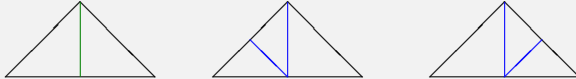
- ▶ Daher muss man
 - ▶ entweder die Stetigkeit der Finite-Element-Funktionen an den hängenden Knoten erzwingen (**hängende Knoten sind keine Freiheitsgrade**)
 - ▶ oder eine zusätzliche irreguläre Unterteilung durchführen.
- ▶ Das Erzwingen der Stetigkeit in den hängenden Knoten kann der Verfeinerung **zuwiderlaufen**.

180 / 264

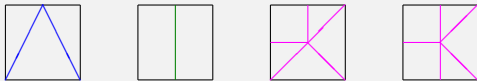


Irreguläre Unterteilung

▶ Dreiecke



▶ Vierecke



181 / 264



Bisektion markierter Kanten

- ▶ Konstruiere die erste Unterteilung so, dass die längste Kante eines jeden Elementes auch die längste Kante des angrenzenden Elementes ist.
- ▶ Markiere die längsten Kanten in der ersten Unterteilung.
- ▶ Unterteile ein Element durch Verbinden des Mittelpunktes seiner markierten Kante mit dem gegenüberliegenden Eckpunkt (**Bisektion**).
- ▶ Bei Unterteilung einer Kante eines Elementes werden die anderen Kanten die markierten Kanten der resultierenden neuen Elemente.



182 / 264



Gittervergrößerung

- ▶ Eine Vergrößerung von Unterteilungen ist erforderlich,
 - ▶ um die Optimalität des adaptiven Algorithmus zu gewährleisten, d.h. um eine gegebene Toleranz mit einer minimalen Anzahl an Freiheitsgraden zu erreichen,
 - ▶ um zeitlich veränderliche Singularitäten zu erfassen.
- ▶ Die Gittervergrößerung erfolgt durch das Zusammenfassen von benachbarten Elementen mit zu kleinem Fehler.

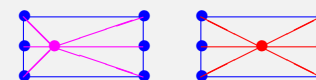
183 / 264



Gitterglättung

- ▶ Verbessere die **Qualität** einer Unterteilung \mathcal{T} durch **Verschieben** der Elementeckpunkte unter Beibehaltung der Nachbarschaftsbeziehungen.
- ▶ Die Qualität wird durch eine **Qualitätsfunktion** q gemessen, wobei ein größerer Wert von q einer besseren Qualität entspricht.
- ▶ Die Qualität wird durch einen Gauß-Seidel artigen **Glättungsprozess** verbessert:

Durchlaufe alle Elementeckpunkte z in \mathcal{T} , fixiere alle durch eine Kante mit z verbundenen Elementeckpunkte und suche einen neuen Punkt \tilde{z} mit $\min_{z \in \tilde{K}} q(\tilde{K}) > \min_{z \in K} q(K)$.



184 / 264



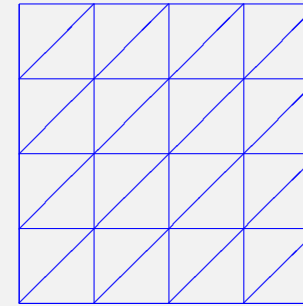
Effiziente Löser

- ▶ Motivation
- ▶ Geschachtelte Iteration
- ▶ Klassische iterative Verfahren
- ▶ CG-Verfahren
- ▶ Mehrgitterverfahren
- ▶ Vergleichsverfahren
- ▶ Unsymmetrische, indefinite und nichtlineare Probleme

185 / 264



Ein typisches Beispiel



- ▶ **Poisson-Gleichung**
 $-\Delta u = f$ in Ω , $u = 0$ auf Γ
- ▶ $\Omega = (0, 1)^2$ Einheitsquadrat
- ▶ **Courant-Triangulation**
bestehend aus $2n^2$
rechtwinklig
gleichschenkligen Dreiecken
mit Katheten der Länge
 $h = n^{-1}$
- ▶ Lineare Finite-Elemente
- ▶ Zahl N der Unbekannten
ist $\approx n^2 = h^{-2}$.

186 / 264



Eigenschaften der Steifigkeitsmatrix

- ▶ Sie ist symmetrisch, positiv definit.
- ▶ Sie hat höchstens 5 von Null verschiedene Elemente pro Zeile.
- ▶ Sie hat Bandstruktur mit einer Bandbreite $\approx h^{-1} \approx N^{\frac{1}{2}}$.
- ▶ Das Gaußsche Eliminationsverfahren erfordert $\approx N^2$ Operationen.
- ▶ Eine Matrix-Vektor-Multiplikation erfordert $\approx 5N$ Operationen.
- ▶ Der kleinste Eigenwert ist ≈ 1 .
- ▶ Der größte Eigenwert ist $\approx h^{-2} \approx N$.

187 / 264



Eigenschaften direkter Löser

- ▶ Für ein diskretes Problem mit N Unbekannten in d Raumdimensionen benötigen sie $\approx N^{2-\frac{1}{d}}$ Speicherplätze.
- ▶ Sie benötigen $\approx N^{3-\frac{2}{d}}$ arithmetische Operationen.
- ▶ Bis auf Rundungsfehler liefern sie die exakte Lösung des diskreten Problems.
- ▶ Sie liefern eine Approximation für die Lösung der Differentialgleichung mit einem Fehler $\approx h^\alpha \approx N^{-\frac{\alpha}{d}}$ (typischerweise: $\alpha \in \{1, 2\}$).

188 / 264



Eigenschaften klassischer iterativer Löser

- ▶ Sie benötigen $\approx N$ Speicherplätze.
- ▶ Sie benötigen $\approx N$ arithmetische Operationen pro Iteration.
- ▶ Ihre Konvergenzrate (Faktor um den der Fehler pro Iteration reduziert wird) verschlechtert sich mit wachsender Konditionszahl des diskreten Problems, die typischerweise $\approx h^{-2} \approx N^{\frac{2}{d}}$ ist.
- ▶ Um einen Anfangsfehler um den Faktor 0.1 zu reduzieren benötigt man typischerweise folgende Anzahl arithmetischer Operationen:
 - ▶ $\approx N^{1+\frac{2}{d}}$ für den Gauß-Seidel Algorithmus,
 - ▶ $\approx N^{1+\frac{1}{d}}$ für das Konjugierte-Gradienten- (CG-) Verfahren,
 - ▶ $\approx N^{1+\frac{1}{2d}}$ für das CG-Verfahren mit SSOR-Vorkonditionierung.

189 / 264



Schlussfolgerungen

- ▶ Direkte Verfahren erfordern zu viel Speicherplatz und Rechenzeit.
- ▶ Es ist vollkommen ausreichend, eine Näherungslösung des diskreten Problems zu bestimmen, die verglichen mit der exakten Lösung der Differentialgleichung eine ähnliche Genauigkeit hat wie die exakte Lösung des diskreten Problems.
- ▶ Iterative Löser sind überlegen, wenn es gelingt, ihre Konvergenzrate zu verbessern und eine gute Startnäherung zu finden.

190 / 264



Geschachtelte Gitter

- ▶ Häufig muss man eine Folge diskreter Probleme $L_k u_k = f_k$ lösen, die zunehmend genaueren Diskretisierungen entsprechen.
- ▶ In der Regel kennt man einen Interpolationsoperator $I_{k-1,k}$, der Funktionen der $(k-1)$ -ten Diskretisierung in solche der k -ten Diskretisierung abbildet.
- ▶ Die Interpolierende einer vernünftigen Näherungslösung des $(k-1)$ -ten diskreten Problem ist dann ein guter Startwert für jeden iterativen Löser für das k -te diskrete Problem.
- ▶ Häufig muss der Anfangsfehler nur um einen Faktor 0.1 reduziert werden.

191 / 264



Geschachtelte Iteration

- ▶ Berechne

$$\tilde{u}_0 = u_0 = L_0^{-1} f_0.$$

- ▶ Für $k = 1, 2, \dots$ berechne sukzessive eine Näherungslösung \tilde{u}_k für $u_k = L_k^{-1} f_k$ durch Anwenden von m_k Iterationen eines iterativen Löser auf das Problem

$$L_k u_k = f_k$$

mit Startnäherung $I_{k-1,k} \tilde{u}_{k-1}$.

- ▶ m_k wird dabei implizit bestimmt durch das Abbruchkriterium

$$\|f_k - L_k \tilde{u}_k\| \leq \varepsilon \|f_k - L_k (I_{k-1,k} \tilde{u}_{k-1})\|.$$

192 / 264



Problemstellung

- ▶ Zu lösen ist ein lineares Gleichungssystem $Lu = f$ mit N Unbekannten.
- ▶ L ist symmetrisch, positiv definit.
- ▶ κ bezeichnet die **Konditionszahl** von L , d.h. das Verhältnis des größten Eigenwertes von L zum kleinsten Eigenwert.
- ▶ Es ist $\kappa \approx N^{\frac{2}{d}}$.

