Inhaltsverzeichnis

Einführung in das Wissenschaftliche Rechnen

- Wie wird Wasserverteilung im Boden beschrieben?
- Wie wird die Ausbreitung von Verschmutzung beschrieben?
- Wie wird ein biologischer Abbauprozess beschrieben?
- Wie werden Unsicherheiten (Materialparameter) quantifiziert?
- Wie können die Parameter gemessen werden?
- Wie kann der Prozess (etwa durch Abpumpen) gesteuert werden? Vereinfachung möglich?

Numerische Methodenentwicklung:

- Wie können die Modelle approximiert werden?
- Ist das approximierte Modell eindeutig lösbar?
- Konvergieren die Approximationen?
- Wie berechnet man die Approximation effizient?
- Wie werden die Algorithmen realisiert?
- Wie werden die Ergebnisse dargestellt?
- Wie kann man die Effizienz/Qualität der Approximation/Algorithmen quantifizieren?

Struktur der Lösung einer Praktikumsaufgabe:

1. Problemstellung
2. Mathematisches Modell
3. Konfiguration eines Beispiels (Daten)
4. Numerische Methoden
5. Visualisierung der Ergebnisse
6. Qualitative Auswertung
7. Diskussion der Ergebnisse (Insgesamt 1-2 Seiten)

Ergänzung B.Sc.-Arbeit (Literaturrecherche)

- Eigenschaften des Modells
- Eigenschaften der numerischen Verfahren (Insgesamt 40-60 Seiten)

Anforderung Promotion: neue Methode / neues Ergebnis.

1 Die Potentialströmung

Problem

$\Omega \subset \mathbb{R}^3$ beschreibe eine poröse Bodenschicht. Von oben sickert Regenwasser langsam nach unten ins Grundwasser. Ziel ist die Bestimmung der transportierten Wassermenge sowie deren Verteilung.

Modell

Sei $q\colon \Omega \times [0,T] \longrightarrow \mathbb{R}^3$ der Fluss mit $\operatorname{div} q = 0$ und $\varrho\colon \Omega \times [0,T] \longrightarrow \mathbb{R}_{\geq 0}$ die Massendichte, d. h. für jedes Kontrollvolumen $Y \subset \Omega$ ist durch

$$\int\limits_Y \varrho(x,t)\,\mathrm{d}x$$

die Wassermenge in Y zum Zeitpunkt t gegeben. Die zeitliche Änderung der Wassermenge in Y ergibt sich durch den Zu- und Abfluss am Rand von Y. Dieser Zusammenhang lässt sich mithilfe der Bilanzgleichung

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_Y \varrho(x,t)\,\mathrm{d}x = -\int\limits_{\partial Y} \varrho(x,t)q(x,t)\cdot n\,\mathrm{da}\,.$$

mathematisch formulieren. Dabei bezeichne ∂Y den Rand von Y und n die nach außen orientierte Einheitsnormale auf ∂Y. Der Gaußsche Integralsatz liefert

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_Y \varrho(x,t)\,\mathrm{d}x = -\int\limits_{\partial Y} \operatorname{div}(\varrho q)\,\mathrm{da}\,.$$

Damit erhält man die partielle Differentialgleichung

$$\partial_t\varrho + \operatorname{div}(\varrho q) = 0.$$

Im Fall $\partial_t\varrho \equiv 0$ spricht man von einer stationären Strömung. Ist außerdem die Dichte $\varrho \equiv \varrho_0$ konstant, vereinfacht sich die partielle Differentialgleichung zu $\operatorname{div} q = 0$.

(1.1) Lemma
Sei $\operatorname{div} q = 0$ *in* Ω*. Definiere* $\Gamma_{\text{in}} = \{x \in \partial\Omega : q(x)\cdot n(x) < 0\}$ *und* $\Gamma_{\text{out}} = \{x \in \partial\Omega : q(x)\cdot n(x) > 0\}$*. Dann gilt*

$$\int\limits_{\Gamma_{in}} q\cdot n\,\mathrm{da} = -\int\limits_{\Gamma_{out}} q\cdot n\,\mathrm{da}$$

Beweis.

$$0 = \int\limits_\Omega \operatorname{div} q\,\mathrm{d}x = \int\limits_{\partial\Omega} q\cdot n\,\mathrm{da} = \int\limits_{\Gamma_{in}} q\cdot n\,\mathrm{da} + \int\limits_{\Gamma_{out}} q\cdot n\,\mathrm{da} \qquad \square$$

Darcy-Gesetz

Das Darcy-Gesetz

$$q = -\kappa(\nabla p - G)$$

beschreibt den Fluss eines Fluids in einem porösen Medium, dabei sind

$$\begin{array}{ll} p\colon \Omega \longrightarrow \mathbb{R} & \text{hydrostatischer Druck} \\ \kappa\colon \Omega \longrightarrow \mathbb{R}^{3\times 3}_{\mathrm{sym}} & \text{Permeabilitätstensor} \\ G = (0, 0, \varrho_0 g_0)^T & \text{Gravitation.} \end{array}$$

Ist die Permeabilität richtungsunabhängig, d. h. $\kappa = \kappa_0 I$, spricht man von isotropem Verhalten. Wenn zusätzlich $\kappa \equiv \kappa_0$ konstant ist, dann ist die Permeabilität ortsunabhängig (= homogen).

Mit der Hilfsgröße $u(x) = -p(x) + \varrho_0 g_0 x_3$ vereinfacht sich das Darcy-Gesetz zu $q = \kappa\nabla u$. Damit ergibt sich die partielle Differentialgleichung

$$\operatorname{div} \kappa\nabla u = 0.$$

Bemerkung
Im Fall ortsunabhängiger Permeabilität entspricht das der Laplace-Gleichung

$$(\operatorname{div} \nabla u =)\Delta u = 0.$$

Randwertaufgabe

Sei $\partial\Omega = \Gamma_{\mathrm{N}} \cup \Gamma_{\mathrm{D}}$. Durch die Vorgabe von Funktionen

$$\begin{array}{rl} u_{\mathrm{D}}\colon \Gamma_{\mathrm{D}} \longrightarrow \mathbb{R} & \text{(Dirichlet-Werte)} \\ g_{\mathrm{N}}\colon \Gamma_{\mathrm{N}} \longrightarrow \mathbb{R} & \text{(Neumann-Werte)} \end{array}$$

ergibt sich die Randwertaufgabe: Bestimme $u\colon \overline{\Omega} \longrightarrow \mathbb{R}$ mit

$$\begin{array}{rcll} \operatorname{div} \kappa\nabla u & = & 0 & \text{in } \Omega \\ u & = & u_{\mathrm{D}} & \text{auf } \Gamma_{\mathrm{D}} \\ \kappa\nabla u \cdot n & = & g_{\mathrm{N}} & \text{auf } \Gamma_{\mathrm{N}}. \end{array}$$

2D-Reduktion

Sei $\Omega_{2\mathrm{D}} = \{(x_1, x_3) \in \mathbb{R}^2 : (x_1, 0, x_3) \in \Omega_{3\mathrm{D}}\}$ ein Querschnitt von $\Omega_{2\mathrm{D}}$. Die Lösung sei lokal bei Ω_0 symmetrisch um $x_2 = 0$. Dann gilt $\partial_2 u_{3\mathrm{D}} = 0$. Mit $u_{2\mathrm{D}}(x_1, x_3) = u_{3\mathrm{D}}(x_1, 0, x_3)$ folgt

$$\operatorname{div} \kappa_{2\mathrm{D}}\nabla u_{2\mathrm{D}} = 0,$$

wobei $\kappa_{2\mathrm{D}}(x_1, x_3) \in \mathbb{R}^{2\times 2}_{\mathrm{sym}}$. Man kann also allgemein von $\Omega \subset \mathbb{R}^D$ und $\kappa \in \mathbb{R}^{D\times D}_{\mathrm{sym}}$ ausgehen, wobei D die Werte 2 und 3 annehmen kann.

Beispiel
Sei $\Omega = (0,1)^2, \Gamma_{\mathrm{D}} = [0,1] \times \{0\}, \Gamma_{\mathrm{N}} = \partial\Omega \setminus \Gamma_{\mathrm{D}}, u_D \equiv 0, \kappa = I$ *und*

$$g_{\mathrm{N}}(x) = \begin{cases} -1 & \textit{falls } x_3 = 1 \\ 0 & \textit{sonst} \end{cases}.$$

Die exakte Lösung der zugehörigen Randwertaufgabe lautet $u(x_1, x_3) = -x_3$*, denn dann gilt*

$$\Delta u = 0,$$

sowie

$$\nabla u \cdot n = \begin{cases} -1 & \textit{falls } x_2 = 1 \\ 1 & \textit{falls } x_2 = 0 \\ 0 & \textit{sonst} \end{cases}.$$

Schwache Formulierung

(1.2) Lemma

a) Sei $\operatorname{div} q = 0$ *in* Ω*, dann gilt für jede Testfunktion* $\phi\colon \Omega \longrightarrow \mathbb{R}$ *mit* $\phi = 0$ *auf* Γ_{D}

$$\int\limits_{\Omega} q \cdot \nabla\phi \,\mathrm{d}x = \int\limits_{\partial\Omega} q \cdot n\phi \,\mathrm{da}$$

b) Sei

$$\int\limits_{\Omega} q \cdot \nabla\phi \,\mathrm{d}x = \int\limits_{\partial\Omega} g\phi \,\mathrm{da}$$

für alle Testfunktionen. Dann gilt $\operatorname{div} q = 0$ *in* Ω *und* $q \cdot n = g$ *auf* $\partial\Omega$*.*

Beweis. a)

$$\operatorname{div}(\phi q) = \nabla\phi \cdot q + \phi \operatorname{div} q$$

$$\Rightarrow \int\limits_{\partial\Omega} \phi q \cdot n \,\mathrm{da} = \int\limits_{\Omega} \nabla\phi \cdot q \,\mathrm{d}x + \int\limits_{\Omega} \phi \operatorname{div} q \,\mathrm{d}x$$

b)

$$\int\limits_{\partial\Omega} \phi(q \cdot n - g) \,\mathrm{da} = \int\limits_{\Omega} \phi \operatorname{div} q \,\mathrm{d}x$$

$\Rightarrow \operatorname{div} q = 0$ in Ω
$\Rightarrow q \cdot n - g = 0$ auf $\partial\Omega$. □

Anwendung auf Randwertaufgabe

Starke Form:

$$\operatorname{div} q = 0, q = -\kappa \nabla u \text{ in } \mathit{\Omega}$$
$$-q \cdot n = g_{\mathrm{N}} \text{ auf } \mathit{\Gamma}_{\mathrm{N}},\, u = u_{\mathrm{D}} \text{ auf } \mathit{\Gamma}_{\mathrm{D}}$$

Schwache Form:

$$\int\limits_{\mathit{\Omega}} \kappa \nabla u \cdot \nabla \phi \,\mathrm{d}x = \int\limits_{\mathit{\Gamma}_{\mathrm{N}}} g_{\mathrm{N}} \phi \,\mathrm{d}a$$

$\forall \phi$ mit $\phi = 0$ auf $\mathit{\Gamma}_{\mathrm{D}}$

2 Finite-Elemente-Approximation

Um numerisch eine näherungsweise Lösung der schwachen Formulierung der Potentialströmung zu bestimmen, werden ein endlich-dimensionaler Funktionenraum V_h konstruiert und ein $u_h \in V_h$ mit $u_h \approx u_\mathrm{D}$ auf Γ_D und

$$\int\limits_{\Omega} \kappa \nabla u_h \cdot \nabla \phi_h \,\mathrm{d}x = \int\limits_{\Gamma_\mathrm{N}} g_\mathrm{N} \phi_h \,\mathrm{d}a \quad \forall \phi_h \in V_h \text{ mit } \phi_h = 0 \text{ auf } \Gamma_\mathrm{D} \tag{1}$$

bestimmt.

Wir betrachten im Folgenden den Spezialfall, dass $\Omega \subset \mathbb{R}^2$ ein Polygongebiet ist. Die hier vorgestellte Vorgehensweise lässt sich auf dreidimensionale Problemstellungen übertragen. Es sei eine Zerlegung $\overline{\Omega} = \bigcup_{K \in \mathcal{K}} \overline{K}$ von Ω in Dreiecke oder Vierecke gegeben. Diese sogenannten Zellen werden über

$$\overline{K} = \operatorname{conv}(\mathcal{V}_K)$$

mit

$$\mathcal{V}_K = \begin{cases} \{z_{K,0}, z_{K,1}, z_{K,2}\} \subset \mathbb{R}^2 & \text{Dreieck} \\ \{z_{K,0}, z_{K,1}, z_{K,2}, z_{K,3}\} \subset \mathbb{R}^2 & \text{Viereck} \end{cases}$$

beschrieben. Die Menge aller Eckpunkte der Zerlegung wird mit

$$\mathcal{V}_\mathcal{K} = \bigcup_{K \in \mathcal{K}} \mathcal{V}_K = \{z_1, \dots, z_N\}$$

notiert. Im Folgenden sei die Zerlegung $\mathcal{K}$ zulässig.

(2.1) Definition
Eine Zerlegung $\mathcal{K}$ heißt zulässig, wenn für alle $K, K' \in \mathcal{K}$

$$\operatorname{conv}(\mathcal{V}_K \cap \mathcal{V}_{K'}) = \operatorname{conv} \mathcal{V}_K \cap \operatorname{conv} \mathcal{V}_{K'}$$

gilt, d. h. wenn der Durchschnitt von zwei (verschiedenen) Zellen leer, eine Ecke oder eine Kante ist.

Idee

Wähle Knotenwerte $u(z)$ für $z \in \mathcal{V}_\mathcal{K}$ und interpoliere die Werte in jedem $K \in \mathcal{K}$ linear (bei Dreiecken) oder bilinear (bei Vierecken). Um effizienter rechnen zu können, wird zunächst eine Referenzzelle festgelegt.

Referenzdreieck

Bei einer Zerlegung in Dreiecke wird das Referenzdreieck

$$\hat{K} = \operatorname{conv}\left\{ \hat{z}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \hat{z}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \hat{z}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

betrachtet. Die affin-lineare Koordinatentransformation $\varphi_K \colon \hat{K} \longrightarrow \overline{K}$ sei durch

$$\varphi_K(\xi) = z_{K,0} + B_K \xi \text{ mit } B_k = \mathrm{D}\varphi_K$$

gegeben und liefert für jedes $K \in \mathcal{K}$ den Übergang vom Referenzdreieck $\hat{K}$ auf K. Auf dem Referenzdreieck $\hat{K}$ wird der Ansatzraum

$$\hat{V} = \operatorname{span}\{\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2\} = \mathbb{P}_1(\hat{K})$$

mit

$$\begin{aligned} \hat{\lambda}_0(\xi) &= 1 - \xi_1 - \xi_2 \\ \hat{\lambda}_1(\xi) &= \xi_1 \\ \hat{\lambda}_2(\xi) &= \xi_2 \end{aligned}$$

definiert. Nach Konstruktion gilt dann

$$\hat{\lambda}_i(\hat{z}_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Für den Ansatzraum auf dem Dreieck $K \in \mathcal{K}$ ergibt sich über die Koordinatentransformation φ_K

$$V_K = \operatorname{span}\{\hat{\lambda}_i \circ \varphi_K^{-1} \colon i = 0, 1, 2\} = \mathbb{P}_1(K).$$

Die Kroneckereigenschaft

$$\left(\hat{\lambda}_i \circ \varphi_K^{-1}\right)(z_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

bleibt so erhalten.

Referenzviereck

Bei einer Zerlegung in Vierecke werden auf die gleiche Weise ein Referenzviereck

$$\hat{K} = \operatorname{conv}\left\{\hat{z}_0 = \begin{pmatrix}0\\0\end{pmatrix}, \hat{z}_1 = \begin{pmatrix}1\\0\end{pmatrix}, \hat{z}_2 = \begin{pmatrix}1\\1\end{pmatrix}, \hat{z}_3 = \begin{pmatrix}0\\1\end{pmatrix}\right\}$$

sowie die zugehörige Koordinatentransformation $\varphi_K \colon \hat{K} \longrightarrow \overline{K}$

$$\varphi_K(\xi) = z_{K,0} + B_K(\xi) \text{ mit } B_K(\xi) = \mathrm{D}\varphi_K(\xi)$$

definiert. Mit den Ansatzfunktionen

$$\begin{aligned} \hat{\lambda}_{00}(\xi) &= (1 - \xi_1)(1 - \xi_2) \\ \hat{\lambda}_{01}(\xi) &= (1 - \xi_1)\xi_2 \\ \hat{\lambda}_{10}(\xi) &= \xi_1(1 - \xi_2) \\ \hat{\lambda}_{11}(\xi) &= \xi_1\xi_2 \end{aligned}$$

ergeben sich wie eben die lokalen Ansatzräume $\hat{V} = \operatorname{span}\{\hat{\lambda}_{00}, \hat{\lambda}_{01}, \hat{\lambda}_{10}, \hat{\lambda}_{11}\}$ und $V_K = \operatorname{span}\{\hat{\lambda}_{ij} \circ \varphi_K^{-1} \colon i, j = 0, 1\}$.