193 / 264



Richardson-Iteration

- ▶ **Iterationsschritt:** $u \mapsto u + \frac{1}{\omega}(f - Lu)$
- ▶ ω heißt **Dämpfungsparameter**.
- ▶ ω muss die gleiche Größenordnung haben wie der größte Eigenwert von L .
- ▶ Die **Konvergenzrate** ist $\frac{\kappa-1}{\kappa+1} \approx 1 - N^{-\frac{2}{d}}$.

194 / 264



Jacobi-Iteration

- ▶ **Iterationsschritt:** $u \mapsto u + D^{-1}(f - Lu)$
- ▶ D ist die Diagonale von L .
- ▶ Die **Konvergenzrate** ist $\frac{\kappa-1}{\kappa+1} \approx 1 - N^{-\frac{2}{d}}$.
- ▶ In jedem Iterationsschritt werden sukzessive alle Gleichungen durchlaufen, und die i -te Gleichung wird exakt nach der i -ten Unbekannten aufgelöst, **ohne dabei nachfolgende Gleichungen entsprechend anzupassen**.

195 / 264



Gauß-Seidel-Iteration

- ▶ **Iterationsschritt:** $u \mapsto u + \mathcal{L}^{-1}(Lu - f)$
- ▶ \mathcal{L} ist der Teil von L unterhalb der Diagonalen einschließlich der Diagonalen
- ▶ In jedem Iterationsschritt werden sukzessive alle Gleichungen durchlaufen, die i -te Gleichung exakt nach der i -ten Unbekannten aufgelöst **und das Ergebnis unmittelbar in alle nachfolgenden Gleichungen eingesetzt**.
- ▶ Die **Konvergenzrate** ist $\frac{\kappa-1}{\kappa+1} \approx 1 - N^{-\frac{2}{d}}$.

196 / 264



Idee des CG-Verfahrens

- ▶ Die Lösung des **symmetrischen, positiv definiten** Gleichungssystems $Lu = f$ ist das eindeutige **Minimum** der quadratischen Funktion $J(u) = \frac{1}{2}u \cdot (Lu) - f \cdot u$.
- ▶ Der negative Gradient $-\nabla J(v) = f - Lv$ von J an der Stelle v ist die Richtung des **steilsten Abstieges**.
- ▶ J nimmt auf der Geraden $t \mapsto v + td$ sein eindeutiges Minimum an der Stelle $t^* = \frac{(f-Lv) \cdot d}{d \cdot (Ld)}$ an.
- ▶ Bei sukzessiver Minimierung in Richtung der negativen Gradienten verlangsamt sich das Verfahren zunehmend, da die Suchrichtungen nahezu parallel werden.
- ▶ Das Verfahren wird durch Wahl **L -orthogonaler** Suchrichtungen ($d_i \cdot (Ld_j) = 0$ für $i \neq j$) beschleunigt.
- ▶ L -orthogonale Suchrichtungen können während des Verfahrens rekursiv mit bestimmt werden.

197 / 264



Das CG-Verfahren

0. Gegeben: eine Startnäherung u_0 und eine Toleranz $\varepsilon > 0$.
1. Berechne $r_0 = f - Lu_0$, $d_0 = r_0$, $\gamma_0 = r_0 \cdot r_0$. Setze $i = 0$.
2. Falls $\gamma_i < \varepsilon^2$ ist, gebe u_i als Näherungslösung aus; **stopp**. Andernfalls gehe zu Schritt 3.
3. Berechne $s_i = Ld_i$, $\alpha_i = \frac{\gamma_i}{d_i \cdot s_i}$, $u_{i+1} = u_i + \alpha_i d_i$,
 $r_{i+1} = r_i - \alpha_i s_i$, $\gamma_{i+1} = r_{i+1} \cdot r_{i+1}$, $\beta_i = \frac{\gamma_{i+1}}{\gamma_i}$,
 $d_{i+1} = r_{i+1} + \beta_i d_i$. Erhöhe i um 1 und gehe zu Schritt 2.

198 / 264



Eigenschaften

- ▶ Das CG-Verfahren benötigt nur Skalarprodukte und Matrix-Vektor-Multiplikationen.
- ▶ Die **Konvergenzrate** ist $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \approx 1 - N^{-\frac{1}{d}}$.
- ▶ Das CG-Verfahren kann nur auf **symmetrische, positiv definite** Gleichungssysteme angewandt werden und **bricht** für **unsymmetrische oder indefinite** Systeme in der Regel **zusammen**.

199 / 264



Idee der Vorkonditionierung (preconditioning)

- ▶ Löse statt des ursprünglichen Systems $Lu = f$ das äquivalente System $\widehat{L}\widehat{u} = \widehat{f}$ mit $\widehat{L} = H^{-1}LH^{-t}$, $\widehat{f} = H^{-1}f$, $\widehat{u} = H^t u$ und einer invertierbaren Matrix H .
- ▶ Wähle H so, dass:
 - ▶ die Konditionszahl von \widehat{L} wesentlich kleiner ist als diejenige von L ,
 - ▶ Gleichungssysteme der Form $Cv = d$ mit $C = HH^t$ wesentlich leichter zu lösen sind als das ursprüngliche Problem $Lu = f$.
- ▶ Wende das CG-Verfahren auf das neue Gleichungssystem $\widehat{L}\widehat{u} = \widehat{f}$ an und drücke alle Größen durch die ursprünglichen Daten L , f und u aus.

200 / 264



Das PCG-Verfahren

0. Gegeben: eine Startnäherung u_0 und eine Toleranz $\varepsilon > 0$.
1. Berechne $r_0 = f - Lu_0$, löse $Cz_0 = r_0$ und berechne $d_0 = z_0$, $\gamma_0 = r_0 \cdot z_0$. Setze $i = 0$.
2. Falls $\gamma_i < \varepsilon^2$ ist, gebe u_i als Näherungslösung aus; **stopp**. Andernfalls gehe zu Schritt 3.
3. Berechne $s_i = Ld_i$, $\alpha_i = \frac{\gamma_i}{d_i \cdot s_i}$, $u_{i+1} = u_i + \alpha_i d_i$, $r_{i+1} = r_i - \alpha_i s_i$, löse $Cz_{i+1} = r_{i+1}$ und berechne $\gamma_{i+1} = r_{i+1} \cdot z_{i+1}$, $\beta_i = \frac{\gamma_{i+1}}{\gamma_i}$, $d_{i+1} = z_{i+1} + \beta_i d_i$. Erhöhe i um 1 und gehe zu Schritt 2.

201/ 264



Eigenschaften

- ▶ Die **Konvergenzrate** des PCG-Verfahrens ist $\frac{\sqrt{\hat{\kappa}}-1}{\sqrt{\hat{\kappa}}+1}$, wobei $\hat{\kappa}$ die Konditionszahl von \hat{L} ist.
- ▶ Bei geschickter Wahl von C , z.B. bei der **SSOR-Vorkonditionierung**, ist $\hat{\kappa} = N^{\frac{1}{d}}$, was die Konvergenzrate $1 - N^{-\frac{1}{2d}}$ ergibt.

202/ 264



SSOR-Vorkonditionierung

0. Gegeben: r und ein Relaxationsparameter $\omega \in (0, 2)$.
Gesucht: $z = C^{-1}r$.
1. Setze $z = 0$.
2. Berechne für $i = 1, \dots, N$

$$z_i = z_i + \omega L_{ii}^{-1} \left\{ r_i - \sum_{j=1}^N L_{ij} z_j \right\}.$$
3. Berechne für $i = N, \dots, 1$

$$z_i = z_i + \omega L_{ii}^{-1} \left\{ r_i - \sum_{j=1}^N L_{ij} z_j \right\}.$$

203/ 264



Die zugrundeliegende Idee

- ▶ Klassische iterative Löser wie die Gauß-Seidel-Iteration dämpfen schnell schwingende Fehlerkomponenten stark.
- ▶ Klassische iterative Löser wie die Gauß-Seidel-Iteration dämpfen dagegen langsam schwingende Fehlerkomponenten nur sehr schlecht.
- ▶ Langsam schwingende Fehlerkomponenten können auf einem größeren Gitter mit weniger Unbekannten gut approximiert werden.