Legt man zu jedem $K \in \mathcal{K}$ mithilfe einer Abbildung

$$\mu_K \colon I_{\hat{K}} \longrightarrow I_h = \{1, \ldots, N\}$$

und
$$I_{\hat{K}} = \begin{cases} \{0,1,2\} & \text{Dreieck} \\ \{0,1,2,3\} & \text{Viereck} \end{cases}$$

die Zuordnung zwischen den Ecken von K und denen der Referenzzelle $\hat{K}$ fest, ergibt sich in K der Zusammenhang

$$\phi_k(x) = \hat{\lambda}_i \circ \varphi_K^{-1}(x)$$

mit $\mu_K(i) = k$. Folglich gilt in K

$$\mathrm{D}\phi_k(x) = \hat{\mathrm{D}}\hat{\lambda}_i(\hat{x})\mathrm{D}\varphi_K^{-1}(x) \text{ mit } \varphi_K(\hat{x}) = x$$

bzw.
$$\nabla\phi_k(x) = B_K^{-T}(\hat{x})\hat{\nabla}\hat{\lambda}_i(\hat{x}).$$

Finite-Elemente-Approximation

Zu einer gegebenen Zerlegung $\mathcal{K}$ mit der maximalen Kantenlänge $h = \max_{K\in\mathcal{K}} \operatorname{diam}(K)$ werden die Finite-Elemente-Ansatzräume

$$V_h = \{\phi_h \in C(\overline{\Omega})\colon \phi_h|_K \in V_K \ \forall K \in \mathcal{K}\}$$

und
$$V_h(u_\mathrm{D}) = \{\phi_h \in V_h\colon \phi_h(z) = u_\mathrm{D}(z), z \in \mathcal{V}_\mathcal{K} \cap \Gamma_\mathrm{D}\}$$
definiert. Aus der schwachen Formulierung (1) ergibt sich damit die Finite-Elemente-Formulierung: Bestimme $u_h \in V_h(u_\mathrm{D})$ mit

$$\sum_{K\in\mathcal{K}} \int_K \kappa\nabla u_h \cdot \nabla\phi_h \,\mathrm{d}x = \int_{\Gamma_\mathrm{N}} g_\mathrm{N}\phi_h \,\mathrm{d}a \quad \forall\phi_h \in V_h(0).$$

Unter Verwendung der Knotenbasis $\{\lambda_1, \dots, \lambda_N\}$ von V_h mit

$$\hat{\lambda}_n(\hat{z}_k) = \begin{cases} 1 & n = k \\ 0 & n \neq k \end{cases}$$

erhält man die äquivalente algebraische Formulierung: Sei $\mathcal{I} = \{1, \dots, N\}$ und $\mathcal{I}_\mathrm{D} = \{h \in \mathcal{I}\colon z_h \in \Gamma_\mathrm{D}\}$ die Indexmenge der Knoten, die auf dem Dirichletrand Γ_D liegen. Bestimme $\underline{u} \in \mathbb{R}^N$ mit $u_h = \sum \underline{u}_n\lambda_n \in V_h(u_\mathrm{D})$ und

$$\sum_{n=1}^N \underline{u}_n \sum_{K\in\mathcal{K}} \int_K \kappa\nabla\lambda_n \cdot \nabla\lambda_k \,\mathrm{d}x = \int_{\Gamma_\mathrm{N}} g_\mathrm{N}\lambda_k \,\mathrm{d}a, \quad \forall k \in \mathcal{I} \setminus \mathcal{I}_\mathrm{D}.$$

Wir definieren nun die Steifigkeitsmatrix $\underline{A} \in \mathbb{R}^{N\times N}$ und den Lastvektor $\underline{b} \in \mathbb{R}^N$ über

$$\underline{A}[k,n] = \begin{cases} \sum_{K\in\mathcal{K}} \int_K \kappa\nabla\lambda_n \cdot \nabla\lambda_k \,\mathrm{d}x, & k \notin \mathcal{I}_\mathrm{D} \\ 1, & n = k \in \mathcal{I}_\mathrm{D} \\ 0, & n \neq k \in \mathcal{I}_\mathrm{D} \end{cases}$$

und

$$\underline{b}[k] = \begin{cases} \int\limits_{\Gamma_\mathrm{N}} g_\mathrm{N} \lambda_k \, \mathrm{d}a & k \notin \mathcal{I}_\mathrm{D} \\ u_\mathrm{D}(z_k) & k \in \mathcal{I}_\mathrm{D} \end{cases} .$$

Damit ergibt sich das zu lösende lineare Gleichungssystem

$$\underline{A}\ \underline{u} = \underline{b}.$$

Nach Konstruktion gilt $\underline{u}_k = u_\mathrm{D}(z_k)$ für $k \in \mathcal{I}_\mathrm{D}$. Zur Bestimmung von A wird auf den Zellen der Zerlegung $\mathcal{K}$ und den Basisfunktionen λ_n mit Einschränkungen $\lambda_{K,n}$ auf K gerechnet:

$$\underline{A}[n,k] = \int\limits_\Omega \kappa \nabla \lambda_{K,k} \cdot \nabla \lambda_{K,n} \, \mathrm{d}x = \sum_{K \in \mathcal{K}} \int\limits_K \kappa \nabla \lambda_{K,k} \cdot \nabla \lambda_{K,n} \, \mathrm{d}x.$$

Den Vorgang des Zusammenfügens von A bezeichnet man als *Assemblieren*. Die Integration auf jeder Zelle wird über

$$\sum_{K \in \mathcal{K}} \int\limits_K \kappa \nabla \lambda_{K,k} \cdot \nabla \lambda_{K,n} \, \mathrm{d}x = \sum_{K \in \mathcal{K}} \int\limits_{\hat{K}} \det \mathrm{D}\varphi_K \kappa \nabla \lambda_{K,k} \circ \varphi_K \cdot \nabla \lambda_{K,n} \circ \varphi_K \, \mathrm{d}\hat{x}$$

auf die Referenzzelle transformiert. Dort wird mit Hilfe eines Quadraturverfahrens numerisch integriert. Falls $\kappa|_K$ für alle $K \in \mathcal{K}$ konstant ist, ergibt sich

$$\underline{A}[n,k] = \sum_{K \in \mathcal{K}} \sum_{\xi \in \Xi} w_\xi \det B_K(\xi) \kappa \circ \varphi_K B_K^{-T}(\xi) \hat{\nabla} \hat{\lambda}_{K,k}(\xi) \cdot B_K^{-T}(\xi) \hat{\nabla} \hat{\lambda}_{K,n}(\xi).$$

Für das Referenzdreieck werden

$$\Xi = \left\{ \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \right\} \text{ und } \omega_\xi = \frac{1}{2}$$

verwendet, für das Referenzviereck

$$\Xi = \left\{ \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}, \begin{pmatrix} 1-\alpha \\ \alpha \end{pmatrix}, \begin{pmatrix} \alpha \\ 1-\alpha \end{pmatrix}, \begin{pmatrix} 1-\alpha \\ 1-\alpha \end{pmatrix} \right\} \text{ mit } \alpha = \frac{\sqrt{2}-1}{2\sqrt{2}} \text{ und } \omega_\xi \ = \ \frac{1}{4}.$$

Algorithmus

A0) $\underline{A} = 0, \underline{b} = 0$

A1) $\forall K \in \mathcal{K}$:
$\underline{A} = \underline{A} + \underline{R}_K^T \underline{A}_K \underline{R}_K,\ \underline{b} = \underline{b} + \underline{R}_K^T \underline{b}_K$ d.h. $\underline{A}[r_{k,n}, r_{k,k}] = \underline{A}[r_{k,n}, r_{k,k}] + \underline{A}_K[n,k]$, $\underline{b}[r_{k,n}] = \underline{b}[r_{k,n}] + \underline{b}_K[n]$

A2) $\forall k \in \mathcal{I}_\mathrm{D} : \underline{b}[k] = u_\mathrm{D}(z_K),\ \underline{A}[k,k] = 1,\ \underline{A}[k,n] = 0$ für $n \neq k$

A3) $\forall k \in \mathcal{I}_\mathrm{D},\ n \neq k : \underline{b}[n] = \underline{b}[n] - \underline{A}[n,k] u_\mathrm{D}(z_k)$

(2.2) Lemma
Sei κ uniform symmetrisch positiv definit, d. h. $\kappa y \cdot y \geq \kappa_0 |y|^2\ \forall y \in \mathbb{R}^D$ mit $\kappa_0 > 0$ und sei $\Gamma_\mathrm{D} \cap \mathcal{V}_\langle \neq \emptyset$. Dann ist A regulär.

Beweis. Per Konstruktion gilt: $\underline{A}$ ist symmetrisch positiv definit, also zeige

$$\underline{A}\,\underline{v} \cdot \underline{v} = 0 \Leftrightarrow \underline{v} = 0$$

Sei also $\underline{A}\,\underline{v} \cdot \underline{v} = 0$. Außerdem gilt

$$\begin{aligned}
\underline{A}\,\underline{v} \cdot \underline{v} &= \sum_K \sum_\xi w_\xi |\det B_K| \kappa \nabla \underline{v}(\varphi_K(\xi)) \cdot \nabla \underline{v}(\varphi_K(\xi)) \\
&\geq \kappa_0 \sum_K \sum_\xi w_\xi \det B_K |\nabla \underline{v}(\varphi_K(\xi))|^2 \\
&= \kappa_0 \int\limits_\Omega |\nabla v|^2 \,\mathrm{d}x
\end{aligned}$$

$\Rightarrow \nabla \underline{v} = 0$ in Ω
$\Rightarrow \underline{v} \equiv \text{const}$ in $\overline{\Omega}$
Mit $z_K \in \Gamma_\mathrm{D} \cap \mathcal{V}_h \;\Rightarrow\; \underline{v}[k] = 0 \;\Rightarrow\; \underline{v}(z_k) = 0 \;\Rightarrow\; \underline{v} \equiv 0.$ □

Bemerkung
Im Folgenden sei $\underline{A}_0 \in \mathbb{R}^{N \times N}$ gegeben durch

$$\underline{A_0}[n,k] = \begin{cases} \underline{A}[n,k] & n,k \notin \mathcal{I}_\mathrm{D} \\ 1 & n = k \in \mathcal{I}_\mathrm{D} \\ 0 & n \neq k \in \mathcal{I}_\mathrm{D} \\ 0 & k \neq n \in \mathcal{I}_\mathrm{D} \end{cases}.$$

Nach Satz 2.2 ist A_0 positiv definit. Zerlegt man u in

$$\underline{u} = \underline{u_0} + \underline{u_\mathrm{D}}$$

ergibt sich das äquivalente Gleichungssystem

$$\underline{A_0 u_0} = \underline{b} - \underline{A}\underline{u_\mathrm{D}}.$$

Man kann also ohne Einschränkung annehmen, dass A positiv definit ist.

3 Lösen von dünn besetzten LGS

Ein Vorteil der finiten Elemente Methode ist, dass die auftretenden Matrizen in der Regel dünnbesetzt sind. Deswegen ist es von Vorteil spezielle Lösungsverfahren für dünnbesetzte Gleichungssysteme zu verwenden.

Betrachte eine Familie von linearen Gleichungssystemen $Au = b$ in $\mathbb{R}^N$. Sei $\mathcal{I} = \{1, ..., N\}$ die Index- Menge, $|\mathcal{I}| = N$ und $G = G_A = \{(n,k) \in \mathcal{I} \times \mathcal{I} : A[n,k] \neq 0\}$ der Matrixgraph.

(3.1) Definition
Eine Matrix heißt dünn besetzt , wenn $|G_A| = O(N)$. Sonst heißt sie voll besetzt.

(3.2) Lemma
Sei $\mathcal{I}_K = \{k \in \mathcal{I} : K \subset \mathit{supp}\phi_k\} = \{k \in \mathcal{I} : z_k \in \mathcal{V}_K\}$.

a) *Dann gilt für den Graph der FE- Matrix*

$$\begin{aligned} & G \subset \bigcup \mathcal{I}_K \times \mathcal{I}_K \\ \textit{und} \qquad & |G| \leq N(\max_{k \in \mathcal{I}} |\{K : z_k \in \mathcal{V}_K\}| + 1). \end{aligned}$$

   Bemerkung (D=2): Sei γ_0 der kleinste Innenwinkel aller Zellen K. Dann ist $|G| \leq (\frac{2\pi}{\gamma_0} + 1)N$.

b) *Das Assemblieren benötigt $O(N)$ Operationen.*

c) *Die Multiplikation Au benötigt $O(N)$ Operationen.*

d) *Das compressed row storage format benötigt $O(N)$ Speicher: Sei $M = |G|$ die Anzahl der Matrixeinträge ungleich Null. Sei außerdem*

$$\begin{gathered} a[m], m = 1..., M, \textit{ Vektor der Matrixeinträge } \neq 0 \\ c[m]\ in \{1, ..., N\} \textit{ Spaltenindex von } a[m] \\ d[n] \in \{1, ..., M+1\}, n = 1, N+1 \textit{ mit } d[1] = 1, \\ d[N+1] = M+1 \textit{ Diagonalpointer.} \end{gathered}$$

   Dann werden die Einträge der Matrix A folgendermaßen gespeichert:

$$\begin{aligned} A[n, c[m]] &= a[m] \textit{ für } m = d[n], ..., d[n+1]-1, n = 1, ..., N \\ \Rightarrow (Au)[n] &= \sum_{m=d[n]}^{d[n+1]-1} a[m]u[c[m]]. \end{aligned}$$

Assemblieren

$$\begin{aligned}\text{Sei } A_K[i,j] &= \int_K K\nabla(\hat{\lambda}_j \circ \varphi_K^{-1})\,\mathrm{d}x \\ \mathcal{V}_K &= \{z_{\mu_K(i)} : 1 \le i \le V_K\} \subset \mathcal{V}, V_K = |\mathcal{V}_K| \\ A &= \sum_K R_K^T A_K R_K \text{ mit } R_K \in \mathbb{R}^{V_K \times V}\end{aligned}$$

$$R_K[i,n] = \begin{cases} 1 & \mu_K(i) = n \\ 0 & \text{sonst} \end{cases}.$$

Nested Dissection

Sei $\overline{\Omega} = \overline{\bigcup_{p\in\mathcal{P}} \Omega^p}, \mathcal{P} = \{1, ..., P\}, \Omega^p \cap \Omega^q = \emptyset$ mit $p, q \in \mathcal{P}, p \neq q$ eine nichtüberlappende Gebietszerlegung $(\Omega^p \cap \Omega^q \neq \emptyset)$ mit Skeleton $\Gamma_{\mathcal{P}} = \Omega \cap (\cup \partial \Omega^p)$.
Seien $\mathcal{I}^p = \{n : z_n \in \overline{\Omega^p} \setminus \Gamma_{\mathcal{P}}\}$ Indizes zu inneren Knotenpunkten in Ω^p und $\mathcal{I}_\Gamma = \{n \in \mathcal{I} : z_n \in \Gamma_{\mathcal{P}}\}$ Interface- Indizes. Dann gilt: $A[n,k] = 0$ für $n \in \Omega^p, k \in \Omega^q (p \neq q)$.

$$A = \left(\begin{array}{ccc|c} A_{11} & & 0 & A_{\Gamma 1} \\ & \ddots & & \vdots \\ 0 & & A_{pp} & A_{\Gamma p} \\ \hline A_{1\Gamma} & \cdots & A_{p\Gamma} & A_{\Gamma\Gamma} \end{array}\right)$$

Aufwand in 2D: $O((\frac{N}{P})^2) + O((P(\frac{N}{P})^{\frac{1}{2}})^3)$, falls $p < \sqrt{N}$.
Das folgende Lemma zeigt, dass das Schurkomplement die Invertierbarkeit der ursprünglichen Matrix erbt.

(3.3) Lemma
Wenn A symmetrisch positiv definit ist, dann ist das Schurkomplement

$$S = A_{\Gamma\Gamma} - \sum_{p=1}^{P} A_{\Gamma p} A_{pp}^{-1} A_{p\Gamma}$$

ebenfalls symmetrisch positiv definit, und es gilt für $u = A^{-1}b$:

$$\begin{aligned} u_\Gamma &= S^{-1}(b_\Gamma - \sum A_{pp}^{-1} A_{p\Gamma} b_p), \\ u_p &= A_{pp}^{-1}(b_p - A_{p\Gamma} u_\Gamma). \end{aligned}$$

Beweis. Aus $A_{pp}u_p = b_p - A_{p\Gamma}u_\Gamma$ folgt

$$\begin{aligned} A_{\Gamma\Gamma}u_\Gamma &= b_\Gamma - \sum A_{\Gamma p}u_p = b_\Gamma - \sum A_{\Gamma p}A_{pp}^{-1}(b_p - A_{p\Gamma}u_\Gamma) \\ \Rightarrow S u_\Gamma &= b_\Gamma - \sum A_{\Gamma p}A_{pp}^{-1} b_p. \end{aligned}$$

Sei $v_\Gamma \neq 0$ und $v_p = -A_{pp}^{-1}A_{p\Gamma}v_\Gamma$

$$\begin{aligned} \Rightarrow S v_\Gamma \cdot v_\Gamma &= A_{\Gamma\Gamma}v_\Gamma \cdot v_\Gamma - \sum (A_{pp}^{-1}A_{p\Gamma}v_\Gamma)(A_{p\Gamma} \cdot v_\Gamma) \\ &= A_{\Gamma\Gamma}v_\Gamma \cdot v_\Gamma - \sum v_p(A_{pp}v_p) \\ &= Av \cdot v \end{aligned}$$

$\Rightarrow Sv_\Gamma \cdot v_\Gamma = Av \cdot v = 0$
$\Rightarrow v = 0 \;\Rightarrow\; v_\Gamma = 0$
$\Rightarrow$ S ist positiv definit. $\square$

Nun wähle

$$P = 2^S, \Omega^{0,1} = \Omega$$
$$\overline{\Omega}^{1,1} \cup \overline{\Omega}^{1,2} = \overline{\Omega}$$

rekursiv

$$\begin{aligned}\overline{\Omega}^{s,q} &= \overline{\Omega}^{s+1,2q-1} \cup \overline{\Omega}^{s+1,2q}, q = 1, ..., 2^S \\ \Omega^p &= \Omega^{S,p}, p = 1, ..., 2^S = P.\end{aligned}$$

Nested Dissection: Rekursive Anwendung von Lemma 3.3 für $s = S-1, S-2, ..., 0$.

Beispiele für Partitionierungsverfahren

Sobald die Finite Elemente Methode parallel ausgeführt wird, was in der Praxis der Regel entspricht, muss das zugrunde liegende Gebiet auf die einzelnen Prozesse verteilt werden. Dies geschieht mit Hilfe eines Partitionierungsverfahrens.

A) Rekursive Koordination - Bisektion (RCB)

Sei $z_K = \frac{1}{|K|} \sum\limits_{z \in \mathcal{V}_K} z \in \mathbb{R}^2$ Zellenmittelpunkt
Definiere
$z <_1 y \Leftrightarrow z[1] < y[1]$ oder $z[1] = y[1]$ und $z[2] < y[2])$
$z <_2 y \Leftrightarrow z[2] < y[2]$ oder $z[2] = y[2]$ und $z[1] < y[1])$

$s = 0$: $\mathcal{I}^0 = 1, ..., N, \mathcal{V}_\mathcal{K} = z_1, ..., z_N$ für $N = |\mathcal{K}|, z_n = z_{K_n}$
$\mathcal{I}^0 = \mathcal{I}^{1,1} \cup \mathcal{I}^{2,2}$ mit $z <_1 y$ für $z \in \mathcal{I}^{1,1}, y \in \mathcal{I}^{1,2}$

$s = 2$: $\mathcal{I}^0 = \mathcal{I}^{2,1} \cup \mathcal{I}^{2,2} \cup \mathcal{I}^{2,3} \cup \mathcal{I}^{2,4}$ mit $\mathcal{I}^{s-1,t} = \mathcal{I}^{s,2t-1} \cup \mathcal{I}^{s,2t}$ und $z <_2 y$

Fahre nun mit abwechselnder Sortierung rekursiv fort, bis das Gebiet weit genug zerlegt ist. Das Verfahren ist sehr schnell, allerdings nicht invariant gegenüber Gebietstransformationen und in ungünstigen Fällen wird das Interface $\mathcal{I}_\Gamma$ sehr groß.

B) Rekursiv Trägheits- Bisektion (RIB)
Definiere diskreten Laplace- Operator zum Graphen

$$\begin{aligned}\Gamma_\mathcal{K} &= \{(K, K') : \partial K \cap \partial K' = f \text{ gemeinsame Seite}\} \subset \mathcal{K} \times \mathcal{K} \\ d_K &= |\{K' \neq K : (K, K') \in \Gamma_\mathcal{K}\}|\end{aligned}$$

$$L[K, K'] = \begin{cases} d_K & K = K' \\ -1 & (K, K') \in \Gamma_\mathcal{K} \\ 0 & \text{sonst} \end{cases}$$

$L \cdot e = 0$ mit $e = (1, ..., 1)^\top \in \mathbb{R}^{N_\mathcal{K}}$ $\Rightarrow$ symmetrisch, positiv definit
Wir approximieren den Eigenvektor $w \in W = v \in R^{N_\mathcal{K}} : v \cdot e = 0$ zum kleinsten Eigenwert $\mu > 0$ von $Lw = \mu w$.
$w \cdot e = 0$ $\Rightarrow$ Definiere

$$\mathcal{K}^{1,1} = \{K : w[K] > 0\}, \mathcal{K}^{1,2} = \mathcal{K} \setminus \mathcal{K}^{1,1}.$$

Inverse Iteration zur Eigenwert- Berechnung:

I0) Wähle $w^0 \in \mathbb{R}^{N_\mathcal{K}} \setminus \{0, e\}, k = 0, \varepsilon > 0$

I1) Approximiere $w^k \approx (L|w)^{-1} w^{k-1}$ durch einige cg-Schritte.

I2) berechne $w^k := \frac{1}{|w^k|} w^k, \mu_k = L w^k \cdot w^k$

I3) falls $|\mu_k - \mu_{k-1}| < \varepsilon$ STOP

I4) gehe zu I1)

Dieses Verfahren ist invariant gegenüber Gebietstransformationen.

C) Space-Filling Curve

Definiere eine Kurve $\gamma : [a, b] \to \Omega$, die jede Zelle genau einmal durchläuft. Damit wird eine Nummerierung in $\mathcal{K}$ definiert. Setze

$$\mathcal{I}^q = \{K_{[(q-1)\frac{M}{p}]+1}, ..., K_{[q\frac{M}{p}]}\}.$$

Damit ensteht eine hierarchische Konstruktion, die auf unregelmäßige Gitter erweiterbar ist. Allerdings ist das Verfahren aufwendig und nicht optimal.

Netzgenerierung

Wenn ein Gebiet Ω gegeben ist, muss dieses erst in Zellen zerlegt werden, bevor die Finite Elemente Methode anwendbar ist. Da nicht jedes Gebiet eine regelmäßige Struktur hat, sind verschiedene Methoden zur Netzgenerierung nötig.