204/ 264



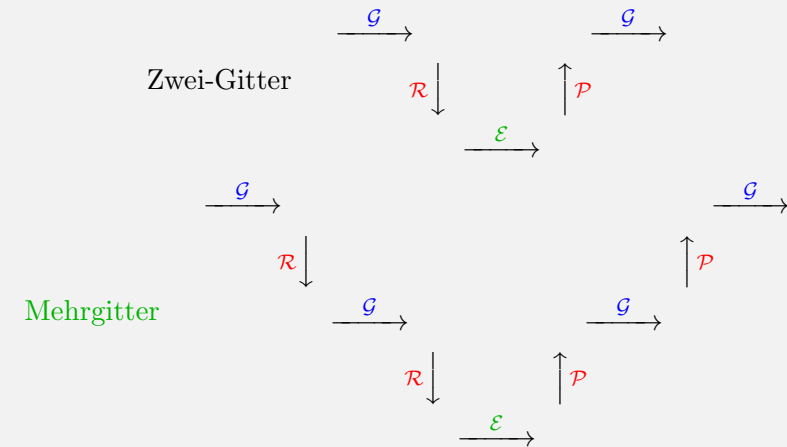
Der Zwei-Gitter-Algorithmus

- ▶ Führe auf dem aktuellen Gitter mehrere Schritte eines klassischen iterativen Verfahrens durch.
- ▶ Korrigiere die aktuelle Näherungslösung wie folgt:
 - ▶ Berechne das aktuelle Residuum.
 - ▶ Schränke das aktuelle Residuum auf das nächst größere Gitter ein.
 - ▶ Löse das resultierende Problem auf dem größeren Gitter exakt.
 - ▶ Setze die Grobgitter-Lösung durch Interpolation auf das nächst feinere Gitter fort.
- ▶ Führe auf dem aktuellen Gitter mehrere Schritte eines klassischen iterativen Verfahrens durch.

205 / 264



Schematische Darstellung



206 / 264



Ingredienzien

- ▶ Eine Folge \mathcal{T}_k zunehmend (gleichmäßig oder adaptiv) verfeinerter Unterteilungen mit zugehörigen diskreten Problemen $L_k u_k = f_k$.
- ▶ Ein **Glättungsoperator** M_k , der leicht auswertbar ist und der gleichzeitig eine passable Näherung für L_k^{-1} liefert.
- ▶ Ein **Restriktionsoperator** $R_{k,k-1}$, der Funktionen zur Unterteilung \mathcal{T}_k in solche zur nächst größeren Unterteilung \mathcal{T}_{k-1} abbildet.
- ▶ Ein **Prolongationsoperator** $I_{k-1,k}$, der Funktionen zur Unterteilung \mathcal{T}_{k-1} in solche zur nächst feineren Unterteilung \mathcal{T}_k abbildet.

207 / 264



Der Mehrgitter-Algorithmus

0. Gegeben: das aktuelle Niveau k , Parameter μ, ν_1, ν_2 , die Matrix L_k , die rechte Seite f_k , eine Startnäherung u_k .
Gesucht: eine verbesserte Näherungslösung u_k .
1. Falls $k = 0$ ist, berechne $u_0 = L_0^{-1} f_0$; **stopp**.
2. (**Vor-Glättung**) Führe ν_1 Schritte des Iterationsverfahrens $u_k \mapsto u_k + M_k(f_k - L_k u_k)$ durch.
3. (**Grobgitter-Korrektur**)
 - 3.1 Berechne $f_{k-1} = R_{k,k-1}(f_k - L_k u_k)$ und setze $u_{k-1} = 0$.
 - 3.2 Führe μ Iterationen des Mehrgitter-Algorithmus mit Parametern $k-1, \mu, \nu_1, \nu_2, L_{k-1}, f_{k-1}, u_{k-1}$ durch und bezeichne das Ergebnis mit u_{k-1} .
 - 3.3 Ersetze u_k durch $u_k + I_{k-1,k} u_{k-1}$.
4. (**Nach-Glättung**) Führe ν_2 Schritte des Iterationsverfahrens $u_k \mapsto u_k + M_k(f_k - L_k u_k)$ durch.

208 / 264



Typische Werte der Parameter

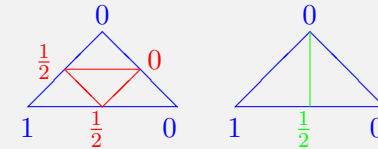
- ▶ $\mu = 1$ **V-Zyklus** oder
 $\mu = 2$ **W-Zyklus**
- ▶ $\nu_1 = \nu_2 = \nu$ oder
 $\nu_1 = \nu, \nu_2 = 0$ oder
 $\nu_1 = 0, \nu_2 = \nu$
- ▶ $1 \leq \nu \leq 4$.

209/ 264



Prolongation and Restriktion

- ▶ Die Prolongation ist typischerweise bestimmt durch die natürliche Inklusion der Finite-Element-Räume, d.h. eine Finite-Element-Funktion zu einer größeren Unterteilung kann durch die Basisfunktionen der aktuellen, feineren Unterteilung ausgedrückt werden.



- ▶ Die Restriktion wird üblicherweise dadurch berechnet, dass Basisfunktionen zur größeren Unterteilung in das diskrete Problem zur feineren Unterteilung als Testfunktionen eingesetzt werden.

210/ 264



Glättung

- ▶ Gauß-Seidel-Iteration
- ▶ SSOR-Iteration:
 - ▶ Durchlaufe die Gleichungen in aufsteigender Nummerierung mit einer Gauß-Seidel-Iteration als Vor-Glättung.
 - ▶ Durchlaufe die Gleichungen in absteigender Nummerierung mit einer Gauß-Seidel-Iteration als Nach-Glättung.
- ▶ **ILU**-Glättung:
 - ▶ Bestimme eine unvollständige LR -Zerlegung (**incomplete lower upper decomposition**) von L_k , indem Nullelemente der Matrix unterdrückt werden (**Unterdrückung von fill-in**).
 - ▶ Das Ergebnis ist eine näherungsweise Zerlegung $\mathcal{L}_k \mathcal{U}_k \approx L_k$.
 - ▶ Berechne $v_k = M_k u_k$ durch Lösen des Systems $\mathcal{L}_k \mathcal{U}_k v_k = u_k$.

211/ 264



Benötigte arithmetische Operationen

- ▶ Nehme an, dass
 - ▶ ein Glättungsschritt $O(N_k)$ Operationen erfordert,
 - ▶ die Prolongation in $O(N_k)$ Operationen berechnet werden kann,
 - ▶ die Restriktion $O(N_k)$ Operationen erfordert,
 - ▶ $\mu \leq 2$ ist,
 - ▶ $N_k > \mu N_{k-1}$ gilt.
- ▶ Dann benötigt eine Iteration des Mehrgitterverfahrens $O(N_k)$ Operationen.

212/ 264



Konvergenzraten

- ▶ Die Konvergenzrate ist kleiner als 1 gleichmäßig für alle Unterteilungen.
- ▶ Die Konvergenzrate ist durch $\frac{c}{c+\nu_1+\nu_2}$ beschränkt, wobei c nur von den Formparametern der Unterteilungen abhängt.
- ▶ In der Praxis werden typischerweise Konvergenzraten zwischen 0.1 und 0.5 erzielt.



Arithmetische Operationen

Beispiel: Lineare Finite-Element-Diskretisierung auf Courant-Triangulierung der Poisson-Gleichung im Einheitsquadrat; Anfangsfehler wird um Faktor 0.05 reduziert

h	Gauß-El.	GS	CG	PCG	MG
$\frac{1}{16}$	$7.6 \cdot 10^5$	$2.6 \cdot 10^5$	$2.7 \cdot 10^4$	$1.6 \cdot 10^4$	$1.2 \cdot 10^4$
$\frac{1}{32}$	$2.8 \cdot 10^7$	$4.5 \cdot 10^6$	$2.2 \cdot 10^5$	$8.6 \cdot 10^4$	$4.9 \cdot 10^4$
$\frac{1}{64}$	$9.9 \cdot 10^8$	$7.6 \cdot 10^7$	$1.9 \cdot 10^6$	$5.0 \cdot 10^5$	$2.1 \cdot 10^5$
$\frac{1}{128}$	$3.3 \cdot 10^{10}$	$1.2 \cdot 10^9$	$1.5 \cdot 10^7$	$3.2 \cdot 10^6$	$8.4 \cdot 10^5$



Iterationen

Beispiel: Lineare Finite-Element-Diskretisierung auf Courant-Triangulierung der Poisson-Gleichung im Einheitsquadrat; Anfangsfehler wird um Faktor 0.05 reduziert

h	GS	CG	PCG	MG
$\frac{1}{16}$	236	12	4	1
$\frac{1}{32}$	954	23	5	2
$\frac{1}{64}$	3820	47	7	2
$\frac{1}{128}$	15287	94	11	1