A) Strukturierte Gitter

1) Wähle Box $\Omega \supset [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]$.
Wähle $h > 0$ und Gitter $\Gamma_h = h\mathbb{Z}^2 \cap \Omega$.
Konstruiere regelmäßiges Vierecksgitter zu Γ_h.

2) Wähle weitere Punkte auf dem Rand $\partial\Omega$ und konstruiere (unregelmäßige) Vierecke am Rand.

3) Verschiebung einiger Knoten, um gleichmäßigere Vierecke zu erhalten (bzw. Einfügen von Dreiecken).

Dieses Verfahren hat den Vorteil, dass es schnell durchführbar ist, allerdings ist es nicht geeignet den Rand gut zu approximieren.

B) Advancing Front

1) Wähle (gleichmäßige) Punkte auf dem Rand.

2) Konstruiere eine Elementschicht entlang dem Rand, so dass ein neuer innerer Rand entsteht. Wiederhole diesen Prozess so lange die inneren Ränder nicht zusammenstoßen.
3) Verbinde die inneren Ränder durch Elemente.
4) Glätten der entstandenen Zellen.

Bei diesem Verfahren entstehen bessere Elemente am Rand, allerdings ist die Kollisionskontrolle im Inneren aufwendig.

C) Voronoi- Konstruktion

1) Wähle gleichmäßig verteilte Punkte z_n auf dem Rand und im Inneren.
2) Bilde Voronoi-Zellen

$$C_n = \{z \in \overline{\Omega} : |z - z_n| \leq |z - z_k| \; \forall k \neq n\} \text{ konvex!}$$

3) Bilde Kanten:

$$\mathcal{E} = \{(z_n, z_k) : n \neq k, C_n \cap C_k \text{ gemeinsame Seite}\}$$

4) Bilde zugehöriges Dreiecksnetz (Tetraedernetz).

Der Vorteil dieses Verfahrens sind die guten Netze die entstehen, allerdings ist man auf Dreiecke eingeschränkt.

Iterative Lösungsverfahren

Sei $B \in \mathbb{R}^{N \times N}$ ein Vorkonditionierer zu $A = L + D + R$ mit $B \approx A^{-1}$, z.B.

A) Jacobi, $B_{\text{Jac}} = D^{-1}$

B) Gauß-Seidel $B_{\text{GS}} = (L + D)^{-1}$

C) Sym. Gauß-Seidel $c = B_{\text{SGS}} r$ mit

$$\begin{aligned} c_1 &= B_{\text{GS}} r \\ r_1 &= r - A c_1 \\ c &= c_1 + B_{GS}^\top r_1 \\ \Rightarrow B_{\text{SGS}} &= B_{\text{GS}} + B_{\text{GS}}^\top - B_{\text{GS}}^\top A B_{\text{GS}} \end{aligned}$$

D) Parallele Vorkonditionierer zu $A = \sum_{p \in \mathcal{P}} A_p$

Wähle B^p zu $R_p A R_p^\top$. Definiere $c = B_{\mathcal{P}} r$ mit

$$\begin{aligned} c_p &= A_p^{-1} R_p r \\ c &= \sum R_p^\top c_p \end{aligned}$$

E) Zweigitter- Verfahren mit $A = A_1, A_0 = R_0 A_1 R_0^\top$ und
Restriktionen $R_0 \colon \mathbb{R}^N \to \mathbb{R}^{N_0}$ mit $N_0 < N$ und Prolongation $R_0^\top : \mathbb{R}^{N_0} \to \mathbb{R}^N$
Definiere $c = B_{\text{ZG}} r$ durch:
Glätten:

$$\begin{aligned} c_1 &= B_{\text{GS}} r \\ r_1 &= r - Ac_1 \\ r_0 &= R_0 r_1 \end{aligned}$$

Grobgitterkorrektur:

$$\begin{aligned} c_0 &= A_0^{-1} r_0 \\ c &= c_1 + R_0^\top c_0 \end{aligned}$$

Symmetrische Variante: Nachglätten mit $B_{\text{GS}}^\top$.

F) Mehrgitter: $A = A_L, A_{L-1} = R_{L-1} A_L R_{L-1}^\top, ...$ rekursiv glätten.

G) Least Squares:

1) Wähle $\varepsilon > 0$ als Abbruchbedingung und $\Theta > 0$ als Reduktionsfaktor. Sei u_0 gegeben. Setze

$$k = 0, r^0 = b - Au^0, \varphi_0 = |r^0|$$

2) falls $\varrho_k < \max\{\Theta\varrho_0, \varepsilon\}$ STOP
3) Wähle

$$\begin{aligned} c^{k-1} &= Br^k \\ r^{k+1} &= r^k - Ac^k \\ u^{k+1} &= u^k + c^k \end{aligned}$$

4) $k := k + 1$, gehe zu 2)

Es gilt

$$\begin{aligned} r^k &= (I - BA)^k r^0 \\ &\Rightarrow \varrho_k \leq \|I_N - BA\|^k \varrho_0 \\ &\Rightarrow \text{ Falls } \|I_N - BA\| < 1, \text{ so konvergiert LS} \end{aligned}$$

Falls $I_N - BA$ normal und $|\cdot| = |\cdot|_2$, so gilt

$$\frac{\varrho_k}{\varrho_0} \to \varrho(I_N - BA)$$

Krylov- Raum- Verfahren

Krylov- Raum- Verfahren sind eine Klasse von iterativen Verfahren zum Lösen von dünnbesetzten Gleichungssystemen.

Berechne eine ONB im Krylov- Raum $\mathcal{K}_k(BA, Br) = \{P(BA)Br \, : \, P \in P_{k-1}\}$, wobei A, B symmetrisch, positiv definit sind. Es gilt in der Energienorm $|u|_A = \sqrt{Au \cdot u}$:

1) Wähle $\varepsilon, \Theta > 0$, gegeben x^0
Setze

$$k = 0, r^0 = b - Ax^0, \varphi_0 = 1, p^0 = u^0$$

2) Falls $|r^k| < \max\{\Theta|r^0|, \varepsilon\}$ STOP

3) Setze

$$\begin{aligned}
c^k &= Br^k \\
\varrho_1 &= c \cdot r \\
p^k &:= \frac{\varrho_1}{\varrho_0} p^k \\
p^{k+1} &:= p^k + c^k \\
\varrho_0 &= \varrho_1 \\
q^{k+1} &= Ap^{k+1} \\
\alpha &= \frac{\varrho_0}{p^{k+1} \cdot q^{k+1}} \\
u^{k+1} &= u^k + \alpha_k p^{k+1} \\
r^{k+1} &= r^k - \alpha q^{k+1}
\end{aligned}$$

4) $k := k + 1$, gehe zu 2)

Es gilt in der Energienorm $|u|_A = \sqrt{Au \cdot u}$:

$$|u^k - u^0|_A \leq 2 \left(\frac{\sqrt{\varkappa} - 1}{\sqrt{\varkappa} + 1} \right)^k |u^1 - u^0|_A \text{ mit } \varkappa = \varkappa(BA).$$

Das Folgende wurde im SS'15 in der Vorlesung notiert nach Space-Filling Curve:

Subspace Correction Methode
Sei $A \in \mathbb{R}^{N \times N},\ N_p < N,\ p \in 1, ..., P,\ R_p \in \mathbb{R}^{N_p \times N},\ A_p = R_p A R_p^\top$.
Beispiel:
A ist FE-Matrix, sei weiter $N = |\mathcal{V}_\mathcal{K}|,\ V_h = span\{\lambda_1, ..., \lambda_N\}, \mathcal{V}_\mathcal{K} = \{z_1, ..., z_N\} = \cup\, \mathcal{V}_\mathcal{K}$. Sei zudem $\mathcal{K} = \bigcup\limits_{p \in \mathcal{P}} \mathcal{K}^p$ nicht überlappend und $\mathcal{V}^P = \bigcup\limits_{k \in \mathcal{K}^P} \mathcal{V}_K$ überlappend.

$$\begin{aligned}
R_P[n, k] &= \{z_{P,1}, ..., z_{P,N_P}\} \\
&= \begin{cases} 1 & z_{P,n} = z_K \\ 0 & \text{sonst} \end{cases}
\end{aligned}$$

Nun wähle $B_P \approx A_P^{-1}$.
Beispiel:

1) $B_P = A_P^{-1}$ LU-Zerlegung

2) $B_P = (\operatorname{diag} A_P)^{-1}$ Jacobi-Verfahren

3) $B_P = (\operatorname{diag} A_P + \ lower\ A_P)^{-1}$ Gauß-Seidel

Parallel Subspace Correction

P0) Wähle $u^0 \in \mathbb{R}^N$ $(u^0[n] = u_\mathrm{D}(z_n),\ z_n \in \Gamma_\mathrm{D})$
Berechne $r^0 = b - Au^0,$ setze $k = 0,$ und wähle $\varepsilon > 0,\ \theta > 0$

P1) Falls $|r^k| < \varepsilon$ STOP

P2) $\forall p:$ berechne

$$c_p = B_P R_P r^k \text{ und } c^k = \sum R_P^\top c_P$$

P3)

$$\begin{aligned} u^{k+1} &= u^k + \theta c^k \\ r^{k+1} &= r^k - \theta A c^k \end{aligned}$$

P4) Setze $k := k + 1,$ gehe zu P1).

Sucsessive Subspace Correction

P2)* Setze $\tilde{r} = r^k$ für $p = 1, ..., P:$ $c_P = B_P R_P \tilde{r}$ dann $\tilde{r} := \tilde{r} - AR_P^\top c_P$.

Symmetrisierung
Für ungerade k ersetze B_P durch $B_P^\top$.
$\Rightarrow$ falls A, B_P positiv definit: CG-Verfahren.

Poisson Problem

Sei $\Omega \subset \mathbb{R}^N$ beschränkt. Suche $u : \Omega \to \mathbb{R}$, sodass $u \in C^2(\Omega) \cap C(\overline{\Omega})$ und $\partial\Omega = \Gamma_\mathrm{D} \cup \Gamma_\mathrm{N}$ mit

$$\begin{aligned} -\Delta u &= f \text{ in } \Omega \\ u &= g \text{ auf } \Gamma_\mathrm{D} \\ \nabla u n &= h \text{ auf } \Gamma_\mathrm{N} \end{aligned} \tag{P}$$

wobei $f \in C(\Omega)$ und $g, h \in C(\partial\Omega)$ gegeben.
Spezialfall: Für $f = 0$ $\Rightarrow$ Laplace-Gleichung

Schwache Formulierung
$u \in L_1(\Omega)$ ist schwache Lösung von (P), falls

$$\begin{aligned}
\int_\Omega -\Delta u v \,\mathrm{d}x &= \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\partial\Omega} \nabla u n \cdot v \,\mathrm{d}a \\
&= \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x + \int_{\Gamma_\mathrm{N}} \nabla u n \cdot v \,\mathrm{d}a \\
&= \int_\Omega f \cdot v \,\mathrm{d}x
\end{aligned}$$

$\forall v \in V(0)$ wobei $V(0) = H^1_{0,\Gamma_\mathrm{D}} := \{v \in V : v = 0 \text{ auf } \Gamma_\mathrm{D}\}$ mit $v = H^1(\Omega) := \{v \in L_2(\Omega) : v$ besitzt schwache Ableitung $\partial_i v \in L_2(\Omega)\}$.
Suche $u \in V(g) := \{u \in V : u = g \text{ auf } \Gamma_\mathrm{D}\}, sodass$

$$\begin{aligned}
a(u, v) &:= \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x \\
&= \int_\Omega f v \,\mathrm{d}x + \int_{\Gamma_\mathrm{N}} \nabla u n \cdot v \,\mathrm{d}a \qquad \text{(VP)} \\
&:= \langle l, v \rangle \qquad \forall v \in V(0).
\end{aligned}$$

Galerkin-Verfahren
Wähle endlich dimensionalen Ansatzraum bzw. Testraum $V_h \subset V$ mi Knotenbasis $\{\lambda_z : z \in \mathcal{V}\}$, $\mathcal{V} = \{z_1, ..., z_n\}$. Dann ist

$$\begin{aligned}
v_h &= \sum_{z \in \mathcal{V}} \underline{v}_z \lambda_z, \\
a(u_h, v_h) &= \langle l, v_h \rangle \qquad \forall v_h \in V_h(0)
\end{aligned}$$

mit $u_h = g$ auf Γ_D und $u_z = g(z) \, \forall z \in \Gamma_\mathrm{D} \cap \mathcal{V}$.
Das heißt $\underline{u}_N^\top \underline{A}_N = \underline{l}^\top$ mit $\underline{A}_N = (a(\lambda_z, \lambda_{\overline{z}}))_{\overline{z}, z}$ und $\underline{l} = (\langle l, \lambda_{\overline{z}} \rangle)_{\overline{z}}$.

$$\underline{Au} = \begin{pmatrix} \underline{A}_N & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \underline{u}_N \\ \underline{u}_D \end{pmatrix} = \begin{pmatrix} \underline{l} \\ g(z_1^D) \\ \vdots \\ g(z_n^D) \end{pmatrix}$$

Ritz-Galerkin-Verfahren
Genau dann ist $u_h \in V_h$ Lösung von (VP), wenn sie das Energiefunktional

$$\begin{aligned}
\mathcal{E}(u_h) &= \frac{1}{2} a(u_h, u_h) - \langle l, u_h \rangle \\
&= \frac{1}{2} \underline{u}^\top \underline{Au} - \underline{l}^\top \underline{u} \qquad \text{mit } u_h = g \text{ auf } \Gamma_\mathrm{D}
\end{aligned}$$

minimiert, das heißt suche Nullstelle von

$$\begin{aligned}
\nabla \mathcal{E}(u_h) &= \underline{A}^\top \underline{u} - \underline{l} \overset{!}{=} 0 \\
\nabla^2 \mathcal{E}(u_h) &= \underline{A}^\top = A \qquad \text{symmetrisch positiv definit.}
\end{aligned}$$

Dann linearisieren:

$$\nabla\mathcal{E}(u^k) \approx \nabla\mathcal{E}(u^{k-1}) + \nabla^2\mathcal{E}(u^{k-1})\Delta w^k \stackrel{!}{=} 0$$

Im Newton-Schritt:

$$\Delta w^k = -(\nabla^2\mathcal{E}(u^{k-1}))^{-1}\nabla\mathcal{E}(u^{k-1})$$

Setze $u^k = u^{k-1} - \Delta w^k$
STOP: $\|\nabla\mathcal{E}(u^k)\| < \varepsilon$.

(3.4) Satz
Seien A, B_p symmetrisch positiv definit und sei $B \in \mathbb{R}^{N\times N}$ mit $u^{k+1} = u^k + \theta Br^k$, d.h.

$$\textit{additiv/parallel: } B = \sum R_p^\top B_p R_p$$
$$\textit{multiplikativ/successive: } I_N - BA = \prod_{p=1}^{P} (I_N - R_p^\top BR_0 a).$$

Dann existiert

$$\Theta_0 > 0 \ \textit{ mit } \ \varrho = ||I_N - \Theta BA||_A < 1 \ \textit{ für } \ \Theta \in (0, \Theta_0) \ \textit{ und es gilt}$$

$$|u - u^k|_A \leq \varrho^k |u - u^0|_A$$

mit $|y|_A = \sqrt{y^\top Ay}$, $||C||_A = \sup\limits_{y\neq 0} \dfrac{|C_y|_A}{|y|_A} = ||L^\top CL^{-T}||_2$ *und* $A = LL^\top$ *Cholesky-Zerlegung.*

Beweis.

$$|u - u^k|_A = |(I_N - \Theta BA)^k(u - u^0)|_A \leq ||I_N - \Theta BA||_A^k \, |u - u^0|$$

$$||I_N - \Theta BA||_A = ||I_N - \Theta L^\top BL||_2 = \sup_{\lambda\in\sigma(L^\top BL)} |1 - \Theta\lambda|$$

wobei gilt $\sigma(L^\top BL) = \sigma(BA) \in [\lambda_{min}, \lambda_{\max}] \subset [0, \infty]$
$\Rightarrow \ \Theta_{opt} = \dfrac{2}{\lambda_{min} + \lambda_{\max}}$
$\Rightarrow \ ||I_N - \Theta BA||_A < 1$ für $\Theta < \dfrac{1}{\lambda_{\max}}$. □

Krylovraum:
$(BA)u = Bb$ soll berechnet werden
$\Rightarrow$ berechne $u^k \in u^0$ und $span\{P(BA)Br^0 : \ P \in \mathbb{P}_{k-1}\}$
$\Rightarrow$ berechne optimales P

(3.5) Satz

a) *Seien A, B symmetrisch positiv definit und $\alpha y^\top Ay \leq y^\top ABAy \leq Cy^\top Ay$. Dann gilt für das cg-Verfahren:*

$$|u^k - u|_A \leq 2\left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\right)^k \ |u - u^0|_A$$

mit $\kappa = \dfrac{C}{\alpha}$

b) Seien A, B regulär und $\alpha y^\top y \leq y^\top BAy$, $||BA||_2 \leq C$. Dann gilt für GMRES

$$|u^k - u|_2 \leq \kappa_2(BA)(1-\kappa^2)^{\frac{k}{2}}|u - u^0|_2.$$

Multilevel-Vorkonditionierer
Sei $V_H \subset V_h$ ein Finite-Elemente-Raum in einer grÃ¶beren Triangulierung. Wir eliminieren alle Dirichlet-Freiheitsgrade
$\{\lambda_1, ..., \lambda_N\}$ Knotenbasis $V_h(0)$ $\quad N = \dim V_h(0)$
$\{\lambda_1^H, ..., \lambda_{N_0}^H\}$ Knotenbasis $V_H(0)$ $\quad N_0 = \dim V_H(0) < N$.
Es existiert ein $R_0 \in \mathbb{R}^{N_0 \times N}$ mit $\lambda_k^H = \sum_{n=1}^{H} R_0[k,n]\lambda_n$.
Wir haben $A \in \mathbb{R}^{N\times N}$ und $A_0 \in \mathbb{R}^{N_0 \times N}$ mit Einträgen

$$\begin{aligned} A[k,n] &= \int \kappa \cdot \nabla\lambda_n \cdot \nabla\lambda_k \,\mathrm{d}x =: a(\lambda_n, \lambda_k) \\ A_0[k,n] &= a(\lambda_n^H, \lambda_k^H) \end{aligned}$$

(3.6) Lemma
a) Es gilt für die so definierten Matrizen $A_0 = R_0 A R_0^\top$ (Galerkin-Produkt)

b) $P_0 = R_0^\top A_0^{-1} R_0 A$ (Galerkin-Projektion) Orthogonal-Projektion bzgl.

$$\langle y, z\rangle_A = y^\top A z.$$

Beweis. zu a)

$$\begin{aligned} A_0[k,n] = a(\lambda_n^H, \lambda_k^H) &= a\left(\sum_{n'} R_0[n,n']\,\lambda_{n'}, \sum_{k'} R_0[k,k']\,\lambda_{k'}\right) \\ &= \sum_{n',k'} R_0[n,n']\,R_0[k,k']\,a(\lambda_{n'}, \lambda_{k'}) \\ &= (R_0 A R_0^\top)[k,n] \end{aligned}$$

zu b)

$$\begin{aligned} P_0^2 = (R_0^\top A_0^{-1} R_0 A)^2 &= R_0^\top A_0^{-1} R_0 A R_0^\top A_0^{-1} R_0 A \\ &= R_0^\top A_0^{-1} R_0 A = P_0 \end{aligned}$$

Da $P_0^2 = P$: Projektion auf $P_0(\mathbb{R}^N) = R_0^\top(\mathbb{R}^{N_0})$
$(I_N - P_0)^2 = I_N - 2P_0 - P_0^2 = I_N - P_0$
Ist es auch Orthogonalprojektion bzgl- Skalarprodukt?

$$\begin{aligned} \langle P_0 v, (I_N - P_0)w\rangle_A &= v^\top P_0^\top A (I_N - P_0) w \\ &= v^\top (P_0^\top A - P_0^\top A P_0) w \end{aligned}$$

Nebenrechnung:

$$P_0^\top A P_0 = A R_0^\top A_0^{-1} R_0 A R_0^\top A_0^{-1} R A = P_0^\top A \qquad \square$$

Zweilevel-Verfahren

Z0) Wähle $u^0 \in \mathbb{R}^N$, berechne $r^0 = b - Au^0$, setze $k = 0$, wähle $\varepsilon > 0$ und $\Theta > 0$

Z1) Falls $|r^k| < \varepsilon$ STOP

Z2) $c_0^k = A_0^{-1} R_0 r^k$ (Grobgitterkorrektur)

Z3) $c^k = R_0^\top c_0^k + B(r^k - AR_0^\top c_0^k)$ (Glätten)

Z4) $u^{k+1} = u^k + c^k,\ r^{k+1} = r^k - Ac^k$ und setze $k = k + 1$, gehe zu Z1)

(3.7) Satz

a) Es gilt

$$u - u^{k+1} = (\mathcal{I}_\mathrm{N} - BA)(\mathcal{I}_\mathrm{N} - P_0)(u - u^k)$$

b) Sei A, B symmetrisch positiv definit, $||BA||_A \le 1$ und $v^\top Av \le Cv^\top ABAv$ für $P_0 v = 0$. Dann gilt

$$||(\mathcal{I}_\mathrm{N} - BA)(\mathcal{I}_\mathrm{N} - P_0)||_A \le \sqrt{1 - \frac{1}{c}}.$$

Bemerkung

$C > 0$ *ist unabhängig von* N, *wenn* $B \approx A^{-1}$ *auf* $\mathcal{N}(P_0) = \{v \in \mathbb{R}^N : P_0 v = 0\}$.