Iterationen und Konvergenzraten

Beispiel: Lineare Finite-Element-Diskretisierung mit adaptiver Verfeinerung einer Reaktions-Diffusions-Gleichung im Einheitsquadrat mit innerer Grenzschicht; Anfangsfehler wird um Faktor 0.05 reduziert

DOF	CG		PCG		MG	
	It.	κ	It.	κ	It.	κ
9	4	0.10	3	0.2	4	0.3
47	10	0.60	7	0.5	3	0.3
185	24	0.80	12	0.7	5	0.2
749	49	0.90	21	0.8	5	0.4
2615	94	0.95	37	0.9	6	0.4
5247	130	0.96	55	0.9	5	0.4



CG-Verfahren für unsymmetrische oder indefinite Probleme

- ▶ Das CG-Verfahren bricht für unsymmetrische oder indefinite Probleme (**Steifigkeitsmatrix hat Eigenwerte mit positivem und mit negativem Realteil**) in der Regel zusammen.
- ▶ Ein naiver Ausweg ist das Anwenden des CG-Verfahrens auf die symmetrischen, positiv definiten **Normalgleichungen** $L^T L u = L^T f$.
- ▶ Dadurch verdoppelt sich der Aufwand, weil sich die Konditionszahl bei Übergang zu den Normalgleichungen quadriert.
- ▶ Einen besseren Ausweg bieten spezielle Varianten des CG-Verfahrens wie das **stabilisierte bi-konjugierte Gradienten Verfahren (Bi-CG-Stab-Verfahren)**.

217/ 264



Bi-CG-Stab-Verfahren

0. Gegeben: eine Startnäherung u_0 und eine Toleranz $\varepsilon > 0$.
1. Berechne $r_0 = b - L u_0$ und setze $\bar{r}_0 = r_0$, $v_{-1} = 0$, $p_{-1} = 0$, $\alpha_{-1} = 1$, $\rho_{-1} = 1$, $\omega_{-1} = 1$, sowie $i = 0$.
2. Falls $r_i \cdot r_i < \varepsilon^2$ ist, gebe u_i als Näherungslösung aus; **stopp**. Andernfalls gehe zu Schritt 3.
3. Berechne $\rho_i = \bar{r}_i \cdot r_i$, $\beta_{i-1} = \frac{\rho_i \alpha_{i-1}}{\rho_{i-1} \omega_{i-1}}$. Falls $|\beta_{i-1}| < \varepsilon$ ist, liegt ein möglicher Abbruch vor; **stopp**. Andernfalls berechne $p_i = r_i + \beta_{i-1} \{p_{i-1} - \omega_{i-1} v_{i-1}\}$, $v_i = L p_i$, $\alpha_i = \frac{\rho_i}{\bar{r}_0 \cdot v_i}$. Falls $|\alpha_i| < \varepsilon$ ist, liegt ein möglicher Abbruch vor; **stopp**. Andernfalls berechne $s_i = r_i - \alpha_i v_i$, $t_i = L s_i$, $\omega_i = \frac{t_i \cdot s_i}{t_i \cdot t_i}$, $u_{i+1} = u_i + \alpha_i p_i + \omega_i s_i$, $r_{i+1} = s_i - \omega_i t_i$. Erhöhe i um 1 und gehe zu Schritt 2.

218/ 264



Eigenschaften

- ▶ Das Bi-CG-Stab-Verfahren versucht, simultan das ursprüngliche Problem $L u = f$ und das **adjungierte** Problem $L^T v = f$ zu lösen.
- ▶ Es benötigt nur die Steifigkeitsmatrix L des ursprünglichen Problems.
- ▶ Es erfordert nur Skalarprodukte und Matrix-Vektor-Multiplikationen.
- ▶ Das Bi-CG-Stab-Verfahren kann vorkonditioniert werden; mögliche Vorkonditionierer sind das SSOR-Verfahren oder die ILU-Zerlegung angewandt auf den symmetrischen Anteil $\frac{1}{2}(L + L^T)$ von L .

219/ 264



Mehrgitterverfahren für unsymmetrische oder indefinite Probleme

- ▶ Mehrgitterverfahren können direkt auf unsymmetrische oder indefinite Probleme angewendet werden.
- ▶ Unter Umständen muss man spezielle Glätter verwenden.
- ▶ Die Richardson-Relaxation angewandt auf die Normalgleichungen ist ein robuster Glätter, der allerdings zu Konvergenzraten von etwa 0.8 für das Mehrgitterverfahren führt.
- ▶ Die ILU-Zerlegung ist ebenfalls robust, aber aufwändiger und führt zu Konvergenzraten von etwa 0.5.

220/ 264



Nichtlineare Probleme

- ▶ Nichtlineare Probleme werden typischerweise mit einem (gedämpften) Newton-Verfahren gelöst.
- ▶ In jeder Iteration des Newton-Verfahrens ist ein lineares Problem zu lösen.
- ▶ Dies kann mit iterativen Lösern geschehen, das Ergebnis der vorhergehenden Newton-Iteration ist dann meist ein guter Startwert für die innere Iteration.
- ▶ Bei Mehrgitterverfahren kann auch die Rolle von äußerer und innerer Iteration vertauscht werden, dann werden wenige Newton-Iterationen verbunden mit einer wenig genauen Lösung der linearen Hilfsprobleme als Glätter verwendet.

221 / 264



Parabolische Differentialgleichungen

- ▶ Diskretisierungsmethoden
- ▶ Raum-Zeit Finite-Elemente
- ▶ Charakteristikenmethode
- ▶ Adaptivität

222 / 264



Modellproblem: Lineare parabolische Differentialgleichung zweiter Ordnung

$$\begin{aligned} \frac{\partial u}{\partial t} - \operatorname{div}(A\nabla u) + \mathbf{a} \cdot \nabla u + \alpha u &= f && \text{in } \Omega \times (0, T] \\ u &= 0 && \text{auf } \Gamma \times (0, T] \\ u(\cdot, 0) &= u_0 && \text{in } \Omega \end{aligned}$$

- ▶ Ω ein Polyeder in \mathbb{R}^d mit $d = 2$ oder $d = 3$
- ▶ $A(x, t)$ eine symmetrische, positiv definite, $d \times d$ Matrix für jedes x in Ω , t in $(0, T]$
- ▶ $\mathbf{a}(x, t)$ ein Vektor in \mathbb{R}^d für jedes x in Ω , t in $(0, T]$
- ▶ $\alpha(x, t)$ eine nicht-negative Zahl für jedes x in Ω , t in $(0, T]$
- ▶ $\alpha(x, t) - \frac{1}{2} \operatorname{div} \mathbf{a}(x, t) \geq 0$ für jedes x in Ω , t in $(0, T]$

223 / 264



Gebräuchliche Diskretisierungsmethoden

- ▶ Es gibt drei Hauptzugänge:
 - ▶ Linien-Methode,
 - ▶ Rothe-Verfahren,
 - ▶ Raum-Zeit Finite-Elemente.
- ▶ Für klassische nicht adaptive Unterteilungen liefern sie häufig die selben diskreten Lösungen.
- ▶ Die Linien-Methode ist unflexibel und für Adaptivität nicht geeignet.
- ▶ Die Analyse des Rothe-Verfahrens ist knifflig, da sie Differenzierbarkeitseigenschaften bzgl. der Zeitvariablen benötigt, die häufig nicht zur Verfügung stehen.
- ▶ Raum-Zeit Finite-Elemente erlauben a posteriori Fehlerabschätzungen und sind für Raum-Zeit-Adaptivität gut geeignet.