Beweis. $A = LL^\top$ Cholesky-Zerlegung , $w_1, ..., w_\mathrm{N}$ Orthonormalbasis zu Eigenwert μ_n von $(L^\top BL)w_n = \mu_n w_n > 0$.

$$\begin{aligned} 1 \ge ||BA||_A^2 &= \sup_{v \ne 0} \frac{v^\top ABABAv}{v^\top Av} = \sup_{w \ne 0} \frac{w^\top L^* BLw}{w^\top w} \\ &= \max \mu_n \quad \in (0, 1] \end{aligned}$$

$$\begin{aligned} \Rightarrow\ |(\mathcal{I}_\mathrm{N} - BA)v|_A^2 &= v^\top (A - 2ABA + ABABA)v \\ &= w^\top (\mathcal{I}_\mathrm{N} - 2L^\top BL + (L^\top BL)^2)w \\ &= \sum (w_n^\top w)^2 (1 - 2\mu_n + \mu_n^2) \\ &\le \sum w_n^\top w(1 - \mu_n) \\ &= w^\top (\mathcal{I}_\mathrm{N} - L^\top BL)w \\ &= v^\top (A - ABA)v \end{aligned}$$

Betrachte weiter

$$\begin{aligned} ||(\mathcal{I}_\mathrm{N} - BA)(\mathcal{I}_\mathrm{N} - P_0)||_A &= \sup_{w \ne 0} \frac{|(\mathcal{I}_\mathrm{N} - BA)(\mathcal{I}_\mathrm{N} - P_0)w|^2}{|w|_A^2} \\ &= \sup_{v = (1 - P_0)w} \frac{|(\mathcal{I}_\mathrm{N} - BA)v|_A^2}{|v|_A^2 + |v - w|_A^2} \\ &\le \sup_{P_0 v = 0,\ v \ne 0} \frac{|v|_A^2 - v^\top ABAv}{|v|_A^2} \\ &\le 1 - \frac{1}{c} \end{aligned}$$

Anwendung:
$B = \Theta \operatorname{diag}(A^{-1})$ mit $\Theta \in (0,1),\ M \in \mathbb{R}^{N\times N},\ M[n,k] = \int_\Omega \lambda_k \lambda_n \,\mathrm{d}x$. V_h ist ein Finite-Elemente-Raum zu regelmäßigen Gittern $h = \operatorname{diam} K,\ H = 2h$. Dann gilt:

a) $\inf\limits_{\phi_H \in V_H} ||\phi_h - \phi_H||_0 \le ch||\nabla\phi_h||_0 \qquad \forall\, \phi_h \in V_h,\ ||\phi_0|| = \sqrt{\int_\Omega \phi^2 \,\mathrm{d}x}$

b) $||\lambda_h||_0 \le c_1 h^{-1} ||\nabla\lambda_h||_0$
$||h^{-1}||_0 \le c_2 ||\lambda_h||_0$

$$\Rightarrow v^\top M v \le c_1 h^{-2} v^\top B^{-1} v \ \text{ und } \ h^2 v^\top B^{-1} v \le c_2 v^\top M v$$
$$\text{mit } \ B^{-1}[n,n] = A[n,n] = \int \kappa \nabla\lambda_n \cdot \nabla\lambda_n \,\mathrm{d}x$$

c) $\phi_h = \sum v[n]\lambda_n \ \Rightarrow cv^\top A v \le ||\nabla\phi_h||_0^2 \le C v^\top A v,$
$cv^\top M v \le ||\phi_h||^2 \le C v^\top M v$

d) $\phi_h = \sum v[n]\lambda_n,\ P_0 v = 0$

$$\begin{aligned}
\Rightarrow a(\phi_h,\phi_h) = v^\top A v &= \inf_{w_0 \in \mathbb{R}^{N_0}} v^\top A(v - R_0^\top w_0) \\
&= \inf_{w_0} \ v^\top A B^{\frac12} B^{-\frac12}(v - R_0^\top w_0) \\
&\le \sqrt{v^\top ABAv}\, \underbrace{\sqrt{\inf_{w_0} \ (v - R_0^\top w_0) B^{-1} (v - R_0^\top w_0)}}_{} \\
&\qquad\qquad \le Ch^{-1} \ \inf \ ||\phi_h - \phi_H||_0 \\
&\qquad\qquad \le C||\nabla\phi_H||_0 \\
&\qquad\qquad \le C v^\top A v
\end{aligned}$$
□

Multilevel-Verfahren
$V_0(0) \subset V_1(0) \subset \cdots \subset V_L(0),\ h_l = h_0 2^{-l},\ l = 0, ..., L$
$N_l = \dim V_l(0),\ A_l \in \mathbb{R}^{N_l \times N_l},\ R_l \in \mathbb{R}^{N_l \times N_{l+1}},\ A_l = R_l A_{l+1} R_l^\top$

M0) Wähle $u_L^0,\ r_L^0 = b_L - A_L u_L$

M1) Falls $|r_L^k| < \varepsilon$ STOP

M2) Setze $l = L,\ c_L = 0,\ r_L = r_L^k$

M3) Für $\nu = 1, ..., \nu_l:\ w_l = B_l r_l,\ c_l := c_l + w_l,\ r_l := r_l - A_l w_l$

M4) $r_{l-1} = R_{l-1} r_l$, setze $l := l - 1$

M5) Falls $l > 0$ gehe zu M3)

M6) $c_0 = A_0^{-1} r_0$

M7) Setze $l := l + 1,\ c_l = R_{l-1}^\top c_{l-1}$

M8) Glätten (wie (M3))

M9) Falls $l < L$ gehe zu (M7)

M10) $u_L^{k+1} = u_L^k + c_L,\ r_L^{k+1} = r_L^k - A_L c_L,\ k := k + 1$
$\rightarrow$ gehe zu (M1)

4 Gemischte und hybride Finite Elemente

Sei ein Differentialgleichungssystems gegeben, das von mehreren Variablen abhängt. Hier kann es sinnvoll sein, die einzelnen Unbekannten in verschiedenen Räumen zu diskretisieren. In diesem Fall spricht man von gemischten Finiten Elementen. Betrachte hierzu das Problem

$$\begin{aligned} \operatorname{div} q &= 0 \text{ in } \Omega \\ q &= -\kappa\nabla u \text{ in } \Omega \\ -q\cdot n &= g_{\mathrm{N}} \text{ auf } \Gamma_{\mathrm{N}} \\ u &= u_{\mathrm{D}} \text{ auf } \Gamma_{\mathrm{D}} \end{aligned}$$

wobei κ symmetrisch positiv definit ist. Es gilt

$$\int_\Omega \operatorname{div} q\phi \,\mathrm{d}x = 0 \qquad \text{für alle Testfunktionen } \phi : \Omega \to \mathbb{R}.$$

und

$$\int_\Omega \kappa^{-1}(q+\kappa\nabla u)\cdot\psi \,\mathrm{d}x = 0 \qquad \text{für alle Testfunktionen } \psi : \Omega \to \mathbb{R}^2$$

Mit dem Satz von Gauß folgt zudem für $u\psi$:

$$\int_{\partial\Omega} (u\psi)\cdot n \,\mathrm{d}a = \int_\Omega \operatorname{div}(u\psi)\,\mathrm{d}x = \int_\Omega \nabla u\cdot\psi\,\mathrm{d}x + \int_\Omega u\cdot\operatorname{div}\psi\,\mathrm{d}x.$$

Dies führt auf das Sattelpunktproblem: Bestimme (q,u) mit $q\cdot n = -g_{\mathrm{N}}$ auf Γ_{N} und

$$\begin{aligned} \int_\Omega \kappa^{-1} q\cdot\psi\,\mathrm{d}x - \int_\Omega u\operatorname{div}\psi\,\mathrm{d}x &= -\int_{\Gamma_{\mathrm{D}}} u_{\mathrm{D}}\psi\cdot n\,\mathrm{d}a \\ \int_\Omega \operatorname{div} q\phi\,\mathrm{d}x &= 0 \end{aligned}$$

für alle (ψ,ϕ) mit $\psi\cdot n = 0$ auf Γ_{N}.

Diskretisierung

Sei $\mathcal{K}$ eine Triangulierung und $\mathcal{F}$ die Menge aller Seiten $\partial K\cap\partial K'$ oder $\partial K\cap\partial\Omega$.
Wähle FE-Räume W_h, Q_h und

$$W_h(g) = \{\psi_h\in W_h : \int_F \psi_h\cdot n\,\mathrm{d}a = \int_F g\,\mathrm{d}a \quad \forall F\subset\Gamma_{\mathrm{N}}\}.$$

Bestimme $(q_h,u_h)\in W_h(-g_{\mathrm{N}})\times Q_h$ mit

$$\begin{aligned} &\int_\Omega \kappa^{-1} q_h\cdot\psi_h\,\mathrm{d}x - \int_\Omega u_h\operatorname{div}\psi_h\,\mathrm{d}x = -\int_{\Gamma_{\mathrm{D}}} u_{\mathrm{D}}\psi_h\cdot n\,\mathrm{d}a \\ &-\int_\Omega \operatorname{div} q_h\phi_h\,\mathrm{d}x = 0 \quad \forall(\psi_h,\phi_h)\in W_h(0)\times Q_h. \end{aligned}$$

Zu $\hat{K} = \mathrm{conv}\{\binom{0}{0}, \binom{1}{0}, \binom{0}{1}\}$ definiere $\hat{W} = \mathrm{span}\{\hat{\psi}_0, \hat{\psi}_1, \hat{\psi}_2\}$ und

$$\begin{aligned}
\hat{F}_0 &= \mathrm{conv}\left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}\\
\hat{F}_1 &= \mathrm{conv}\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}\\
\hat{F}_2 &= \mathrm{conv}\left\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\right\}
\end{aligned}$$

mit

$$\begin{aligned}
\hat{\psi}_0(\xi) &= \begin{pmatrix}\xi_1\\ \xi_2 - 1\end{pmatrix}\\
\hat{\psi}_1(\xi) &= \begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}\\
\hat{\psi}_2(\xi) &= \begin{pmatrix}\xi_1 - 1\\ \xi_2\end{pmatrix}.
\end{aligned}$$

Zu $\hat{K} = \mathrm{conv}\{\binom{0}{0}, \binom{1}{0}, \binom{1}{1}, \binom{0}{1}\}$ definiere $\hat{W} = \mathrm{span}\{\hat{\psi}_0, \hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3\}$ und

$$\begin{aligned}
\hat{F}_0 &= \mathrm{conv}\left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\}\\
\hat{F}_1 &= \mathrm{conv}\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}\\
\hat{F}_2 &= \mathrm{conv}\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}\\
\hat{F}_3 &= \mathrm{conv}\left\{\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}\right\}
\end{aligned}$$

mit

$$\begin{aligned}
\hat{\psi}_0(\xi) &= \begin{pmatrix}0\\ \xi_2 - 1\end{pmatrix}\\
\hat{\psi}_1(\xi) &= \begin{pmatrix}\xi_1\\ 0\end{pmatrix}\\
\hat{\psi}_2(\xi) &= \begin{pmatrix}0\\ \xi_2\end{pmatrix}\\
\hat{\psi}_3(\xi) &= \begin{pmatrix}\xi_1 - 1\\ 0\end{pmatrix}.
\end{aligned}$$

Es gilt

$$\int\limits_{\hat{F}_k} \hat{\psi} \cdot \hat{n}\, \mathrm{d}\hat{a} = \begin{cases} 1 & j = k \\ 0 & \text{sonst.} \end{cases}$$

(4.1) Lemma
Sei $\hat{n}$ Normale an $\partial\hat{K}$. Sei $\varphi_K : \hat{K} \to \bar{K}$ linear affin mit

$$\varphi_K(\xi) = z_{0,K} + B_K\xi,\; J_K = \det B_K > 0.$$

Dann ist $n = \frac{1}{|B_K^{-\top}\hat{n}|} B_K^{-\top}\hat{n}$ *Normale an* ∂K.

Beweis. $\hat{K} = \operatorname{conv}\hat{V}$, $\hat{n}\cdot(\hat{z}_j - \hat{z}_k) = 0$ mit $\hat{z}_j, \hat{z}_k \in \hat{V}$, $z_j = \varphi(\hat{z}_j)$, $z_k = \varphi(\hat{z}_k)$

$$\Rightarrow 0 = n\cdot(z_j - z_k) = n^\top(B_K\hat{z}_j - B_K\hat{z}_k) = (B_K^\top n)^\top(\hat{z}_j - \hat{z}_k)$$

$$\begin{aligned}&\Rightarrow \hat{n} = \lambda B_K^\top n \qquad \text{mit } \lambda\in\mathbb{R}\\ &\Leftrightarrow n = \frac{1}{\lambda}B_K^{-\top}\hat{n}\end{aligned}$$

□

Zu $\hat{\psi}$ mit $\hat{\psi}\cdot\hat{n}$ als Freiheitsgrad:
Definiere $\psi = J_K^{-1}B_K\hat{\psi}\circ\varphi_K \;\Rightarrow\; \psi\cdot n$ ist Freiheitsgrad ($\hat{\psi}\cdot\hat{n} = 0 \Leftrightarrow \psi\cdot n = 0$)
Definiere $W_h = \{\psi_h : \Omega\to\mathbb{R}^2 : \psi_h|_K = J_K^{-1}B_K\hat{\psi}\circ\varphi \text{ mit } \hat{\psi}\in\hat{W}\}$

$$\begin{aligned}W_h &= \operatorname{span}\{\psi_h : h = 1, ..., |\mathcal{F}|\}\\ \mathcal{F} &= \{F_1, F_2, ...\},\ Q_h = \prod_{K\in\mathcal{K}} P_0\\ \Omega &= \bigcup_{K\in\mathcal{K}} K \cup \bigcup_{F\in\mathcal{F}} F \cup \bigcup_{z\in\mathcal{V}}\{z\}\end{aligned}$$

Sei $\overline{F} = \operatorname{conv}\{z_j, z_k\}$ mit $z_j < z_k$, definiere Facenormale

$$u_F = \frac{1}{|z_j - z_k|}$$

mit $z_k[2] - z_j[2]$ und $z_k[1] - z_j[1]$.

$$\begin{aligned}s_{K,F} &= \operatorname{sgn}(u_F\cdot B_K^{-\top}\hat{u}) \quad \in\{\pm 1\}\\ \Rightarrow \psi(x) &= s_{K,F_n}J_K^{-1}B_K\hat{\psi}_{\hat{F}}(\xi) \quad \text{mit } x = \psi_K(\xi),\ F_n = \psi_K(\hat{F}).\end{aligned}$$

Für die Matrizen wähle Nummerierung:

$$\mathcal{F} = \{F_1, ..., F_{\mathrm{N}}\},\ \mathcal{K} = \{K_1, ..., K_M\}$$

und Basis

$$W_h = \operatorname{span}\{\psi_1, ..., \psi_{\mathrm{N}}\},\ Q_h = \{\mu_1, ..., \mu_M\}.$$

Es gilt für $\underline{A}\in\mathbb{R}^{N\times N}$ mit $\mu_m(x) = \begin{cases}1 & x\in K_m\\ 0 & \text{sonst}\end{cases}$:

$$\begin{aligned}\underline{A}[n,k] &= \int_\Omega \kappa^{-1}\psi_K\cdot\psi_n\,\mathrm{d}x\\ \underline{B}[m,k] &= -\int_\Omega \mu_m\operatorname{div}\psi_k\,\mathrm{d}x, \quad \underline{B}\in\mathbb{R}^{M\times N}\\ \underline{b}[k] &= -\int_{\Gamma_{\mathrm{D}}} u_{\mathrm{D}}\psi_k n\,\mathrm{d}a, \quad \underline{b}\in\mathbb{R}^N\end{aligned}$$

Randbedingungen: $\underline{W}(g) = \{\psi_h\in W_h : \psi_h[k] = \frac{1}{|F_h|}\int_{F_h} g\,\mathrm{d}a \quad \forall F_h\in\Gamma_{\mathrm{N}}\}$

(4.2) Lemma
Sei bei $(\underline{q},\underline{u}) \in \underline{W}(-g_N) \times \mathbb{R}^M$ *Sattelpunkt von* $L(\underline{q},\underline{u})$:

$$\begin{aligned} L(\underline{q},\underline{u}) &= \frac{1}{2}\underline{q}^\top \underline{A}\underline{q} - \underline{b}^\top \underline{q} + \underline{u}^\top \underline{B}\underline{q} \\ \textit{d.h. } L(\underline{q},\underline{v}) &\le L(\underline{q},\underline{u}) \le L(\underline{r},\underline{u}) \quad \forall (r,v) \in \underline{W}(-g_N) \times \mathbb{R}^M \end{aligned}$$

dann löst

$$q_K = \sum \underline{q}[m]\psi_m,\ u_h = \sum \underline{u}[m]\mu_m$$

das gemischte FE-Problem.