224 / 264



Linien-Methode

- ▶ Wähle eine feste Unterteilung \mathcal{T} von Ω und einen zugehörigen Finite-Element-Raum $X(\mathcal{T})$ (Ortsdiskretisierung); bezeichne mit $A_{\mathcal{T}}$ und $f_{\mathcal{T}}$ die Steifigkeitsmatrix und den Lastvektor.
- ▶ Dann liefert die Ortsdiskretisierung das folgende System gewöhnlicher Differentialgleichungen:

$$\frac{du_{\mathcal{T}}}{dt} = f_{\mathcal{T}} - A_{\mathcal{T}}u_{\mathcal{T}}.$$

- ▶ Wende hierauf ein Standardverfahren für gewöhnliche Differentialgleichungen an (impliziter Euler, Crank-Nicolson, Runge-Kutta, ...) (Zeitdiskretisierung).
- ▶ Das Crank-Nicolson-Verfahren ergibt z.B. die Vorschrift

$$\frac{u_{\mathcal{T}}^n - u_{\mathcal{T}}^{n-1}}{\tau} = \frac{1}{2}(f_{\mathcal{T}}^n - A_{\mathcal{T}}u_{\mathcal{T}}^n + f_{\mathcal{T}}^{n-1} - A_{\mathcal{T}}u_{\mathcal{T}}^{n-1}).$$



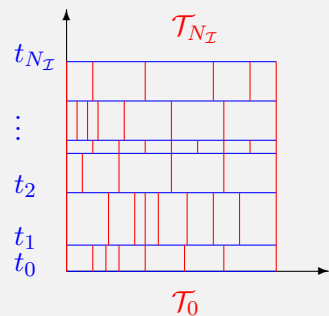
Rothe-Verfahren

- ▶ Interpretiere das parabolische Problem als eine gewöhnliche Differentialgleichung und wende hierauf ein Standardverfahren an (impliziter Euler, Crank-Nicolson, Runge-Kutta, ...) (Zeitdiskretisierung).
- ▶ Jeder Zeitschritt erfordert dann die Lösung einer stationären elliptischen Differentialgleichung, die mit einem üblichen Finite-Element-Verfahren diskretisiert wird (Ortsdiskretisierung).
- ▶ Das Crank-Nicolson-Verfahren ergibt z.B. die elliptischen Differentialgleichungen

$$\frac{u^n - u^{n-1}}{\tau} + \frac{1}{2}(-\operatorname{div}(A\nabla u^n) + \mathbf{a} \cdot \nabla u^n + \alpha u^n - \operatorname{div}(A\nabla u^{n-1}) + \mathbf{a} \cdot \nabla u^{n-1} + \alpha u^{n-1}) = \frac{1}{2}(f^n + f^{n-1}).$$



Raum-Zeit-Gitter



- ▶ $\mathcal{I} = \{[t_{n-1}, t_n] : 1 \leq n \leq N_{\mathcal{I}}\}$: Unterteilung von $[0, T]$ mit $0 = t_0 < \dots < t_{N_{\mathcal{I}}} = T$ (Die t_n werden im Laufe des Verfahrens sukzessive bestimmt.)
- ▶ $\tau_n = t_n - t_{n-1}$
- ▶ \mathcal{T}_n : Unterteilungen von Ω
- ▶ $X_n = X(\mathcal{T}_n)$: zugehörige Finite-Element-Räume



Raum-Zeit Finite-Element-Diskretisierung

Berechne eine Interpolierende $u_{\mathcal{T}_0}^0 \in X_0$ von u_0 und bestimme für $n = 1, 2, \dots$ sukzessive $u_{\mathcal{T}_n}^n \in X_n$ (Ansatzfunktion) so, dass mit $u^{n\theta} = \theta u_{\mathcal{T}_n}^n + (1 - \theta)u_{\mathcal{T}_{n-1}}^{n-1}$ für alle $v_{\mathcal{T}_n}^n \in X_n$ (Testfunktion) gilt:

$$\begin{aligned} & \int_{\Omega} \frac{1}{\tau_n} (u_{\mathcal{T}_n}^n - u_{\mathcal{T}_{n-1}}^{n-1}) v_{\mathcal{T}_n} dx + \int_{\Omega} \nabla u^{n\theta} \cdot A \nabla v_{\mathcal{T}_n} dx \\ & + \int_{\Omega} \mathbf{a} \cdot \nabla u^{n\theta} v_{\mathcal{T}_n} dx + \int_{\Omega} \alpha u^{n\theta} v_{\mathcal{T}_n} dx \\ & = \int_{\Omega} f v_{\mathcal{T}_n} dx \end{aligned}$$



Wahl von θ

- ▶ $\theta = \frac{1}{2}$ entspricht dem Crank-Nicolson-Verfahren.
- ▶ $\theta = 1$ entspricht dem impliziten Euler-Verfahren.
- ▶ $\theta = 0$ entspricht dem expliziten Euler-Verfahren.
- ▶ Aus Stabilitätsgründen sollte $\theta \geq \frac{1}{2}$ gewählt werden.



Eigenschaften

- ▶ Für $\theta > 0$ ist in jedem Zeitschritt ein lineares Gleichungssystem zu lösen, das der Finite-Element-Diskretisierung einer elliptischen Differentialgleichung entspricht.
- ▶ Für $\mathbf{a} \neq 0$ ist die Steifigkeitsmatrix unsymmetrisch und indefinit.
- ▶ Bei Verwenden eines iterativen Löser ist $u_{T_{n-1}}^{n-1}$ ein guter Startwert für die Berechnung von $u_{T_n}^n$.
- ▶ Der Fehler der Diskretisierung verhält sich wie $h^2 + \tau^\gamma$ mit $\gamma = 2$ für $\theta = \frac{1}{2}$ und $\gamma = 1$ für $\theta \neq \frac{1}{2}$ (h maximale Ortsgitterweite, τ maximale Zeitschrittweite).



Idee

- ▶ Für jeden Punkt $(x^*, t^*) \in \Omega \times (0, T]$ besitzt die folgende **Charakteristengleichung** eine eindeutige Lösung auf dem Intervall $(0, t^*)$

$$\frac{d}{dt}x(t; x^*, t^*) = \mathbf{a}(x(t; x^*, t^*), t), \quad x(t^*; x^*, t^*) = x^*.$$

- ▶ Für $U(x^*, t) = u(x(t; x^*, t^*), t)$ gilt

$$\frac{dU}{dt} = \frac{\partial u}{\partial t} + \mathbf{a} \cdot \nabla u.$$

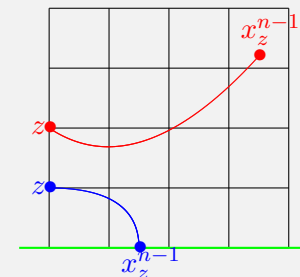
- ▶ Daher kann das Modellproblem geschrieben werden als

$$\frac{dU}{dt} - \text{div}(\mathbf{A} \nabla u) + \alpha u = f.$$

- ▶ Die **Reaktions-Diffusions-Gleichung** und die **Materialableitung** werden separat diskretisiert.



Re-Interpolation



- ▶ \mathcal{N}_n bezeichne die globalen Freiheitsgrade von X_n .
- ▶ Wende für jedes n und jedes $z \in \mathcal{N}_n$ ein klassisches Verfahren (impliziter Euler, Crank-Nicolson, Runge-Kutta, ...) auf die Charakteristengleichung zu $(x^*, t^*) = (z, t_n)$ an und bezeichne mit x_z^{n-1} die resultierende Näherung für $x(t_{n-1}; z, t_n)$.



Charakteristikenmethode

- ▶ Bestimme $\tilde{u}_{\mathcal{T}_n}^{n-1} \in X_n$ so, dass $\tilde{u}_{\mathcal{T}_n}^{n-1}(z) = u_{\mathcal{T}_{n-1}}^{n-1}(x_z^{n-1})$ ist für alle $z \in \mathcal{N}_n$.
- ▶ Bestimme $u_{\mathcal{T}_n}^n \in X_n$ so, dass für alle $v_{\mathcal{T}_n}^n \in X_n$ gilt

$$\begin{aligned} \frac{1}{\tau_n} \int_{\Omega} u_{\mathcal{T}_n}^n v_{\mathcal{T}_n}^n dx + \int_{\Omega} \nabla u_{\mathcal{T}_n}^n \cdot A \nabla v_{\mathcal{T}_n}^n dx + \int_{\Omega} \alpha u_{\mathcal{T}_n}^n v_{\mathcal{T}_n}^n dx \\ = \frac{1}{\tau_n} \int_{\Omega} \tilde{u}_{\mathcal{T}_n}^{n-1} v_{\mathcal{T}_n}^n dx + \int_{\Omega} f v_{\mathcal{T}_n}^n dx \end{aligned}$$



Eigenschaften

- ▶ Die **Charakteristikenmethode**, alias **Transport-Diffusions-Algorithmus**, ist besonders geeignet für die Diskretisierung parabolischer Gleichungen mit einem großen Konvektionsterm.
- ▶ Sie entkoppelt die Diskretisierung der Zeit- und Konvektionsableitungen von der Diskretisierung der restlichen Terme.
- ▶ Sie erfordert die Lösung einer Folge **gewöhnlicher Differentialgleichungen** und von **Reaktions-Diffusions-Gleichungen** mit **symmetrisch, positiv definiten Steifigkeitsmatrix**.