Beweis. i)

$$\begin{aligned} & L(\underline{q},\underline{u}+\underline{v}) \le L(\underline{q},\underline{u}) \quad \forall\, v \in \mathbb{R}^M \\ \Rightarrow\ & (\underline{u}+\underline{v})^\top \underline{B}\underline{q} \le \underline{u}^\top \underline{B}\underline{q} \\ \Rightarrow\ & \pm \underline{v}^\top \underline{B}\underline{q} \le 0 \\ \Rightarrow\ & \underline{B}\underline{q} = 0 \end{aligned}$$

ii) $L(\underline{q},\underline{u}) \le L(\underline{q}+t\cdot\underline{r},\underline{u}) \quad \forall\ \underline{r} \in \underline{W}(0)$
$\Rightarrow\ \frac{1}{2}\underline{q}^\top \underline{A}\underline{q} - \underline{b}^\top \underline{q} + \underline{u}^\top \underline{B}\underline{q} \le \frac{1}{2}(\underline{q}+t\cdot\underline{r})^\top\ \underline{A}(\underline{q}+t\cdot\underline{r}) - (\underline{q}+t\cdot\underline{r})^\top b + \underline{uB}(\underline{q}+t\cdot\underline{r})$

$$\begin{aligned} \Rightarrow\ 0 &\le t\underline{q}\underline{Ar} + \frac{1}{2}\underline{r}^\top \underline{Ar} + t\underline{r}^\top(\underline{B}^\top \underline{u} - b) \\ &= \frac{t^2}{2}\underline{r}^\top \underline{A} + t\cdot S \quad \text{ mit } S = \underline{r}^\top(\underline{A}\underline{q} + \underline{B}^\top \underline{u} - b) \end{aligned}$$

noch zu zeigen: $S = 0$
Annahme: $S \ne 0$:

$$\Rightarrow\ \underline{r} \ne 0\ \Rightarrow\ \underline{r}^\top \underline{Ar} \ne 0,\ A \text{ positiv definit.}$$

Wähle

$$t = -\frac{S}{\underline{r}^\top \underline{Ar}}\ \Rightarrow\ 0 \le -\frac{S^2}{2\underline{r}^\top \underline{Ar}} \quad ⚡$$

□

Optimierungstheorie
$\underline{W}(-g_\mathrm{N}) \subset \mathbb{R}^N$ linear affin $\Rightarrow$ konvex

$$L(\underline{q},\underline{u}) = \mathcal{E}(\underline{q}) + \underline{u}^\top \underline{B}\underline{q} \text{ und } \mathcal{E}(\underline{q}) = \frac{1}{2}\underline{q}^\top \underline{A}\underline{q} + b\underline{q}$$

Es gilt: $(\underline{q},\underline{u})$ ist Sattelpunkt von $L(.,.)\ \Rightarrow\ q \min \mathcal{E}(.)$ unter der Nebenbedingung $\underline{B}\underline{q} = 0$, $\underline{u}$ Lagrange-Multiplikator

(4.3) Satz
Sei $\operatorname{div} q = 0,\ q + \kappa\nabla u = 0 \in \Omega$,
$u = u_D$ *auf* Γ_D *mit* u_D *stückweise linear,*
$q \cdot n = -g_N$ *auf* Γ_N *mit* g_N *stückweise konstant auf* $\partial\Omega \supset F$.
Berechne "primale "Lösung $u_h \in V_h(u_D)$ *und "duale "Lösung* $q_h \in W_h(-g_N)$ *mit*

$\operatorname{div} q_h = 0$ *in* $K \; \forall\, K \in \mathcal{K}$.
Dann gilt für den Fehler

$$||q - q_h||_{\kappa^{-1}} \leq ||q_h + \kappa \nabla u_h||_{\kappa^{-1}}$$

mit $||\psi||^2_{\kappa^{-1}} = \int \kappa^{-1} \psi \cdot \psi \, \mathrm{d}x$.

Beweis.

$$\begin{aligned}
||q - q_h||^2_{\kappa^{-1}} = & \int\limits_{\Omega} \kappa^{-1}(q - q_h) \cdot (q + \kappa \nabla u - q_h - \kappa \nabla u_h) \, \mathrm{d}x \\
& + \int\limits_{\Omega} \kappa^{-1}(q - q_h) \cdot (\kappa \nabla u_h - \kappa \nabla u) \, \mathrm{d}x \\
= & - \int\limits_{\Omega} \kappa^{-1}(q - q_h) \cdot (q_h + \kappa \nabla u_h) \, \mathrm{d}x - \int\limits_{\Omega} (q - q_h) \cdot \nabla (u - u_h) \, \mathrm{d}x \\
\leq & \; ||q - q_h||_{\kappa^{-1}} ||q_h + \kappa \nabla u_h||_{\kappa^{-1}} - \sum_{K \in \mathcal{K}} \int\limits_{\Omega} \operatorname{div}(q - q_h)(u - u_h) \, \mathrm{d}x \\
& + \int\limits_{\partial K} (q - q_h) \cdot n (u - u_h) \, \mathrm{d}a \\
= & \; 0
\end{aligned}$$

□

Algorithmus

S1) Wähle $\underline{q}_{\mathrm{N}} \in \underline{W}(-g_{\mathrm{N}})$

S2) Bestimme $\underline{q}_0 \in \underline{W}(0)$ und $u \in \mathbb{R}^M$ mit

$$\begin{aligned}
\underline{A} q_0 + \underline{B}^\top \underline{u} &= \underline{b} - \underline{A} \underline{q}_{\mathrm{N}} \\
\overline{B} q_0 &= -\underline{B} \underline{q}_{\mathrm{N}}
\end{aligned}$$

S3) Setze $\underline{q} = \underline{q}_{\mathrm{N}} + q_0$

Hybridisierung

Definiere

$$\begin{aligned}
W_K &= \{\psi : K \to \mathbb{R}^2 : J_K^{-1} F_K \hat{\psi} \circ \varphi_K^{-1} = \psi|_K, \hat{\psi} \in \hat{W}\} \\
W_{\mathcal{K}} &= \prod_{K \in \mathcal{K}} W_K
\end{aligned}$$

Dann gilt

$$\psi_h \in W_h \;\Leftrightarrow\; \psi_h \in W_{\mathcal{K}} \text{ und } (\psi_K - \psi_{K_F}) n_F = 0 \quad \forall\, F \subset \Omega.$$

Zusätzlich

$$\psi_h \in W_h(-g_{\mathrm{N}}) \text{ und } (\psi_K \cdot n + g_{\mathrm{N}}) = 0 \quad \forall\, F \subset \Gamma_{\mathrm{N}}.$$

Berechne $q_h \in W_{\mathcal{K}}$ als Minimum von

$$\mathcal{E}(\psi_h) = \frac{1}{2} \int_{\Omega} \kappa^{-1} \psi_h \cdot \psi_h \,\mathrm{d}x - \int_{\Gamma_\mathrm{D}} u_\mathrm{D} \psi_h \cdot n \,\mathrm{d}a$$

unter der Nebenbedingung

$$\begin{aligned} \int_K \operatorname{div} \psi_h \,\mathrm{d}x &= 0 \quad \forall K \in \mathcal{K} \\ \Leftrightarrow \int_\Omega \operatorname{div} q_h \cdot \eta_h \,\mathrm{d}x &= 0 \quad \forall \eta_h \in Q_h \end{aligned}$$

$$\Rightarrow \sum_{K \in \mathcal{K}} \int_{\partial K} q_h \cdot n_K \cdot \mu_h \,\mathrm{d}a = - \int_{\Gamma_\mathrm{N}} g_\mathrm{N} \cdot \mu_h \,\mathrm{d}a \quad \forall\, \mu_h \in M_h(0) \;\text{ mit }\; M_h = \prod_{\mathcal{F}} P_0$$

$$M_h(u_\mathrm{D}) = \{\mu_h \in M_h : \mu_h = \frac{1}{|F|} \int_F u_\mathrm{D} \,\mathrm{d}a \quad \forall F \in \Gamma_\mathrm{D}\}$$

Dazu bestimme Sattelpunkt $(q_h, u_h, \mu_h) \in W_{\mathcal{K}} \times Q_h \times M_h(u_\mathrm{D})$ von

$$\begin{aligned} L(q_h, u_h, \mu_h) = \frac{1}{2} \int \kappa^{-1} q_h \cdot q_h \,\mathrm{d}x + \int_\Omega \operatorname{div}(q_h) u_h \,\mathrm{d}x - \int_{\Gamma_\mathrm{D}} \mu_h \cdot q_h \cdot n \,\mathrm{d}a \\ + \sum_{K \in \mathcal{K}} \int_{\partial K} q_h \cdot n_K \cdot \mu_h \,\mathrm{d}a - \int_{\Gamma_\mathrm{N}} g_\mathrm{N} \cdot q_h \,\mathrm{d}a \end{aligned}$$

Hybride Finite Elemente

Definiere

$$\begin{aligned} W_K &= \{\psi_K = J_K^{-1} B_K \hat{\psi} \circ \varphi_K^{-1} \in P_1(K) : \psi \in \hat{W} \;\text{ mit }\; \psi_K \cdot n|_F \in P_0(F) \;\text{ und }\; F \in \mathcal{F}_K\} \\ W_{\mathcal{K}} &= \prod W_K \\ V_K &= P_1(K) (\text{bzw. } Q_1) \quad V_{\mathcal{K}} = \prod V_K \quad V_h = V_{\mathcal{K}} \cap C^0(\overline{\Omega}) \\ Q_K &= P_0(K) \quad Q_h = \prod Q_K \subset \mathrm{L}_2(\Omega) \\ M_K &= \prod_{F \in \mathcal{F}_K} P_0(F) \\ \mathcal{F} &= \bigcup_{K \in \mathcal{K}} \mathcal{F}_K \end{aligned}$$

(4.4) Lemma

a) Sei $\psi_h \in W_{\mathcal{K}}$. Dann gilt:

$$\psi_h \in W_h \;\Leftrightarrow\; \int_\Omega \operatorname{div} \psi_h \cdot \phi_h \,\mathrm{d}x = - \int_\Omega \psi_h \cdot \nabla \phi_h \,\mathrm{d}x$$

$\forall\, \phi_h \in V_h(0)$ und $\Gamma_D = \partial\Omega$.

b) Sei $q_h \in W_h$ *mit* $\int_\Omega \operatorname{div} q_h \phi_h \,\mathrm{d}x = 0 \quad \forall \phi_h \in Q_h$. *Dann gilt*

$$\int_{\Gamma_{in}} q_h \cdot n \,\mathrm{da} = - \int_{\Gamma_{out}} q_h \cdot n \,\mathrm{da}$$

mit $\Gamma_{in} = \overline{\bigcup_{q_h \cdot n|_F < 0} F}$ *und* $\Gamma_{out} = \overline{\bigcup_{q_h \cdot n|_F > 0} F}$

c) Für $u_h \in V_h$ *und* $\lambda_h \in M_h$ *gelte*

$$\int_\Omega \nabla_h \cdot \psi_h \,\mathrm{d}x + \int_\Omega u_h \operatorname{div} \psi_h \,\mathrm{d}x = \sum_{K \in \mathcal{K}} \int_{\partial K} \lambda_h \psi_h \cdot n \,\mathrm{da} \quad \forall \psi_h \in W_{\mathcal{K}}$$

Dann gilt:

$$\int_F \lambda_h \,\mathrm{da} = \int_F u_h \,\mathrm{da} \quad \forall F \in \mathcal{F}.$$

Beweis. zu a) Für $\psi_h \in W_{\mathcal{K}}$ gilt:

$$\psi_h \in W_h \Leftrightarrow \psi_K \cdot n_K = \psi_{K_F} \cdot n_K \quad \forall F = \partial K \cup \partial K_F \subset \Omega$$

$$\begin{aligned}
\Rightarrow \; 0 = \int_F (\psi_K - \psi_{K_F}) \cdot n_f \phi_h \,\mathrm{da} &= \sum_K \int_{\partial K} \psi_K \cdot n_K \phi_h \,\mathrm{da} \\
&= \sum_K \int_K \operatorname{div}(\phi_h \psi_K) \,\mathrm{d}x \qquad \text{mit } \psi_K = \psi_h|_K \\
&= \int_\Omega (\operatorname{div} \psi_h \phi_h + \nabla \phi_h \cdot \psi_h) \,\mathrm{d}x
\end{aligned}$$

zu b) Es gilt:

$$\int_\Omega \operatorname{div} q_h \phi_h \,\mathrm{d}x = 0 \quad \phi_h \in Q_h.$$

$$\begin{aligned}
&\Rightarrow \; \forall K: \; 0 = \int_K \operatorname{div} q_h \,\mathrm{d}x = \int_{\partial K} q_h \cdot n_k \,\mathrm{da} \\
&\Rightarrow \; 0 = \sum_K \int_{\partial K} q_h \cdot n \,\mathrm{da} = \int_{\Gamma_{in}} q_h \cdot n \,\mathrm{da} = \int_{\Gamma_{out}} q_h \cdot n \,\mathrm{da}
\end{aligned}$$

zu c) Wähle ψ_K mit $\psi_K \cdot n_K|_F = 1, \; \psi_K \cdot n|_F = 0, \; F' \neq F$

$$\psi_h = \begin{cases} \psi_K & x \in K \\ 0 & \text{sonst} \end{cases}$$

$$\int_F \lambda_h \,\mathrm{d}a = \int_F \lambda_h \cdot \psi_K \cdot n_K \,\mathrm{d}a \overset{c)}{=} \int_K (\nabla u_h \cdot \psi_K + \operatorname{div} \psi_h \cdot n_h) \,\mathrm{d}x$$
$$= \int_K \operatorname{div}(u_h \cdot \psi_h) \,\mathrm{d}x = \int_{\partial K} u_h \cdot \psi_h \cdot n_K \,\mathrm{d}a$$
$$= \int_F n_h \,\mathrm{d}a \qquad \square$$

(4.5) Folgerung
Es ist äquivalent:

(1) $(g_h, u_h) \in W_h(-g_N) \times Q_h$ *löst*

(1.1)

$$\int_\Omega \kappa^{-1} q_h \cdot \psi_h \,\mathrm{d}x - \int_\Omega u_h \operatorname{div} \psi_h dx = \int_{\Gamma_D} u_D \psi_h \cdot n \,\mathrm{d}a$$

(1.2)

$$\int_\Omega \operatorname{div} q_h \phi_h \,\mathrm{d}x = 0 \quad \forall (\psi_h, \phi_h) \in W_h(0) \times Q_h$$

(2) $(q_h, u_h, \lambda_h) \in w_\mathcal{K} \times Q_h \times M_h(u_d)$ *löst*

(2.1)

$$\int_\Omega \kappa^{-1} q_h \cdot \psi_K \,\mathrm{d}x - \int_K u_h \operatorname{div} \psi_K \,\mathrm{d}x = -\int_{\partial K} \lambda_h \psi_K \cdot n_K \,\mathrm{d}a = 0 \quad \forall\, \psi_K \in W_K$$

(2.2)

$$\int_K \operatorname{div} q_h \,\mathrm{d}x = 0$$

(2.3)

$$\sum_K \int_{\partial K} q_h \cdot n \mu_h \,\mathrm{d}a = -\int_{\Gamma_N} g_N \mu_h \,\mathrm{d}a \quad \forall \mu_h(0)$$

Beweis. Es gilt folgenden Äquivalenz:

$$q_h \in w_\mathcal{K} \overset{(2.3)}{\Rightarrow} q_h \in W_h(-g_\mathrm{N}) \quad \text{und}$$
$$\psi_h \in w_h(0) \overset{(2.1)}{\Rightarrow} \sum_{K \in \mathcal{K}} \int_{\partial K} \lambda_h \cdot \psi_h \cdot n \,\mathrm{d}a = \int_{\Gamma_\mathrm{D}} u_\mathrm{D} \psi_h \cdot n \,\mathrm{d}a \qquad \square$$

Algebraische Formulierung

Wähle Basis

$$\begin{aligned} W_K &= \operatorname{span}\{\psi_F : F \in \mathcal{F}_K\} \\ Q_h &= \operatorname{span}\{\eta_K\} \\ M_h &= \operatorname{span}\{\lambda_F\} \end{aligned}$$

mit

$$\psi_F \cdot n|_{F'} = \begin{cases} 1 & F = F' \\ 0 & \text{sonst} \end{cases} \qquad \lambda_F|_{F'} = \begin{cases} 1 & F = F' \\ 0 & \text{sonst} \end{cases} \qquad \eta_K|_{K'} = \begin{cases} 1 & K = K' \\ 0 & \text{sonst} \end{cases}$$

Man kann für die algebraische Formulierung die folgenden Matrizen und Vektoren definieren:

$$\begin{aligned} \underline{A}_K[F,F'] &= \int_K \kappa^{-1}\psi_F \cdot \psi_{F'}\,\mathrm{d}x \\ \underline{B}_K[F,K] &= -\int_K \operatorname{div}\psi_F\,\mathrm{d}x \\ \underline{b}_K[F] &= \int_{\Gamma_\mathrm{N}\cup F} g_\mathrm{N}\,\mathrm{d}a \\ \underline{c}_K[F,F'] &= \int_{F'} \psi_F \cdot n\,\mathrm{d}a \end{aligned}$$

Einsetzen in (2.1)-(2.3) liefert:

$$\begin{aligned} q_K &= \sum \underline{q}_K[F]\psi_F \\ \lambda_K &= \sum \underline{\lambda}_K[F]\lambda_F \end{aligned}$$

Dies führt auf die Formulierung:

$$\Rightarrow \begin{pmatrix} \underline{A}_K & \underline{B}_K^\top \\ \underline{B}_K & 0 \end{pmatrix} \begin{pmatrix} \underline{q}_K \\ \underline{u}_K \end{pmatrix} = \begin{pmatrix} \underline{c}_K\,\underline{\lambda}_K \\ 0 \end{pmatrix} = \begin{pmatrix} \underline{c}_K \underline{R}_K \underline{\lambda}_h \\ 0 \end{pmatrix}$$

mit

$$\underline{\lambda}_K = \underline{R}_K\lambda_h \;\text{ und }\; \underline{R}_K[F,F'] = \begin{cases} 1 & F \neq F' \\ 0 & sonst, \end{cases}$$

also $F \in \mathcal{F}_K,\ F' \in \mathcal{F}$.

$$\Rightarrow \sum_K (\underline{R}_K\underline{\mu}_h)^\top \underline{c}_K \underline{q}_K = \sum_K (\underline{R}_K\underline{\mu}_h)^\top \underline{b}_K$$

(4.6) Satz
Es gilt folgendes Gleichungssystem:

$$\underline{\mu}^\top \underline{S} \cdot \underline{\lambda} = \underline{\mu}^\top \underline{b} \quad \forall\, \underline{\mu} \;\; \textit{mit} \;\; \underline{\mu}[F] = 0, \;\; \textit{also} \;\; F \subset \Gamma_\mathrm{D} \;\; \textit{und}$$

$$\underline{S} = \sum_K \begin{pmatrix} \underline{c}_K \underline{R}_K \\ 0 \end{pmatrix}^\top \begin{pmatrix} \underline{A}_K & \underline{B}_K^\top \\ \underline{B}_K & 0 \end{pmatrix}^{-1} \begin{pmatrix} \underline{c}_K \underline{R}_K \\ 0 \end{pmatrix}$$
$$\underline{b} = \sum_K \begin{pmatrix} \underline{c}_K \underline{R}_K^\top \\ 0 \end{pmatrix} \begin{pmatrix} \underline{b}_K \\ 0 \end{pmatrix}$$

Bemerkung: Falls gilt $\Gamma_\mathrm{D} \neq 0$ ist $\underline{S}$ symmetrisch positiv definit auf $M_h(0)$.

Mehrgitterverfahren für das hybride Problem

Fortsetzung von $\lambda_h \in M_h 0 \prod \mathbb{P}_0(F)$ nach $V_h^{NC} \subset \prod \mathbb{P}_1(K) \subset \mathrm{L}_2(\Omega)$. Definiere dazu

$$V_K^{NC} = \mathbb{P}_1(K) = \mathrm{span}\{u_F : F \in \mathcal{F}\}$$

mit

$$\int_{F'} u_F \,\mathrm{d}a = \begin{cases} 1 & F = F' \\ 0 & sonst \end{cases}$$

wobei gilt

$$\int_{F'} u_F \,\mathrm{d}a = \int_{F'} \lambda_F \,\mathrm{d}a = |F'| u_F(x_{F'}).$$

Definiere weiter für Level $l = 0, ...L : R_l \in \mathbb{R}^{N_l \times N_{l+1}}$ mit

$$\lambda_l[F] = \sum_{F' \in \mathcal{F}_{l+1}} R_l[F, F'] \lambda_{l+1}[F'] \Leftrightarrow \int_\Omega u_l \cdot \phi_l \,\mathrm{d}x = \int_\Omega u_{l+1} \cdot \phi_l \,\mathrm{d}x$$

$\forall \phi_l \in V_l^{NC} \subset \mathrm{L}_2(\Omega)$ mit $\underline{\lambda}_l = \underline{R}_l \underline{\lambda}_{l+1}$.
Konstruiere (M6) mit S_l, Glätter B_l, restriktion R_l, $I_l = R_l^\top$ und $V_l^{NC} \Rightarrow S_l \neq R_l S_{l+1} R_l^\top$.

Ausblick in die Finite-Elemente-Theorie

Seien

$$\kappa = \{2^{-l} h_0 : l \in \mathbb{N}_0\},\ V_h,\ W_h,\ Q_h \text{ FE-Räume zu } h \in \kappa$$
$$u_\mathrm{D} = u_{h_0}|_{\Gamma_\mathrm{D}}$$
$$g_\mathrm{N} = -q_{h_0} \cdot n|_{\Gamma_\mathrm{N}}.$$

$\forall\, u \in \mathrm{L}_2(\Omega)$ gilt:

$$\lim_{h \in \kappa} \inf_{u_l \in Q_l} ||u - u_l||_0 = 0$$

Fluss: $q(v) - \kappa \nabla v$ mit

$$||q|| = \int_\Omega \kappa^{-1} q \cdot q \,\mathrm{d}x$$

$$||v||_v = \left(\int_\Omega \kappa \nabla v \cdot \nabla v \,\mathrm{d}x \right)^{\frac{1}{2}} = ||q(v)||$$

Es gilt: $V(0) = \overline{\bigcup V_h(0)}$ bzgl. $||\cdot||$ Hilbertraum und $V(0) \subset \mathrm{L}_2(\Omega) = \overline{\bigcup\limits_{h\in\kappa} Q}$. Weiter gilt:

$$\begin{aligned} \underline{A}\underline{q} + \underline{B}^\top \underline{u} &= \underline{b} \\ \underline{B}\underline{u} &= 0 \end{aligned}$$

$$\begin{aligned} &\Longleftrightarrow \underline{q} \;\text{ minimiert }\; \frac{1}{2}\underline{q}^\top A\underline{q} - \underline{b}^\top \underline{q} \;\text{ u. d. NB }\; \underline{B}\underline{q} = 0 \\ \underline{S}\cdot\underline{u} = \underline{f} &\Longleftrightarrow \underline{u} \;\text{ minimiert }\; \frac{1}{2}\underline{u}^\top \underline{S}\,\underline{u} - \underline{f}^\top \underline{u}. \end{aligned}$$

(P) $u_h \in V(u_\mathrm{D})$ minimiert $J(u) = \frac{1}{2}\int \kappa\nabla u\cdot\nabla u\,\mathrm{d}x - \int\limits_{\Gamma_\mathrm{D}} g_\mathrm{D} u\,\mathrm{d}a$

(D) $(q_h, u_h) \in W(-g_\mathrm{N}) \times Q_h$ minimiert

$$\mathcal{E}(q_h) = \frac{1}{2}\int \kappa^{-1} q_h\cdot q_h\,\mathrm{d}x - \int\limits_{\Gamma_\mathrm{D}} u_\mathrm{D}\,\mathrm{d}a \;\text{ u. d. NB }\; \operatorname{div} q_h = 0.$$

Optimierungstheorie

(P) Minimiere $f(q,u) = \frac{1}{2}(q,q) - (g_\mathrm{N}, u)_{\Gamma_\mathrm{N}}$ auf $\mathcal{M} = \{(q,\psi) = -(\kappa\nabla u, \psi)\,\forall\,\psi$ und $(u,\mu)_{\Gamma_\mathrm{N}} = (u_\mathrm{D},\mu)_{\Gamma_\mathrm{D}}\,\forall\,\mu\}$.
Es gilt:

$$L(q,u,\psi,\mu) = f(q,u) - (g + \kappa\nabla u, \psi) + (u - u_\mathrm{D}, \mu)_{\Gamma_\mathrm{D}}.$$

(D) Maximiere $f^*(\phi,\mu) = \inf\limits_{(q,u)} L(q,u,\psi,\mu) = -\frac{1}{2}(\phi,\psi) - (u_\mathrm{D}, \psi\cdot u)_{\Gamma_\mathrm{D}}$ auf $\mathcal{N} = \{(\psi,\mu):\ \inf\ L(q,u,\psi,\mu)\} > -\infty = \{\operatorname{div}\psi = 0,\ \psi\cdot n = g_\mathrm{N}$ auf $\Gamma_\mathrm{N}\}$.

Konvexe Analysis

$$u \in V:\ J(u) = \begin{cases} \int \kappa\nabla u\cdot\nabla u\,\mathrm{d}x - \int\limits_{\Gamma_\mathrm{N}} g_\mathrm{N}\cdot u\,\mathrm{d}a & u = u_\mathrm{D} \text{ auf } \Gamma_\mathrm{D} \\ \infty & \text{sonst} \end{cases}$$

$$V^* \text{ bzgl. } \langle q,u\rangle = \int\limits_\Omega q\cdot\nabla u\,\mathrm{d}x$$

$$J^*(q) = \sup_{v\in V}(\langle q,u\rangle - J(u)) = \begin{cases} \int\limits_\Omega \kappa^{-1} q\cdot q\,\mathrm{d}x + \int\limits_{\Gamma_\mathrm{D}} u_\mathrm{D}\; q\cdot n\,\mathrm{d}a & div q = 0,\ q\cdot u = -g_\mathrm{N} \text{ auf } \Gamma_\mathrm{N} \\ \infty & \text{sonst.} \end{cases}$$

5 Finite Volumen für die lineare Transportgleichung

5.1 Die lineare Transportgleichung

Gegeben seien $\Omega \subset \mathbb{R}^D$ und der Fluss $q\colon \Omega \longrightarrow \mathbb{R}^D$ mit $\operatorname{div} q = 0$. Gesucht wird die Dichteverteilung $\rho\colon \Omega \times [0,T] \longrightarrow \mathbb{R}_{\geq 0}$ einer transportierten Substanz zu gegebenem Einfluss $\rho_{\text{in}}\colon \Gamma_{\text{in}} \times [0,T] \longrightarrow \mathbb{R}_{\geq 0}$ mit $\Gamma_{\text{in}} = \{z \in \partial\Omega\colon q(z)\cdot n(z) \leq 0\}$. Aus der Bilanzgleichung

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_Y \rho \,\mathrm{d}x + \int_{\partial Y} \rho q \cdot n \,\mathrm{d}a = 0 \quad \forall Y \subset \Omega$$

ergibt sich die lineare Transportgleichung

$$\partial_t \rho + \operatorname{div} \rho q = 0 \text{ in } \Omega \times [0,T]. \tag{2}$$

Mit der Randbedingung $\rho(x,t) = \rho_{\text{in}}(x,t)$ auf $\Gamma_{\text{in}} \times [0,T]$ und der Anfangsbedingung $\rho(x,0) = \rho_0(x)$ in Ω erhält man eine Anfangs-Randwertaufgabe. Im Folgenden bezeichnet $\rho(t)\colon [0,T] \longrightarrow \mathbb{R}_{\geq 0}$ mit

$$\rho(t)(x) = \rho(x,t) \quad \forall x \in \Omega$$

die Dichteverteilung zum Zeitpunkt t.

(5.1) Lemma
Für die Lösung der linearen Transportgleichung gilt die Massenbilanz:

$$\int_\Omega \rho(t)\,\mathrm{d}x = \int_\Omega \rho_0 \,\mathrm{d}x - \int_{\Gamma_{\text{in}}} \int_0^t \rho_{\text{in}}(t) q \cdot n \,\mathrm{d}t\ \mathrm{d}a - \int_{\Gamma_{\text{out}}} \int_0^t \rho(t) q\cdot n\,\mathrm{d}t\ \mathrm{d}a$$

Beweis.

$$\begin{aligned}\int_\Omega (\rho(t) - \rho_0)\,\mathrm{d}x &= \int_\Omega \int_0^t \partial_t \rho(s)\,\mathrm{d}s\ \mathrm{d}x \\ &= -\int_\Omega \int_0^t \operatorname{div}(\rho q)\,\mathrm{d}t\ \mathrm{d}x - \int_0^t \int_{\partial\Omega} \rho q\cdot n \,\mathrm{d}a\ \mathrm{d}t\end{aligned}$$ $\square$

Beispiel

A) Wandernde Wellen
Sei $q \equiv q_0 \in \mathbb{R}^D$, $n_0 \in \mathbb{R}^D$ und $a \in C^1(\mathbb{R})$ ein gegebenes Amplitudenprofil. Definiere

$$\rho(x,t) = a(ct - n_0\cdot x) \ \textit{mit}\ c = q_0 \cdot n_0$$

Dann gilt:

$$\begin{aligned}\partial_t \rho(x,t) &= c\, a'(ct - n_0 \cdot x) \\ \nabla \rho(x,t) &= -n_0\, a'(ct - n_0\cdot x) \\ \operatorname{div}(\rho(x,t) q_0) &= -n_0 \cdot q_0\, a'(ct - n_0 \cdot x)\end{aligned}$$

$$\Rightarrow\ \partial_t \rho + \operatorname{div}(\rho q) = 0$$

B) Für $q : \Omega \to \mathbb{R}^D$ gilt:

$$\operatorname{div} q = 0 \;\Rightarrow\; \operatorname{div}(\rho q) = \nabla\rho \cdot q.$$

Sei $\chi : [0,T] \to \Omega$ charakteristische Kurve mit $\chi(0) = x,\ \chi'(t) = q(\chi(t))$. Definiere außerdem $\rho(\chi(t),t) = \rho_0(\chi(0))$.

$$\begin{aligned}\Rightarrow\; 0 = \frac{d}{dt}(\rho(\chi(t),t)) &= \nabla\rho(\chi(t),t)\cdot\chi'(t) + \partial_t\rho(\chi(t),t)\\ &= (\partial_t\rho + \operatorname{div}(\rho q))(\chi(t),t)\end{aligned}$$

$\Rightarrow$ Lösung der Transportgleichung

C) Charakteristiken
Sei $x \in \Omega$. Auf Geraden der Form $\chi(t) = tq + x$ gilt

$$\mathrm{d}_t\rho(\chi(t),t) = \partial_t\rho(\chi(t),t) + q\cdot\nabla\rho(\chi(t),t) = \partial_t\rho(\chi(t),t) + \operatorname{div}(\rho(\chi(t),t)q) = 0,$$

das heißt, es gilt $\mathrm{d}_t\rho \equiv 0$ entlang von χ. Folglich ist $\rho(\chi(t),t) \equiv \rho_0(x)$ konstant. Solche Kurven χ werden als Charakteritiken bezeichnet.

Anwendungsbeispiel

Sei $\Omega = (0,1)^2$, $\Gamma_{\text{in}} = [0,1]\times\{1\}$ und $q = (0,-1)^\top$. Dann gilt

$$\rho(x_1,x_2,t) = \begin{cases}\rho(x_1,1,t-(1-x_2)), & t \geq 1-x_2\\ 0, & t < 1-x_2\end{cases}.$$

(5.2) Lemma

a) Es gilt für die Lösung ρ der linearen Transportgleichung

$$\int_\Omega \rho_0\phi(0)\,\mathrm{d}x = \int_0^T \left[-\int_\Omega (\rho\partial_t\phi + \rho q\cdot\nabla\phi)\,\mathrm{d}x + \int_{\Gamma_{\text{in}}} \rho_{\text{in}} q\cdot n\phi\,\mathrm{da} + \int_{\Gamma_{\text{out}}} \rho q\cdot n\phi\,\mathrm{da}\right]\mathrm{d}t$$

Für alle Testfunktionen $\phi\colon \Omega\times[0,T] \longrightarrow \mathbb{R}$ mit $\phi = 0$ auf $\partial\Omega$ und $\phi(T) = 0$.

b) Wenn die obere Gleichung gilt und ρ hinreichend glatt ist, dann ist ρ die Lösung der Transportgleichung.

Bemerkung

Wenn die obere Gleichung für ein $\rho \in L_1(\Omega\times(0,T))$ und alle ϕ mit zusätzlich $\phi|_{\Gamma_{\text{out}}} = 0$ gilt, dann heißt ρ schwache Lösung.

Beweis.

$$\begin{aligned}-\int_\Omega (\rho\phi)(0)\,\mathrm{d}x &= \int_\Omega \Big((\rho\phi)(T) - (\rho\phi)(0)\Big)\,\mathrm{d}x\\ &= \int_\Omega \int_0^T \partial_t(\rho\phi)\,\mathrm{d}t\ \mathrm{d}x\\ &= \int_0^T \int_\Omega (\partial_t\rho\phi + \rho\partial_t\phi)\,\mathrm{d}x\ \mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
-\int_{\Omega} \partial_t \rho \phi \,\mathrm{d}x &= \int_{\Omega} \operatorname{div}(\rho q) \phi \,\mathrm{d}x \\
&= -\int_{\Omega} \rho q \cdot \nabla \phi \,\mathrm{d}x + \int_{\partial\Omega} \rho q \cdot n \phi \,\mathrm{d}a
\end{aligned}
\qquad \square
$$

Das Riemann-Problem

Sei $\Omega = \mathbb{R}^D,\ q_0 \in \mathbb{R}^D,\ n_0 \in \mathbb{R}^D$ und

$$
\rho_0(x) = \begin{cases} \rho_-, & x \in \Omega_- := \{x \in \Omega : x \cdot n_0 < 0\} \\ \rho_+, & x \in \Omega_+ := \{x \in \Omega : x \cdot n_0 > 0\} \end{cases}
$$

$$
\Rightarrow \rho(x,t) = \begin{cases} \rho_-, & x \in Q_- := \{(x,t) \in \Omega \times [0,T] : (x - tq_0) \cdot n_0 < 0\} \\ \rho_+, & x \in Q_+ := \{(x,t) \in \Omega \times [0,T] : (x - tq_0) \cdot n_0 > 0\} \end{cases}
$$

ist schwache Lösung, denn für Testfunktionen $\phi \in C_0^\infty([-1,T] \times \mathbb{R}^D)$ gilt

$$
\begin{aligned}
&\int_{\Omega} \rho_0 \phi(0) \,\mathrm{d}x + \int_{Q_- \cap Q_+} \begin{pmatrix} \rho \\ \rho q_0 \end{pmatrix} \cdot \begin{pmatrix} \partial_t \\ \nabla \end{pmatrix} \phi \,\mathrm{d}x \,\mathrm{d}t \\
&= \int_{\partial Q_- \cap \partial Q_+} \rho_- \phi \begin{pmatrix} 1 \\ q_0 \end{pmatrix} \cdot \begin{pmatrix} -q_0 \cdot n_0 \\ n_0 \end{pmatrix} \mathrm{d}a \ \mathrm{d}t + \int_{\partial Q_- \cap \partial Q_+} \rho_+ \phi \begin{pmatrix} 1 \\ q_0 \end{pmatrix} \cdot \begin{pmatrix} -q_0 \cdot n_0 \\ n_0 \end{pmatrix} \mathrm{d}a \ \mathrm{d}t.
\end{aligned}
$$

5.2 Das Finite-Volumen-Verfahren

Notation:

$$
\begin{aligned}
(u,v)_{0,\Omega} &= \int_{\Omega} uv \,\mathrm{d}x \\
(u,v)_{0,\Gamma} &= \int_{\Gamma} uv \,\mathrm{d}a
\end{aligned}
$$

Diskretisierung im Raum

Seien $\overline{\Omega} = \bigcup \overline{K}$ eine Triangulierung von Ω und $\phi\colon \Omega \longrightarrow \mathbb{R}$ eine Testfunktion. Dann ergibt sich auf einer Zelle K die schwache Formulierung

$$
(\partial_t \rho, \phi)_{0,\Omega} = -(\operatorname{div}(\psi(\rho)), \phi)_{0,\Omega} = (\psi(\rho), \nabla\phi)_{0,\Omega} - \sum_K (\psi(\rho) \cdot n_K, \phi)_{0,\partial K}.
$$

Im Finite-Volumen-Verfahren wird auf jeder Zelle ein konstanter Ansatz gewählt. Anders als bei der Finite-Elemente-Methode verzichtet man auf eine Stetigkeitsforderung an den Elementgrenzen. Gesucht wird also eine zeitabhängige approximierte Dichteverteilung $\rho_h\colon [0,T] \longrightarrow Q_h$ mit $Q_h = \prod_{K \in \mathcal{K}} \mathbb{P}_0$. Aus der schwachen Formulierung ergibt sich die Forderung an ρ_h

$$
\int_{\Omega} \partial_t \rho_h \phi_h \,\mathrm{d}x = \int_{\Omega} \rho q \cdot \nabla \phi \,\mathrm{d}x - \sum_{\tau} \sum_{f \subset \partial\tau} \int_f \mathcal{F}^*_{\tau,f}(\rho_h) \cdot n_\tau \phi_h \,\mathrm{d}a
$$

$$\forall \phi_h \in Q_h = \operatorname{span}\{\eta_K : K \in \mathcal{K}\} \text{ mit } \eta_K(x) = \begin{cases} 1 & x \in K \\ 0 & \text{sonst} \end{cases}$$

und mit dem *numerischen Fluss*

$$\psi^*(\rho_h) = \begin{cases} \rho_K & q \cdot n_K > 0 \\ \rho_{K'} & q \cdot n_K < 0 \text{ und } F = \partial K \cap \partial K' \end{cases}.$$

Im Folgenden bezeichne $N = \dim Q_h = |\mathcal{K}|$ die Anzahl der Zellen der Triangulierung $\mathcal{K}$, $\underline{\rho} = (\rho_K)_{K \in \mathcal{K}} \in \mathbb{R}^N$ den gesuchten Lösungsvektor und

$$\underline{M}[K, K'] = (\eta_K, \eta_{K'})_{0,\Omega} = \operatorname{diag}(|K|)_{K \in \mathcal{K}} \in \mathbb{R}^{N \times N}$$

die Massenmatrix von $\mathcal{K}$. Mit der Flussmatrix

$$\underline{A}[K, K'] = \begin{cases} -\sum\limits_{F \text{ mit } q \cdot n_K|_F > 0} \int\limits_F \eta_K^2 q \cdot n_K \, \mathrm{d}a & K = K' \\ -\int\limits_F \eta_{K'} \, \eta_K q \cdot n_K \, \mathrm{d}a & F = \partial K \cap \partial K' \text{ und } q \cdot n_K|_F < 0 \\ 0 & \text{sonst} \end{cases}$$

und dem Lastvektor

$$\underline{b}[K] = -\int\limits_{\partial K \cap \Gamma_{\text{in}}} \rho_{\text{in}} \, \eta_K \, q \cdot n_K \, \mathrm{d}a$$

ergibt sich die zu lösende gewöhnliche Differentialgleichung

$$\underline{M} \, \partial_t \underline{\rho} = \underline{A}\underline{\rho} + \underline{b}.$$

$$\Rightarrow \underline{\rho}(t) = \exp(-t\underline{M}^{-1} \, \underline{A})(\underline{\rho}_0 + \int\limits_0^t \exp(-s \, \underline{M}^{-1} \, A)\underline{M}^{-1}\underline{b}(s) \, \mathrm{d}s)$$

$\Rightarrow$ Method of lines

Diskretisierung in der Zeit

Eine numerische Lösung der Differentialgleichung ist beispielweise mit den folgenden numerischen Verfahren denkbar. Dabei bezeichne $\underline{\rho}^n$ die näherungsweise Lösung zum Zeitpunkt $t_n = n\Delta t$.

A) Explizites Euler-Verfahren Es gilt

$$\exp(\xi) \approx 1 + \xi \text{ und } \underline{b} \equiv 0.$$

$$\begin{aligned} \underline{\rho}^n = \underline{\rho}^{n-1} + \Delta t \underline{M}^{-1} \underline{A} \underline{\rho}^{n-1} &= (I_N + \Delta t \underline{M}^{-1} \underline{A}) \underline{\rho}^{n-1} \\ &= (I_N + \Delta t \underline{M}^{-1} \underline{A})^n \underline{\rho}^0 \end{aligned}$$

Beobachtung:
Für den Spektralradius ρ gilt

$$\rho(\Delta t\underline{M}^{-1}\underline{A}) = \mathcal{O}(\Delta t)\mathcal{O}(h^{-2})\mathcal{O}(h) = \mathcal{O}\Big(\frac{\Delta t}{h}\Big),$$

d. h. $(\underline{\rho}^n)_n$ ist genau dann beschränkt, wenn

$$\rho(I_N + \Delta t\underline{M}^{-1}\underline{A}) < 1.$$

Daraus ergibt sich die sogenannte CFL-Bedingung $\mathcal{O}(\frac{\Delta t}{h}) = 1$ bzw. $\Delta t < Ch$.