Ein residueller Fehlerschätzer

- ▶ \mathcal{E}_n : Kanten ($d = 2$) bzw. Seitenflächen ($d = 3$) der Elemente in \mathcal{T}_n

- ▶ **Ortsindikator**

$$\begin{aligned} \eta_h^n = \left\{ \sum_{K \in \mathcal{T}_n} h_K^2 \int_K \left| f(x, t_n) - \frac{1}{\tau_n} (u_{\mathcal{T}_n}^n - u_{\mathcal{T}_{n-1}}^{n-1}) \right. \right. \\ \left. \left. + \operatorname{div}(A \nabla u_{\mathcal{T}_n}^n) - \mathbf{a} \cdot \nabla u_{\mathcal{T}_n}^n - \alpha u_{\mathcal{T}_n}^n \right|^2 dx \right. \\ \left. + \sum_{E \in \mathcal{E}_n} h_E \int_E \left| [\mathbf{n}_E \cdot A \nabla u_{\mathcal{T}_n}^n]_E \right|^2 dS \right\}^{\frac{1}{2}} \end{aligned}$$

- ▶ **Zeitindikator**

$$\eta_{\tau}^n = \left\{ \int_{\Omega} \left| \nabla u_{\mathcal{T}_n}^n - \nabla u_{\mathcal{T}_{n-1}}^{n-1} \right|^2 dx \right\}^{\frac{1}{2}}$$



Bemerkungen

- ▶ η_h^n besteht aus den Elementresiduen der diskreten Lösung und aus Sprungtermen über die Elementgrenzen.
- ▶ Bei den Elementresiduen wird die diskrete Lösung elementweise in die Differentialgleichung eingesetzt.
- ▶ Die Sprungterme sind die gleichen wie bei der entsprechenden elliptischen Differentialgleichung (Zeitableitung unterdrückt).
- ▶ η_{τ}^n beschreibt einen Sprungterm bezüglich der Zeitvariablen.



A posteriori Fehlerabschätzungen

- ▶ Bezeichne mit $u_{\mathcal{I}}$ die stetige, bezüglich der Zeit stückweise lineare Funktion, die zur Zeit t_n mit $u_{\mathcal{T}_n}^n$ übereinstimmt.
- ▶ Dann gilt

$$\left\{ \max_{0 \leq t \leq T} \int_{\Omega} |u - u_{\mathcal{I}}|^2 dx + \int_0^T \int_{\Omega} |\nabla u - \nabla u_{\mathcal{I}}|^2 dx dt \right\}^{\frac{1}{2}}$$

$$\approx \left\{ \int_{\Omega} |u_{\mathcal{T}_0}^0 - u_0|^2 dx + \sum_{n=1}^{N_{\mathcal{I}}} \tau_n [(\eta_h^n)^2 + (\eta_{\tau}^n)^2] \right\}^{\frac{1}{2}}$$



Bemerkungen

- ▶ \approx bedeutet obere und untere Schranken modulo multiplikativer Faktoren.
- ▶ Diese Faktoren hängen von dem Polynomgrad der Finite-Element-Funktionen und den Formparametern der Unterteilungen ab.
- ▶ Die oberen Schranken sind global in Ort und Zeit.
- ▶ Die unteren Schranken sind global im Ort und lokal in der Zeit.
- ▶ η_h^n kontrolliert den Fehler der Ortsdiskretisierung.
- ▶ η_{τ}^n kontrolliert den Fehler der Zeitdiskretisierung.



Raum-Zeit Adaptivität

0. Gegeben: Toleranz ε , Unterteilung \mathcal{T}_0 , Zeitschritt τ_1
1. Passe \mathcal{T}_0 so an, dass $\int_{\Omega} |u_{\mathcal{T}_0}^0 - u_0|^2 dx \leq \frac{1}{4} \varepsilon^2$ ist.
Setze $n = 1, t_1 = \tau_1$.
2. Löse das diskrete Problem zur Zeit t_n und bestimme die Indikatoren η_h^n und η_{τ}^n .
3. Falls $\eta_{\tau}^n > \frac{\varepsilon}{2\sqrt{T}}$ ist, ersetze t_n durch $\frac{1}{2}(t_{n-1} + t_n)$ und gehe zu Schritt 2 zurück (Halbieren von τ_n).
4. Passe \mathcal{T}_n so an, dass $\eta_h^n \leq \frac{\varepsilon}{2\sqrt{T}}$ ist.
Falls $\eta_{\tau}^n < \frac{\varepsilon}{4\sqrt{T}}$ ist, verdoppele τ_n .
5. Falls $t_n = T$ ist, stopp.
Andernfalls setze $t_{n+1} = \min\{T, t_n + \tau_n\}$, erhöhe n um 1 und gehe zu Schritt 2 zurück.



Eigenschaften

- ▶ Am Ende des Algorithmus ist

$$\left\{ \int_{\Omega} |u_{\mathcal{T}_0}^0 - u_0|^2 dx + \sum_{n=1}^{N_{\mathcal{I}}} \tau_n [(\eta_h^n)^2 + (\eta_{\tau}^n)^2] \right\}^{\frac{1}{2}} \leq \varepsilon.$$
- ▶ Beim Anpassen von \mathcal{T}_n werden t_n, τ_n und η_{τ}^n fest gehalten.
- ▶ Die Anpassung von \mathcal{T}_n erfordert ggf. das wiederholte Lösen diskreter Probleme und die Neuberechnung von η_h^n .



Finite-Volumen-Methoden

- ▶ Systeme in Divergenzform
- ▶ Grundidee der Finite-Volumen-Verfahren
- ▶ Konstruktion der Gitter
- ▶ Konstruktion der numerischen Flüsse
- ▶ Zusammenhang mit Finite-Element-Methoden



Aufgabenstellung

- ▶ Gebiet: $\Omega \subset \mathbb{R}^d$
- ▶ Quelle: $\mathbf{g} : \mathbb{R}^m \times \Omega \times (0, \infty) \rightarrow \mathbb{R}^m$
- ▶ Masse: $\mathbf{M} : \mathbb{R}^m \rightarrow \mathbb{R}^m$
- ▶ Fluss: $\underline{\mathbf{F}} : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$
- ▶ Anfangswert: $\mathbf{U}_0 : \Omega \rightarrow \mathbb{R}^m$
- ▶ Problem: Finde $\mathbf{U} : \Omega \times (0, \infty) \rightarrow \mathbb{R}^m$ so, dass mit geeigneten **Randbedingungen** gilt

$$\frac{\partial \mathbf{M}(\mathbf{U})}{\partial t} + \operatorname{div} \underline{\mathbf{F}}(\mathbf{U}) = \mathbf{g}(\mathbf{U}, x, t) \quad \text{in } \Omega \times (0, \infty)$$

$$\mathbf{U}(\cdot, 0) = \mathbf{U}_0 \quad \text{in } \Omega$$

$$\operatorname{div} \underline{\mathbf{F}}(\mathbf{U}) = \left(\sum_{j=1}^d \frac{\partial \mathbf{F}(\mathbf{U})_{i,j}}{\partial x_j} \right)_{1 \leq i \leq m}$$



Advektive und Viskose Flüsse

- ▶ Der Fluss $\underline{\mathbf{F}}$ kann in zwei Beiträge aufgespalten werden:
 $\underline{\mathbf{F}} = \underline{\mathbf{F}}_{\text{adv}} + \underline{\mathbf{F}}_{\text{visc}}$.
- ▶ $\underline{\mathbf{F}}_{\text{adv}}$ heißt **advektiver Fluss** und enthält keine Ableitungen.
- ▶ $\underline{\mathbf{F}}_{\text{visc}}$ heißt **viskoser Fluss** und enthält räumliche Ableitungen.
- ▶ Der advektive Fluss modelliert Transport- oder Konvektionsphänomene.
- ▶ Der viskose Fluss modelliert Diffusionsphänomene.