B) Klassisches Runge-Kutta-Verfahren

$$\exp(\xi) \approx 1 + \xi + \frac{1}{2}\xi^2 + \frac{1}{3!}\xi^3 + \frac{1}{4!}\xi^4$$

Mit $\underline{b} = 0$ gilt:

$$\underline{\rho}^n = \underline{\rho}^{n-1} + \Delta t\underline{M}\underline{A}(\underline{\rho}^{n-1} + \frac{\Delta t}{2}\underline{M}^{-1}\underline{A}(\underline{\rho}^{n-1} + \frac{\Delta t}{3}\underline{M}^{-1}\underline{A}(\underline{\rho}^{n-1} + \frac{\Delta t}{4}\underline{M}^{-1}\underline{A}\underline{\rho}^{n-1})))$$

C) Implizites Euler-Verfahren
Mit

$$\exp(\xi) = (1-\xi)^{-1}$$

gilt

$$\begin{aligned} &\underline{\rho}^n = \underline{\rho}^{n-1} + \Delta t\underline{M}^{-1}(\underline{A}\underline{\rho}^n + \underline{b}^n) \\ \Rightarrow\quad &\underline{\rho}^n = (I_N - \Delta t\underline{M}^{-1}\underline{A})^{-1}(\underline{\rho}^{n-1} + \Delta t\underline{M}^{-1}b^n) = (\underline{M} - \Delta t\underline{A})^{-1}(\underline{M}\underline{\rho}^{n-1} + \Delta t\underline{b}^n). \end{aligned}$$

stabil ohne CFL-Bedingung

D) Implizite Mittelpunktsregel
Mit

$$\exp(\xi) = \frac{1+\frac{1}{2}\xi}{1-\frac{1}{2}\xi}$$

gilt

$$\begin{aligned} \underline{\rho}^n &= \underline{\rho}^{n-1} + \Delta t\underline{M}^{-1}(\underline{A}(\frac{1}{2}\underline{\rho}^n + \frac{1}{2}\underline{\rho}^{n-1}) + \frac{1}{2}(\underline{b}^{n-1} + \underline{b}^n)) \\ &= (\underline{M} - \frac{\Delta t}{2}\underline{A})^{-1}((\underline{M} + \frac{\Delta t}{2}\underline{A})\underline{\rho}^{n-1} + \frac{1}{2}\underline{b}^{n-1} + \frac{1}{2}\underline{b}^n) \end{aligned}$$

Hier tritt folglich keine CFL-Bedingung auf. Das Verfahren ist reversibel.

E) Krylovraum-Verfahren

$$\exp(\xi) = P(\xi) \text{ mit } P \in \mathbb{P}_{m-1} \text{ und } b^n = 0.$$

Approximiere $\exp(\Delta t\underline{M}^{-1}\underline{A})\underline{\rho}^{n-1}$ im Krylovraum

$$\{P(\underline{M}^{-1}\underline{A})\underline{\rho}^{n-1} : P \in \mathbb{P}_{m-1}\}.$$

Berechne hierzu eine Orthonormalbasis $\{\underline{v}_1, \ldots, \underline{v}_m\}$ von Krylovraum bezüglich des Skalarprodukts $(\underline{v}, \underline{\widetilde{v}})_{\underline{M}} = \underline{v}^T \underline{M}\underline{\widetilde{v}}$. Mit $\underline{V}_m = (\underline{v}_1 \mid \ldots \mid \underline{v}_m)$ und der Galerkin-Projektion $\underline{A}_m = \underline{V}_m^T \underline{A}\underline{V}_m \in \mathbb{R}^{m\times m}$ und setze

$$\underline{\rho}^n = \underline{V}_m \exp(\Delta t\, A_m)\underline{V}^T \underline{\rho}^{n-1}.$$

Bemerkung
Die CFL-Bedingung erfordert wieder $m = \mathcal{O}(\frac{\Delta t}{h})$.

Wiederholung:
Eine lineare ODE $\dot{u} = Au$ in $\mathbb{R}^N$ ist stabil (d.h. $|u(t) \le c \;\forall t > 0$), wenn für alle Eigenwerte λ von A gilt: $Re\,\lambda \le 0$ und falls $Re\,\lambda = 0$, dann sind geometrische und algebraische Vielfache gleich. Falls $A = -A^T$ schiefsymmetrisch, dann gilt $|u(t)|_2 = |u(0)|_2$, ODE reversibel.
Ein Runge-Kutta-Verfahren heißt A-stabil, wenn die rationale Funktion mit

$$u^n = R(\Delta t A)u^{n-1} : \{\xi \in \mathbb{C} : |R(\xi)| \le 1\} \subset \{\xi \in \mathbb{C} : |\exp(\xi)| \le 1\} = \{\xi \in \mathbb{C} : Re\,\xi \le 0\}$$

gilt.
Wenn zusätzlich gilt $R(\infty) = 0$: L-stabil.
Wenn zusätzlich gilt $R(\xi)R(\xi)^{-1} = 1$: reversibel.

(5.3) Lemma
Sei $q = q_h \in W_h,\ \operatorname{div} q = 0$ *und* $\rho_0 \in Q_h$. *Dann gilt:*

a) Diskrete Massenbilanz:

$$\int_\Omega \rho_h(t)\,\mathrm{d}x = \int_\Omega \rho_0 - \int_0^t \left(\int_{\Gamma_{\text{in}}} \rho_{\text{in}} q \cdot n \,\mathrm{da} + \int_{\Gamma_{\text{out}}} \rho_h q \cdot n\,\mathrm{da}\right)\mathrm{d}\widetilde{t}$$

b) Wenn zusätzlich $\rho_0 \ge 0$ *und* $\rho_{\text{in}} \ge 0$ *gilt, dann ist auch* $\rho_h \ge 0$.

Beweis. zu a)

$$\begin{aligned}\int_\Omega (\rho_h(t) - \rho_0)\,\mathrm{d}x &= \int_0^t \int_\Omega \partial_t \rho_h(\widetilde{t})\,\mathrm{d}x\,\mathrm{d}\widetilde{t} \\ &= -\int_0^t \sum_K \left(\int_{\partial_K \setminus \Gamma_{\text{in}}} \psi^*(\rho_h)\cdot n\,\mathrm{da} + (\rho_{\text{in}} q\cdot n, 1)_{\Gamma_{\text{in}}}\right)\mathrm{d}\widetilde{t} \\ &= -\int_0^t \left(\int_{\Gamma_{\text{in}}} \rho_h q\cdot n\,\mathrm{da} + \int_{\Gamma_{\text{in}}} q\cdot n\,\mathrm{da}\right)\mathrm{d}\widetilde{t} + 0\end{aligned}$$

zu b) Sei $\rho_h(\widetilde{t}) \geq 0$ für $(0, t)$ und $\rho_K = \rho_h|_K$, dann gilt

$$(\partial_t \rho_K(t), 1)_K = -\sum_{F \in \mathcal{F}_K^{\text{in}}} (\rho_{K_F} q \cdot n, 1)_F - \sum_{F \in \mathcal{F}_K^{\text{out}}} (\rho_K q \cdot n, 1)_F - (\rho_{\text{in}} q \cdot n, 1)_{\Gamma_{\text{in}} \cap \partial K}$$

$$\Rightarrow \rho_K(t) = 0 \Rightarrow \partial_t \rho_K(t) \geq 0.$$ □

(5.4) Lemma
Sei $\operatorname{div} q = 0$.

a) Es gilt die Energiebilanz

$$\int_\Omega |\rho(t)|^2 \,\mathrm{d}x = \int_\Omega |\rho(0)|^2 \,\mathrm{d}x + \int_0^T \left[\int_{\Gamma_{in}} |\rho_{\text{in}}(t)|^2 \, |q \cdot n| \,\mathrm{d}a - \int_{\Gamma_{\text{out}}} |\rho(t)|^2 \, |q \cdot n| \,\mathrm{d}a \right] \mathrm{d}t.$$

b) Für die Finite-Volumen-Lösung gilt

$$\begin{aligned} \int_\Omega |\rho_h(t)|^2 \,\mathrm{d}x = \int_\Omega |\rho(0)|^2 \,\mathrm{d}x + \int_0^T \Bigg[\int_{\Gamma_{\text{in}}} |\rho_{\text{in}}(t)|^2 \, |q \cdot n| \,\mathrm{d}a - \int_{\Gamma_{\text{out}}} |\rho_h(t)|^2 \, |q \cdot n| \,\mathrm{d}a \\ - \sum_{F \in \mathcal{F} \cap \Omega} \int_F |\rho_K - \rho_{K_F}|^2 |q \cdot n| \,\mathrm{d}a - \int_{\Gamma_{\text{in}}} |\rho_h - \rho_{\text{in}}|^2 |q \cdot n| \,\mathrm{d}a \Bigg] \mathrm{d}t. \end{aligned}$$

Beweis. zu a)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega |\rho(t)|^2 \,\mathrm{d}x &= 2 \int_\Omega \rho(t) \partial_t \rho(t) \,\mathrm{d}x = -2 \int_\Omega \rho \operatorname{div}(\rho q) \,\mathrm{d}x \\ &= -2 \int_\Omega \operatorname{div}(\rho^2 q) \,\mathrm{d}x = -\int_{\partial\Omega} \rho^2 q \cdot n \,\mathrm{d}x \end{aligned}$$

Mit $q = 0$ folgt:

$$\begin{aligned} 0 = \int_K \operatorname{div}(\rho_h q) \rho_h \,\mathrm{d}x &= \frac{1}{2} \int_K \operatorname{div}(\rho_h^2 q) \,\mathrm{d}x \\ &= \frac{1}{2} \int_{\partial_K} \rho_h^2 q \cdot n \,\mathrm{d}a \end{aligned}$$

zu b)

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_{\Omega}|\rho_h(t)|^2\,\mathrm{d}x &= 2\int\limits_{\Omega}\rho_h\partial_t\rho_h\,\mathrm{d}x\\
&= -2\sum_K\left(\Psi^*(\rho_h)\cdot n,\rho_h\right)_{\partial K\setminus\Gamma_{\text{in}}} - 2\left(\Psi(\rho_{\text{in}})\cdot n,\rho_h\right)_{\Gamma_{\text{in}}}\\
&= -2\sum_K\left(\Psi^*(\rho_h)\cdot n,\rho_h\right)_{\partial K\setminus\Gamma_{\text{in}}} - 2\int\limits_{\Gamma_{\text{out}}}\rho_h^2 q\cdot n\,\mathrm{da} + 2\int\rho_h\rho_{\text{in}}|q\cdot n|\,\mathrm{da}\\
&\quad+\sum_K\int\limits_{\partial K}\rho_h^2 q\cdot n\,\mathrm{da}\\
&= \sum_K\left((\Psi(\rho_h)-2\Psi^*(\rho_h))\cdot n,\rho_h\right)_{\partial K\setminus\partial\Omega} - \int\limits_{\Gamma_{\text{out}}}\rho_h^2|q\cdot n|\,\mathrm{da}\\
&\quad-\int\limits_{\Gamma_{\text{in}}}\rho_{\text{in}}^2|q\cdot n|\,\mathrm{da} - \int\limits_{\Gamma_{\text{in}}}\left(\rho_{\text{in}}-\rho_h\right)^2|q\cdot n|\,\mathrm{da}
\end{aligned}
$$

Mit $q\cdot n_K > 0$ folgt:

$$
\int\limits_F\left[\left(\rho_K-2\rho_K\right)q\cdot n_K\rho_K + \left(\rho_{K_F}-2\rho_K\right)q\cdot n_{K_F}\rho_{K_F}\right]\,\mathrm{da} = \int\limits_F|q\cdot n_K|\left(\rho_K-\rho_{K_F}\right)^2\,\mathrm{da}
$$

□

Bemerkung

A) Der "upwind flux" Ψ^* ist massenerhaltend, aber nicht energieerhaltend.

B) Der "central flux" $\Psi^c = \frac{1}{2}\left(\Psi(\rho_K)+\Psi(\rho_{K_F})\right)$ ist energieerhaltend, aber nicht monoton.

5.3 Discontinuous-Galerkin-Verfahren

Sei $Q_h = \prod\mathbb{P}_k(K)$. Bestimme $\rho_h\colon [0,T]\to Q_h$ mit

$$
\begin{aligned}
\left(\partial_t\rho_h,\phi_h\right)_{0,\Omega} &= -\left(\Psi(\rho_h),\nabla\phi_h\right)_{0,\Omega} - \sum\left(\Psi^*(\rho_h),\phi_h\right)_{0,\partial K\setminus\Gamma_{\text{in}}} - \left(\Psi(\rho_{\text{in}}),\phi_h\right)_{\Gamma_{\text{in}}}\\
&= \sum_K\left(\operatorname{div}\Psi(\rho_h),\phi_h\right)_{0,K} + \left(\Psi(\rho_h)-\Psi^*(\rho_h),\phi_h\right)_{0,\partial K\setminus\Gamma_{\text{in}}} + \left(\Psi(\rho_h)-\Psi(\rho_{\text{in}}),\phi_h\right)_{\Gamma_{\text{in}}}\\
&=: \left(A_h\rho_h,\phi_h\right)_{0,\Omega} + \left(b_h,\phi_h\right)
\end{aligned}
$$

Bemerkung

Es gilt:

$$
\left(A_h\rho_h,\rho_h\right)_{0,\Omega} = -\int\limits_{\partial\Omega}|\rho_h|^2|q\cdot n|\,\mathrm{da} - \frac{1}{2}\sum_{F\in\mathcal{F}\cap\Omega}\int\limits_F|\rho_K-\rho_{K_F}|^2\,|q\cdot n|\,\mathrm{da}
$$

$\Rightarrow$ *$-A$ ist positiv definit.*

Fehlerkontrolle

Wie können wir berechnen/abschätzen, wie groß der Diskretisierungsfehler ist? Welche Größe interessiert uns? Kann ich effizient die Diskretisierung "steuern", um den Fehler zu reduzieren?

Beispiel 1: Potentialströmung

$\int \kappa \nabla u \cdot \nabla \phi \, \mathrm{d}x = \int g_{\mathrm{N}} \phi \, \mathrm{d}a$: Betrachte $J(u) = \int\limits_{\Gamma} \kappa \nabla u \cdot n \, \mathrm{d}a$.

Dualitätstrick: Berechne $w : \Omega \to \mathbb{R}$ mit $w|_{\Gamma_{\mathrm{D}}} = 0$ und

$$\int\limits_{\Omega} \kappa \nabla \phi \cdot \nabla w \, \mathrm{d}x = J(\phi) \quad \forall\, \phi \;\text{ mit }\; \phi|_{\Gamma_{\mathrm{D}}} = 0.$$

$$\begin{aligned}
\Rightarrow J(u - u_h) &= \int\limits_{\Omega} \kappa \nabla (u - u_h) \cdot \nabla w \, \mathrm{d}x \qquad \text{für } \; u_h \in V_h \\
&= \int\limits_{\Omega} \kappa \nabla (u - u_h) \cdot \nabla (w - \Pi_h w) \, \mathrm{d}x \\
&= \int\limits_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}} (w - \Pi_h w) \, \mathrm{d}a - \int\limits_{\Omega} \kappa \nabla u_h \cdot \nabla (w - \Pi_h w) \, \mathrm{d}x \\
&= \int\limits_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}} (w - \Pi_h w) \, \mathrm{d}a - \sum_K \int\limits_{\partial K} \kappa \nabla u_h \cdot n (w - \Pi_h w) \, \mathrm{d}a
\end{aligned}$$

$$\begin{aligned}
\Rightarrow |J(u - u_h)| \leq& \| g_{\mathrm{N}} - \kappa \nabla u_h \|_{\Gamma_{\mathrm{N}}} \| w - \Pi_h w \|_{\Gamma_{\mathrm{N}}} \\
&+ \frac{1}{2} \sum_{F \in \mathcal{F}_h \setminus \Omega} \| \kappa (\nabla u_K - \nabla u_{K_F}) \|_F \, \| w - \Pi_h w \|_F \\
\approx& \sum_F r_F w_F
\end{aligned}$$

mit $\int\limits_{\Omega} \kappa \nabla \phi_h \cdot \nabla w_h \, \mathrm{d}x = J(\phi_h)$,

$$r_F = \begin{cases} \| g_{\mathrm{N}} - \kappa \nabla u_h \|_{\Gamma_{\mathrm{N}}} & \Gamma \subset \Gamma_{\mathrm{N}} \\ 0 & \Gamma \subset \Gamma_{\mathrm{D}} \\ \| \kappa (\nabla u_K - \nabla u_{K_F}) \|_F & F \subset \Omega \end{cases}$$

und $w_F = \| w_h - \Pi_h w_h \|_F$.

Beispiel 2: Transportgleichung

Bestimmt $\rho_h^n \in Q_h$ mit

$$\frac{1}{\Delta t} \left(\rho_h^n - \rho_h^{n-1} \right) = A_h \left(\frac{1}{2} \rho_h^n + \frac{1}{2} \rho_h^{n-1} \right) + \frac{1}{2} b_h^n + \frac{1}{2} b_h^{n-1}.$$

Definiere

$$\rho_h(t) = \frac{t - t_{n-1}}{\Delta t}\rho_h^{n-1} + \frac{t_n - t}{\Delta t}\rho_h^n \quad \text{mit } t \in (t_{n-1}, t_n)$$

$$\Rightarrow \partial_t \rho_h = \frac{1}{\Delta t}\left(\rho_h^n - \rho_h^{n-1}\right)$$
$$\Rightarrow (\partial_t \rho_h, \phi_h)_{\Omega \times (t_{n-1}, t_n)} = (A_h \rho_h + b_h, \phi_h)_{\Omega \times (t_{n-1}, t_n)}.$$

Zielfunktional $J(\rho) = \int\limits_0^T \int\limits_\Gamma \rho\, q \cdot n \,\mathrm{da}$.
Dualitätstrick: Bestimme $w : \Omega \times [0, T] \to \mathbb{R}$ mit $w(T) = 0, w|_{\Gamma_{\text{out}}} \equiv 0$ und

$$-\int\limits_\Omega \int\limits_0^T \phi\left(\partial_t w + q \cdot \nabla w\right)\,\mathrm{d}t\,\mathrm{d}x = J(\phi)$$

$\forall \phi$ mit $\phi(0) = 0, \phi|_{\Gamma_{\text{in}}} \equiv 0$. Daraus folgt

$$\begin{aligned}
J(\rho - \rho_h) &= -\int\limits_0^T \int\limits_\Omega (\rho - \rho_h)(\partial_t w + q \cdot \nabla w)\,\mathrm{d}x\,\mathrm{d}t \\
&= \int\limits_0^T \Bigg[\int\limits_\Omega \left[(\partial_t \rho - \partial_t \rho_h)\, w + \operatorname{div}(\rho_h q)\right]\,\mathrm{d}x - \int\limits_{\partial\Omega} \rho w\, q \cdot n \,\mathrm{da} - \sum_K \int\limits_K \operatorname{div}(\rho_h q) w \,\mathrm{d}x \\
&\quad + \int\limits_{\partial K} \rho_h w\, q \cdot n \,\mathrm{da} \Bigg]\,\mathrm{d}t \\
&= -\int\limits_0^T \Bigg[\int\limits_\Omega \left(\partial_t \rho_h + \operatorname{div}(\rho_h q)\right)(w - w_h)\,\mathrm{d}x + \sum \left(\Psi(\rho_h) - \Psi^*(\rho_h), w - w_h\right)_{\partial K \setminus \Gamma_{\text{in}}} \\
&\quad + \int\limits_{\Gamma_{\text{in}}} (\rho_{\text{in}} - \rho_h) q \cdot n (w - w_h)\,\mathrm{da} \Bigg]\,\mathrm{d}t, \qquad \forall w_h \in Q_h
\end{aligned}$$

Adaptiver Algorithmus:

1) Berechne (primale) Lösung ρ_h ("Vorwärtslösung")

2) Berechne (duale) Lösung w_h ("Rückwärtslösung")

3) Berechne "Residuen" ρ_F aus ρ_h und "Gewichte" w_F aus w_h

4) Verfeinere K (oder erhöhe Polynomgrad) für K mit $\eta_K > \Theta \max\limits_{K'} \eta_{K'}$,
$\eta_K = \sum\limits_{F \in \mathcal{F}_K} \rho_F w_F$.