Beispiele

- ▶ Lineare parabolische Differentialgleichungen zweiter Ordnung:
 - ▶ $\frac{\partial u}{\partial t} - \operatorname{div}(A \nabla u) + \mathbf{a} \cdot \nabla u + \alpha u = f$
 - ▶ $m = 1$
 - ▶ $\mathbf{U} = u$
 - ▶ $\mathbf{M}(\mathbf{U}) = u$
 - ▶ $\underline{\mathbf{F}}_{\text{adv}}(\mathbf{U}) = \mathbf{a}u$
 - ▶ $\underline{\mathbf{F}}_{\text{visc}}(\mathbf{U}) = -A \nabla u$
 - ▶ $\mathbf{g}(\mathbf{U}) = f - \alpha u + (\operatorname{div} \mathbf{a})u$
- ▶ Euler-Gleichungen
- ▶ Kompressible Navier-Stokes-Gleichungen
- ▶ Burger-Gleichung



Erster Schritt

- ▶ Wähle einen Zeitschritt $\tau > 0$.
- ▶ Wähle eine Unterteilung \mathcal{T} von Ω in beliebige, nicht überlappende Polyeder.
- ▶ Betrachte ein $n \in \mathbb{N}^*$ und ein $K \in \mathcal{T}$.
- ▶ Integriere das System über $K \times [(n-1)\tau, n\tau]$:

$$\begin{aligned} & \int_{(n-1)\tau}^{n\tau} \int_K \frac{\partial \mathbf{M}(\mathbf{U})}{\partial t} dxdt + \int_{(n-1)\tau}^{n\tau} \int_K \operatorname{div} \underline{\mathbf{F}}(\mathbf{U}) dxdt \\ &= \int_{(n-1)\tau}^{n\tau} \int_K \mathbf{g}(\mathbf{U}, x, t) dxdt \end{aligned}$$



Zweiter Schritt

Integriere die Terme auf der linken Seite partiell:

$$\begin{aligned} \int_{(n-1)\tau}^{n\tau} \int_K \frac{\partial \mathbf{M}(\mathbf{U})}{\partial t} dxdt &= \int_K \mathbf{M}(\mathbf{U}(x, n\tau)) dx \\ &\quad - \int_K \mathbf{M}(\mathbf{U}(x, (n-1)\tau)) dx \\ \int_{(n-1)\tau}^{n\tau} \int_K \operatorname{div} \underline{\mathbf{F}}(\mathbf{U}) dxdt &= \int_{(n-1)\tau}^{n\tau} \int_{\partial K} \underline{\mathbf{F}}(\mathbf{U}) \cdot \mathbf{n}_K dSdt \end{aligned}$$



Dritter Schritt

- ▶ Nehme an, dass \mathbf{U} bezüglich Ort und Zeit stückweise konstant ist.
- ▶ Bezeichne mit \mathbf{U}_K^n und \mathbf{U}_K^{n-1} den Wert von \mathbf{U} auf K zu den Zeiten $n\tau$ und $(n-1)\tau$:

$$\begin{aligned} \int_K \mathbf{M}(\mathbf{U}(x, n\tau)) dx &\approx |K| \mathbf{M}(\mathbf{U}_K^n) \\ \int_K \mathbf{M}(\mathbf{U}(x, (n-1)\tau)) dx &\approx |K| \mathbf{M}(\mathbf{U}_K^{n-1}) \\ \int_{(n-1)\tau}^{n\tau} \int_{\partial K} \underline{\mathbf{F}}(\mathbf{U}) \cdot \mathbf{n}_K dSdt &\approx \tau \int_{\partial K} \underline{\mathbf{F}}(\mathbf{U}_K^{n-1}) \cdot \mathbf{n}_K dS \\ \int_{(n-1)\tau}^{n\tau} \int_K \mathbf{g}(\mathbf{U}, x, t) dxdt &\approx \tau |K| \mathbf{g}(\mathbf{U}_K^{n-1}, x_K, (n-1)\tau) \end{aligned}$$



Vierter Schritt

Approximiere das Randintegral für den Fluss durch einen numerischen Fluss:

$$\begin{aligned} & \tau \int_{\partial K} \underline{\mathbf{F}}(\mathbf{U}_K^{n-1}) \cdot \mathbf{n}_K dS \\ & \approx \tau \sum_{\substack{K' \in \mathcal{T} \\ \partial K \cap \partial K' \in \mathcal{E}}} |\partial K \cap \partial K'| \mathbf{F}_{\mathcal{T}}(\mathbf{U}_K^{n-1}, \mathbf{U}_{K'}^{n-1}) \end{aligned}$$



Resultierendes Finite-Volumen-Verfahren

- ▶ Berechne für jedes Element $K \in \mathcal{T}$

$$\mathbf{U}_K^0 = \frac{1}{|K|} \int_K \mathbf{U}_0(x).$$

- ▶ Berechne für $n = 1, 2, \dots$ sukzessive für jedes Element $K \in \mathcal{T}$

$$\begin{aligned} \mathbf{M}(\mathbf{U}_K^n) &= \mathbf{M}(\mathbf{U}_K^{n-1}) \\ &\quad - \tau \sum_{\substack{K' \in \mathcal{T} \\ \partial K \cap \partial K' \in \mathcal{E}}} \frac{|\partial K \cap \partial K'|}{|K|} \mathbf{F}_{\mathcal{T}}(\mathbf{U}_K^{n-1}, \mathbf{U}_{K'}^{n-1}) \\ &\quad + \tau \mathbf{g}(\mathbf{U}_K^{n-1}, x_K, (n-1)\tau). \end{aligned}$$



Mögliche Modifikationen

- ▶ Die Zeitschrittweite kann variabel sein.
- ▶ Die Unterteilung von Ω kann sich von Zeitschritt zu Zeitschritt ändern.
- ▶ Die Näherung \mathbf{U}_K^n muss nicht stückweise konstant sein.



Noch offene Aufgaben

- ▶ Konstruiere die Unterteilung \mathcal{T} .
- ▶ Bestimme den numerischen Fluss $\mathbf{F}_{\mathcal{T}}$.
- ▶ Berücksichtige Randbedingungen.



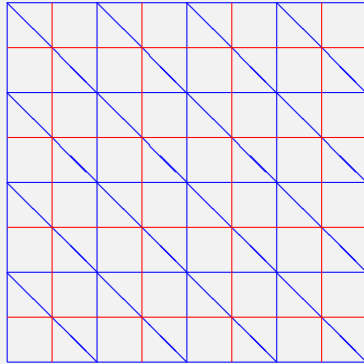
Duale Gitter

- ▶ Die Unterteilung \mathcal{T} wird häufig als **duales Gitter** zu einer zulässigen **primalen Finite-Element-Unterteilung** $\tilde{\mathcal{T}}$ konstruiert.
- ▶ Für Probleme in zwei Raumdimensionen ($d = 2$) gibt es hierfür im wesentlichen zwei verschiedene Ansätze:
 - ▶ Konstruiere für jedes Element $\tilde{K} \in \tilde{\mathcal{T}}$ die Mittelsenkrechten.
 - ▶ Verbinde für jedes Element $\tilde{K} \in \tilde{\mathcal{T}}$ seinen Schwerpunkt mit den Kantenmittelpunkten.

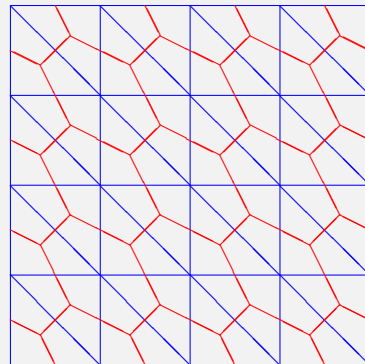


Mittelsenkrechten und Schwerpunkte

Mittelsenkrechten



Schwerpunkte

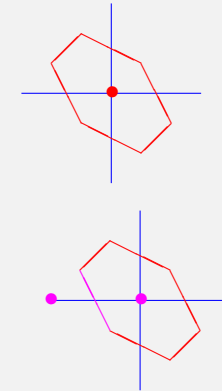


253 / 264



Eigenschaften dualer Gitter

- ▶ Jedes Element in $K \in \mathcal{T}$ entspricht einem Elementeckpunkt x_K von $\tilde{\mathcal{T}}$ und umgekehrt.
- ▶ Zu jeder Kante E von \mathcal{T} gibt es zwei Elementeckpunkte $x_{E,1}$, $x_{E,2}$ von $\tilde{\mathcal{T}}$ so, dass die Strecke $\overline{x_{E,1}x_{E,2}}$ die Kante E schneidet.



254 / 264



Vor- und Nachteile von Mittelsenkrechten

- ▶ Die Strecke $\overline{x_{E,1}x_{E,2}}$ und die Kante E schneiden sich in einem rechten Winkel.
- ▶ Die Mittelsenkrechten eines Dreiecks können sich in einem Punkt außerhalb des Dreiecks schneiden. Der Schnittpunkt der Mittelsenkrechten liegt genau dann im Dreieck, wenn das Dreieck spitzwinklig ist.
- ▶ Die Mittelsenkrechten eines Vierecks brauchen sich nicht zu schneiden. Die Mittelsenkrechten eines Vierecks schneiden sich genau dann in einem Punkt, wenn das Viereck ein Rechteck ist.
- ▶ Die Konstruktion mit Mittelsenkrechten hat kein dreidimensionales Analogon.

255 / 264



Bezeichnungen und Annahmen

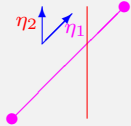
- ▶ \mathcal{T} sei ein **duales Gitter** zu einer primalen Finite-Element-Unterteilung $\tilde{\mathcal{T}}$.
- ▶ Für jede Kante oder Fläche E von \mathcal{T} seien
 - ▶ K_1 und K_2 die **angrenzenden Volumina**,
 - ▶ \mathbf{U}_1 , \mathbf{U}_2 die Werte $\mathbf{U}_{K_1}^{n-1}$ und $\mathbf{U}_{K_2}^{n-1}$,
 - ▶ x_1 , x_2 die Elementeckpunkte von $\tilde{\mathcal{T}}$, für die die Strecke $\overline{x_1x_2}$ die Kante oder Fläche E schneidet.
- ▶ Spalte den numerischen Fluss $\mathbf{F}_{\mathcal{T}}(\mathbf{U}_1, \mathbf{U}_2)$ in einen **viskosen numerischen Fluss** $\mathbf{F}_{\mathcal{T},\text{visc}}(\mathbf{U}_1, \mathbf{U}_2)$ und einen **advektiven numerischen Fluss** $\mathbf{F}_{\mathcal{T},\text{adv}}(\mathbf{U}_1, \mathbf{U}_2)$ auf.

256 / 264



Approximation der viskosen Flüsse

- ▶ Führe ein lokales Koordinatensystem η_1, \dots, η_d so ein, dass die Richtung η_1 parallel ist zur Richtung von $\overline{x_1 x_2}$ und die restlichen Richtungen tangential sind zu E .



- ▶ Drücke alle Ableitungen in \mathbf{F}_{visc} durch partielle Ableitungen bezüglich des neuen Koordinatensystems aus.
- ▶ Unterdrücke alle partiellen Ableitungen, die nicht η_1 betreffen.
- ▶ Approximiere partielle Ableitungen bezüglich η_1 durch Differenzenquotienten der Form $\frac{\varphi_1 - \varphi_2}{|x_1 - x_2|}$.



Spektralzerlegung der advektiven Flüsse

- ▶ Bezeichne mit $C(\mathbf{V}) = D(\mathbf{F}_{\text{adv}}(\mathbf{V}) \cdot \mathbf{n}_{K_1}) \in \mathbb{R}^{m \times m}$ die Ableitung von $\mathbf{F}_{\text{adv}}(\mathbf{V}) \cdot \mathbf{n}_{K_1}$ nach \mathbf{V} .
- ▶ Nehme an, dass diese Matrix diagonalisierbar ist (für Euler- und kompressible Navier-Stokes-Gleichungen erfüllt)

$$Q(\mathbf{V})^{-1} C(\mathbf{V}) Q(\mathbf{V}) = \Delta(\mathbf{V})$$

mit einer invertierbaren Matrix $Q(\mathbf{V}) \in \mathbb{R}^{m \times m}$ und einer Diagonalmatrix $\Delta(\mathbf{V}) \in \mathbb{R}^{m \times m}$.

- ▶ Setze $z^+ = \max\{z, 0\}$, $z^- = \min\{z, 0\}$ und

$$\Delta(\mathbf{V})^\pm = \text{diag}(\Delta(\mathbf{V})_{11}^\pm, \dots, \Delta(\mathbf{V})_{mm}^\pm),$$

$$C(\mathbf{V})^\pm = Q(\mathbf{V}) \Delta(\mathbf{V})^\pm Q(\mathbf{V})^{-1}.$$



Approximation der advektiven Flüsse

- ▶ Steger-Warming

$$\mathbf{F}_{\mathcal{T}, \text{adv}}(\mathbf{U}_1, \mathbf{U}_2) = C(\mathbf{U}_1)^+ \mathbf{U}_1 + C(\mathbf{U}_2)^- \mathbf{U}_2$$

- ▶ van Leer

$$\mathbf{F}_{\mathcal{T}, \text{adv}}(\mathbf{U}_1, \mathbf{U}_2) = \left[C(\mathbf{U}_1) + C\left(\frac{1}{2}(\mathbf{U}_1 + \mathbf{U}_2)\right)^+ - C\left(\frac{1}{2}(\mathbf{U}_1 + \mathbf{U}_2)\right)^- \right] \mathbf{U}_1 + \left[C(\mathbf{U}_2) - C\left(\frac{1}{2}(\mathbf{U}_1 + \mathbf{U}_2)\right)^+ + C\left(\frac{1}{2}(\mathbf{U}_1 + \mathbf{U}_2)\right)^- \right] \mathbf{U}_2$$



Eigenschaften

- ▶ Beide Ansätze erfordern die Berechnung von $D\mathbf{F}_{\text{adv}}(\mathbf{V}) \cdot \mathbf{n}_{K_1}$ und der entsprechenden Eigenwerte und Eigenvektoren für geeignete Werte von \mathbf{V} .
- ▶ Im allgemeinen ist der Ansatz von van Leer aufwändiger als der von Steger-Warming, da er drei statt zwei Auswertungen von $C(\mathbf{V})$ erfordert.
- ▶ Für die kompressiblen Navier-Stokes- und Euler-Gleichungen kann dieser Mehraufwand vermieden werden, da diese Gleichungen die spezielle Struktur $\mathbf{F}_{\text{adv}}(\mathbf{V}) \cdot \mathbf{n}_{K_1} = C(\mathbf{V})\mathbf{V}$ haben.



Ein eindimensionales Beispiel

- ▶ **Burger-Gleichung:** $\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0$
- ▶ $\underline{\mathbf{F}}_{\text{adv}}(u) = \frac{1}{2}u^2$, $C(u) = u$, $C(u)^\pm = u^\pm$
- ▶ **Steger-Warming:**

$$\underline{\mathbf{F}}_{\mathcal{T},\text{adv}}(u_1, u_2) = \begin{cases} u_1^2 & \text{if } u_1 \geq 0, u_2 \geq 0 \\ u_1^2 + u_2^2 & \text{if } u_1 \geq 0, u_2 \leq 0 \\ u_2^2 & \text{if } u_1 \leq 0, u_2 \leq 0 \\ 0 & \text{if } u_1 \leq 0, u_2 \geq 0 \end{cases}$$

- ▶ **van Leer:**

$$\underline{\mathbf{F}}_{\mathcal{T},\text{adv}}(u_1, u_2) = \begin{cases} u_1^2 & \text{if } u_1 \geq -u_2 \\ u_2^2 & \text{if } u_1 \leq -u_2 \end{cases}$$



Zugeordnete Finite-Element-Funktionen

- ▶ \mathcal{T} sei ein duales Gitter zu einer primalen Finite-Element-Unterteilung $\tilde{\mathcal{T}}$.
- ▶ Dann gibt es eine Eins-zu-Eins-Beziehung zwischen stückweise konstanten Funktionen zu \mathcal{T} und stetigen stückweise linearen Funktionen zu $\tilde{\mathcal{T}}$:

$$S^{0,-1}(\mathcal{T})^m \ni \mathbf{U}_{\mathcal{T}} \leftrightarrow \tilde{\mathbf{U}}_{\tilde{\mathcal{T}}} \in S^{1,0}(\tilde{\mathcal{T}})^m$$

$$\mathbf{U}_{\mathcal{T}|_K} = \tilde{\mathbf{U}}_{\tilde{\mathcal{T}}}(x_K) \quad \text{für alle } K \in \mathcal{T}.$$



Ein einfacher adaptiver Algorithmus

- ▶ Zu gegebener Lösung $\mathbf{U}_{\mathcal{T}}$ der Finite-Volumen-Diskretisierung berechne die zugehörige Finite-Element-Funktion $\tilde{\mathbf{U}}_{\tilde{\mathcal{T}}}$.
- ▶ Wende einen gebräuchlichen Fehlerschätzer auf $\tilde{\mathbf{U}}_{\tilde{\mathcal{T}}}$ an.
- ▶ Basierend auf diesem Fehlerschätzer wende eine gebräuchliche Verfeinerungsstrategie auf $\tilde{\mathcal{T}}$ an und konstruiere so eine neue lokal verfeinerte Unterteilung $\hat{\mathcal{T}}$.
- ▶ Benutze $\hat{\mathcal{T}}$ als primales Gitter zur Konstruktion eines neuen dualen Gitters \mathcal{T}' . Dies ist die Verfeinerung von \mathcal{T} .



Literatur

- 📄 D. Braess
Finite Elemente
Springer 2007
- 📄 W. Dahmen, A. Reusken
Numerik für Ingenieure und Naturwissenschaftler
Springer 2006
- 📄 P. Deuffhard, F. Bornemann
Numerische Mathematik II
de Gruyter 2002
- 📄 H. R. Schwarz, N. Köckler
Numerische Mathematik
Vieweg-Teubner 2009