Kontrolle von Strömung und Transport

Wähle Pumpraten $g_1, ..., g_L$ an Brunnen $\Gamma_1, ..., \Gamma_L$ und bestimme $u = u(g)$ mit

$$\int\limits_{\Omega} \kappa \nabla u \cdot \nabla \phi \,\mathrm{d}x = \int\limits_{\Gamma_\mathrm{N}} g_\mathrm{N} \phi \,\mathrm{d}a + \sum_{l=1}^{L} \int\limits_{\Gamma_l} g_l \phi \,\mathrm{d}a$$

$\Rightarrow\ u = u_\mathrm{N} + \sum g_l u_l$ mit

$$\begin{aligned} \int\limits_{\Omega} \kappa \nabla u_l \cdot \nabla \phi \,\mathrm{d}x &= \int\limits_{\Gamma_l} \phi \,\mathrm{d}a \\ \int\limits_{\Omega} \kappa \nabla u_\mathrm{N} \cdot \nabla \phi \,\mathrm{d}x &= \int\limits_{\Gamma_\mathrm{N}} g_\mathrm{N} \phi \,\mathrm{d}a \end{aligned}$$

Sei

$$\begin{aligned} q(g) &= -\kappa \nabla u(g) \\ &= q_\mathrm{N} + \sum g_l q_l \end{aligned}$$

mit $\rho(g)$ Lösung von

$$\begin{aligned} \partial_t \rho + \operatorname{div}(q(g)\rho) &= 0 \\ \rho(0) &= \rho_0 \\ \rho|_{\Gamma_\text{in}} &= \rho_\text{in}. \end{aligned}$$

Aufgabe: Bestimme $g \in \mathbb{R}^l$, so dass $J(\rho) = \int\limits_{\Gamma} \rho q \cdot n \,\mathrm{d}a$ maximal ist, d.h. maximiere

$$\begin{aligned} j(g) &= J(\rho) \\ \text{unter den NB}\quad (\partial_t \rho + \operatorname{div}(\rho q), \phi)_\Omega &= (\rho - \rho_\text{in}, q \cdot n\phi)_{\Gamma_\text{in}} \ \ \forall t \\ q &= q_\mathrm{N} + \sum g_l q_l \\ \rho(0) &= \rho_0. \end{aligned}$$

Die Kritische Stelle des Lagrangefunktionals $\mathcal{L}$ ist das Optimum von j:

$$\mathcal{L}(g, \rho, \eta) = J(\rho) + (\partial_t \rho + \operatorname{div}(\rho q(g)), \eta)_{\Omega \times (0,T)} - (\rho - \rho_\text{in}, q(g) \cdot n\eta)_{\Gamma_\text{in} \times (0,T)} + (\rho(0) - \rho_0, \eta(0))_\Omega$$

mit

$$\begin{aligned} \eta|_{\Gamma_\text{out}} &= 0 \\ \eta(T) &= 0, \end{aligned}$$

d.h.

$$\begin{aligned} 0 &= \partial_l J(\rho) + (\operatorname{div}(\rho q_l), \eta) - (\rho - \rho_\text{in}, q_l \cdot n\eta) \\ &= J(\phi) + (\partial_t \phi + \operatorname{div}(\rho q_l), \eta) - (\phi, q(g) \cdot n\eta) + (\phi(0), \eta(0)) \\ &= (\partial_t \rho + \operatorname{div}(\rho q(g)), \phi) - (\rho - \rho_\text{in}, q(g) \cdot n\phi) + (\rho(0) - \rho_0, \phi(0)) \end{aligned}$$

Beobachtung

ρ löst

$$\begin{aligned} \partial_t \rho + \operatorname{div}(q(g)\rho) &= 0 \\ \rho(0) &= \rho_0 \\ \rho|_{\Gamma_{\text{in}}} &= \rho_{\text{in}} \end{aligned}$$

und η löst

$$\begin{aligned} \partial_t \eta + q(g) \cdot \nabla \eta &= J \\ \eta(T) &= \rho_0 \\ \eta|_{\Gamma_{\text{out}}} &= 0 \end{aligned}$$

rückwärts in der Zeit.

$$\Rightarrow \ \partial_l j(g) = -\left(\operatorname{div}(\rho q_l), \eta\right)_{\Omega \times (0,T)}$$

Taylor:

$$\begin{aligned} j(g + \delta g) &\approx j(g) + \nabla j(g) \delta g \\ \Rightarrow \delta g &= -s \nabla j(g) \end{aligned}$$

ist Abstiegsrichtung.

S0) Wähle Startwert $g^0 = 0$, berechne q_{N}

S1) Berechne $q_1, ..., q_L,\ \rho = \rho(g)$ und $\eta = \eta(g)$

S2) Berechne $\nabla j(g)$ aus $q_l,\ \rho$ und η

S3) Falls $|\nabla j(g)|$ klein STOP

S4) Wähle $s > 0$, so dass $j(g + \delta g) < j(g)$ für $\delta g = -s \nabla j(g)$

Reduziertes Modell (auch für g zeitabhängig)
Zustand x, Kontrolle u, Output y

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned}$$

mit geeigneter Basis $\mathcal{B} \subset W_h \times Q_h$ zur Berechnung von A.

6 Konvektions-Reaktions-Diffusions-Systeme

6.1 Die Konvektions-Reaktions-Diffusions-Gleichung

Sei $\Omega \subset \mathbb{R}^D$, außerdem seien

$$\begin{array}{rll} q\colon & \Omega \longrightarrow \mathbb{R}^D \text{ mit } \operatorname{div} q = 0 & \text{(Flussvektor)} \\ r\colon & \Omega \times [0,T] \times \mathbb{R} \longrightarrow \mathbb{R} & \text{(Reaktionsrate)} \\ \kappa_\mathrm{c}\colon & \Omega \longrightarrow \mathbb{R}^{D\times D} & \text{(Diffusionstensor, symmetrisch, positiv definit)} \end{array}$$

gegeben. Gesucht wird die Konzentration eines Stoffes $c\colon \Omega \times [0,T] \longrightarrow \mathbb{R}$, so dass für alle Kontrollvolumina $Y \subset \Omega$ die Bilanzgleichung

$$\frac{\mathrm{d}}{\mathrm{d}t} \int\limits_Y c(x,t)\,\mathrm{d}x = -\int\limits_{\partial Y} \psi(c) \cdot n\,\mathrm{d}x + \int\limits_Y r(x,t,c(x,t))\,\mathrm{d}x$$

mit dem Fluss $\psi(c) = -\kappa_\mathrm{c} \nabla c + cq$ erfüllt ist. Daraus ergibt sich die Konvektions-Reaktions-Diffusions-Gleichung

$$\partial_t c + \operatorname{div}(-\kappa_\mathrm{c} \nabla c + cq) = r(c)\,. \tag{3}$$

Die Anfangsbedingung $c(x,0) = c_0(x),\ x \in \Omega$, wird vorgegeben. Außerdem werden auf dem Rand $\partial\Omega = \Gamma_\mathrm{D} \cup \Gamma_\mathrm{N} \cup \Gamma_\mathrm{R}$ die folgenden Randbedingungen gefordert:

$$\begin{array}{lrcll} \text{Dirichlet-RB:} & c(x,t) & = & c_\mathrm{D}(x,t), & x \in \Gamma_\mathrm{D}, t \in [0,T] \\ \text{Neumann-RB:} & \kappa_\mathrm{c} \nabla c(x,t) \cdot n(x) & = & g_\mathrm{N}(x,t), & x \in \Gamma_\mathrm{N}, t \in [0,T] \\ \text{gemischte RB:} & \kappa_\mathrm{c} \nabla c(x,t) \cdot n(x) + \alpha c(x,t) & = & g_\mathrm{R}(x,t), & x \in \Gamma_\mathrm{R}, t \in [0,T], \alpha\colon \Gamma_\mathrm{R} \longrightarrow \mathbb{R} \end{array}$$

Spezialfälle:

$$\begin{array}{rcl} r(x,t,c) & = & r(x,t) \begin{cases} > 0, & \text{Quelle} \\ < 0, & \text{Senke} \end{cases} \\ r(x,t,c) & = & r_0(x,t)c \ \text{ mit } r_0(x,t) \begin{cases} > 0, & \text{exponentielles Wachstum} \\ < 0, & \text{exponentieller Verfall} \end{cases}. \end{array}$$

Teilprobleme:

$$\begin{array}{lrcl} \text{Transport/Konvektion} & \partial_t c + \operatorname{div}(cq) & = & 0 \\ \text{Reaktion} & \partial_t c & = & r(c) \\ \text{Diffusion} & \partial_t c & = & \operatorname{div}(\kappa_\mathrm{c} \nabla c) \\ \text{Wärmeleitungsgleichung} & \partial_t c & = & \delta c \\ \text{Poissongleichung} & 0 & = & \delta c + r \end{array}$$

(6.1) Satz

Sei c Lösung der Konvektions-Reaktions-Diffusions-Gleichung (3). Dann gilt für alle ϕ im Testraum V der glatten, auf Γ_D verschwindenen Funktionen ($\phi|_{\Gamma_\mathrm{D}} = 0$)

$$\int\limits_\Omega \big(\partial_t c\,\phi + \kappa_\mathrm{c} \nabla c \cdot \nabla \phi + \nabla c \cdot q \phi\big)\,\mathrm{d}x + \int\limits_{\Gamma_\mathrm{R}} \alpha c \phi\,\mathrm{d}a = \int\limits_\Omega r(c)\phi\,\mathrm{d}x + \int\limits_{\Gamma_\mathrm{N}} g_\mathrm{N} \phi\,\mathrm{d}a + \int\limits_{\Gamma_\mathrm{R}} g_\mathrm{R} \phi\,\mathrm{d}a\,.$$

Beweis.

$$\begin{aligned}
\int\limits_{\Omega} \partial_t c \phi \,\mathrm{d}x &= -\int\limits_{\Omega} \operatorname{div} \psi(c) \phi \,\mathrm{d}x + \int\limits_{\Omega} r(c)\phi \,\mathrm{d}x \\
&= \int\limits_{\Omega} \operatorname{div}(\kappa_{\mathrm{c}} \nabla c) \phi \,\mathrm{d}x - \int\limits_{\Omega} \nabla c \cdot q\phi \,\mathrm{d}x + \int\limits_{\Omega} r(c)\phi \,\mathrm{d}x \\
&= -\int\limits_{\Omega} \kappa_{\mathrm{c}} \nabla c \cdot \nabla \phi \,\mathrm{d}x + \int\limits_{\partial\Omega} \kappa_{\mathrm{c}} \nabla c \cdot n \phi \,\mathrm{d}a - \int\limits_{\Omega} \nabla c \cdot q\phi \,\mathrm{d}x + \int\limits_{\Omega} r(c)\phi \,\mathrm{d}x \\
&= -\int\limits_{\Omega} \kappa_{\mathrm{c}} \nabla c \cdot \nabla \phi \,\mathrm{d}x - \int\limits_{\Gamma_{\mathrm{R}}} \alpha c \phi \,\mathrm{d}a + \int\limits_{\Gamma_{\mathrm{R}}} g_{\mathrm{R}} \phi \,\mathrm{d}a + \int\limits_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}} \phi \,\mathrm{d}a \\
&\quad - \int\limits_{\Omega} \nabla c \cdot q\phi \,\mathrm{d}x + \int\limits_{\Omega} r(c)\phi \,\mathrm{d}x \,.
\end{aligned}$$

Das liefert die gewünschte Aussage. □

Definiere die linearen Operatoren $A_h, M_h \colon V_h \longrightarrow V_h$ und den nichtlinearen, differenzierbaren Operator $R_h \colon V_h \longrightarrow V_h$ durch

$$\begin{aligned}
(M_h V_h, \phi_h)_{\Omega} &= \int\limits_{\Omega} V_h \phi_h \,\mathrm{d}x \\
(A_h V_h, \phi_h)_{\Omega} &= \int\limits_{\Omega} \left(\kappa_{\mathrm{c}} \nabla V_h \cdot \nabla \phi_h + \nabla V_h \cdot q \phi_h\right) \,\mathrm{d}x + \int\limits_{\Gamma_{\mathrm{R}}} \alpha V_h \phi \,\mathrm{d}a \\
(R_h(V_h), \phi_h)_{\Omega} &= \int\limits_{\Omega} r(V_h)\phi_h \,\mathrm{d}x + \int\limits_{\Gamma_{\mathrm{R}}} g_{\mathrm{R}} \phi_h \,\mathrm{d}a + \int\limits_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}} \phi_h \,\mathrm{d}a \,.
\end{aligned}$$

Dann gilt nach Satz 6.1 für eine Lösung c der Reaktions-Diffusions-Gleichung (3)

$$(M_h \partial_t c + A_h c_h, \phi_h)_{\Omega} = (R(c_h), \phi_h)_{\Omega} \quad \forall \phi_h \in V_h \text{ und } t \in (0, T).$$

Mit einem Galerkin-Ansatz ergibt sich das semi-diskrete Problem: Finde $c_h(t) \in V_h(c_{\mathrm{D}}(t))$, so dass

$$(M_h \partial_t c_h + A_h c_h, \phi_h)_{\Omega} = (R(c_h), \phi_h)_{\Omega} \quad \forall \phi_h \in V_h(0).$$

Mit dem impliziten Eulerverfahren und der Zeitdiskretisierung $t_n = n\triangle t$ ergibt sich

$$\frac{1}{\triangle t} \left(M_h \left(c_h^n - c_h^{n-1}\right), \phi_h\right)_{\Omega} + (A_h c_h^n, \phi_h)_{\Omega} = (R(c_h^n), \phi_h)_{\Omega} \quad \forall \phi_h \in V_h(0).$$

Damit ergibt sich das folgende Verfahren:
Löse

$$G_h^n(c) = 0 \;\text{ mit }\; G_h^n(c, \phi_h) = \left(\frac{1}{\triangle t} M_h (c_h^n - c_h^{n-1}) + A_h c_h^n, \phi_h\right)_{\Omega} - (R(c_h^n), \phi_h)_{\Omega}$$

Jacobi-Matrix:

$$\begin{aligned}
(J_h^n(c) v_h, \phi_h)_{\Omega} &= \left(\frac{1}{\triangle t} M_h v_h + A_h v_h, \phi_h\right)_{\Omega} - (R'(c_h^n) v_h, \phi_h)_{\Omega} \\
&= \left(\frac{1}{\triangle t} M_h v_h + A_h v_h, \phi_h\right)_{\Omega} - \int\limits_{\Omega} \partial_3 r(x, t, c(x,t)) v_h, \phi_h \,\mathrm{d}x
\end{aligned}$$

$$\Rightarrow J = (J_h^n(c) \lambda_j, \lambda_m)_{\Omega} \in \mathbb{R}^{N \times N} \;\text{ mit } v_h = span\{\lambda_j\}.$$

S0) Wähle $\triangle t > 0$, setze $n = 1, t_0 = 0, c_h^0(z) = c_0(z)\ \forall z \in \mathcal{N}_h$. Sei $\varepsilon > 0, \Theta > 0$. Assembliere

$$\begin{aligned}
\underline{M}_h &= \left(\int_\Omega \phi_m \phi_k \,\mathrm{d}x \right)_{k,m} = ((M\phi_m, \phi_k)_\Omega)_{k,m} \\
\underline{A}_h &= \left(\int_\Omega (K_\mathrm{c} \nabla\phi_m \cdot \nabla\phi_k + \nabla\phi_m \cdot q\phi_k)\,\mathrm{d}x + \int_{\Gamma_\mathrm{R}} \alpha\phi_m\phi_k \,\mathrm{d}a \right)_{k,m} \\
&= ((A\phi_m, \phi_k)_\Omega)_{k,m}
\end{aligned}$$

mit $V_h(0) = \mathrm{span}\{\phi_1, \dots, \phi_N\}$.

S1) Wähle $c_h^{n,0} \in V_h(c_\mathrm{D}(t_n))$, setze $k = 0$.

S2) Berechne

$$\begin{aligned}
\underline{b}_h^{n,k} &= \Bigg[\int_\Omega \Bigg(K_\mathrm{c} \nabla c_h^{n,k} \cdot \nabla\phi_j + \nabla c_h^{n,k} \cdot q\phi_j + \frac{1}{\triangle t}\left(c_h^{n,k} - c_h^{n-1}\right)\phi_j \\
&\quad + \int_{\Gamma_\mathrm{R}} \alpha c_h^{n,k}\phi_j \,\mathrm{d}a - \int_\Omega r\left(c_h^{n,k}\right)\phi_j\,\mathrm{d}x - \int_{\Gamma_\mathrm{R}} g_\mathrm{R}\phi_j\,\mathrm{d}a - \int_{\Gamma_\mathrm{R}} g_\mathrm{N}\phi_j\,\mathrm{d}a \Bigg)\mathrm{d}x \Bigg]_j \\
&= \left[\left(\frac{1}{\triangle t} M\left(c_h^{n,k} - c_h^{n-1}\right) + Ac_h^{n,k} - R(c_h^{n,k}), \phi_j \right)_\Omega \right]_j \\
&= b_h(c_h^{n,k})
\end{aligned}$$

S3) Falls $|\underline{b}^{n,k}| < \varepsilon$ oder $|\underline{b}^{n,k}| < \Theta|\underline{b}^{n,0}|$ und falls $t_n < T$: Setze $c_h^n = c_h^{n,k},\ n := n + 1$ und gehe zu S1).

S4) Assembliere

$$\underline{J}_h^{n,k} = \left(\left(\left(\frac{1}{\triangle t} M + A - R'(c_h^{n,k}) \right) \phi_i, \phi_j \right)_\Omega \right)_{i,j} \in \mathbb{R}^{N\times N}.$$

S5) Löse approximativ $\underline{d}^{n,k} = -\underline{J}_h^{n,k} \underline{b}_h^{n,k}$.

S6) Bestimme die Schrittweite $s_{n,k} \in \left\{1, \frac{1}{2}, \frac{1}{4}, \dots\right\}$ mit

$$\left| b_h(c_h^{n,k} + s_{n,k}\underline{d}^{n,k}) \right| < \left| \underline{b}_h^{n,k} \right|$$

und setze $c_h^{n,k+1} = c_h^{n,k} + s_{n,k} d_h^{n,k}$, $k := k + 1$ und gehe zu S2).

Bemerkung
Es fehlen noch Details zur Divergenzkontrolle.

(6.2) Satz
Sei r monoton fallend in c, d. h. $\partial_3 r(x,t,c) \le 0$, und $\Gamma_\mathrm{D} \neq \emptyset$. Es gelte $q \cdot n \ge 0$ auf Γ_N und $\alpha + \frac{1}{2} q \cdot n > 0$ auf Γ_R. Dann ist das Newton-Verfahren wohldefiniert und konvergent, d.h. J ist positiv definit, also

$$(J_h^n(c)\phi_h, \phi_h) > 0 \ \textit{für}\ \ \phi \neq 0.$$

Beweis.

$$\begin{aligned}
(J_h^n(c)\phi_h, \phi_h)_\Omega &= \frac{1}{\triangle t}(M_h\phi_h,\phi_h)_\Omega + (A_h\phi_h,\phi_h)_\Omega - (R'(c)\phi_h,\phi_h)_\Omega \\
&\geq \frac{1}{\triangle t}\|\phi_h\|_\Omega^2 + \int_\Omega K_c\nabla\phi_h\cdot\nabla\phi_h\,\mathrm{d}x + \int_\Omega \nabla\phi_h\cdot q\phi_h\,\mathrm{d}x + \int_{\Gamma_\mathrm{R}} \alpha\phi_h^2\,\mathrm{d}a \\
&\geq \frac{1}{\triangle t}\|\phi_h\|_\Omega^2 + C_p^{-2}\|\phi_h\|_\Omega^2 + \frac{1}{2}\int_{\Gamma_\mathrm{N}} \phi_h^2 q\cdot n\,\mathrm{d}a + \int_{\Gamma_\mathrm{R}}\left(\alpha + \frac{1}{2}q\cdot n\right)\phi_h^2\,\mathrm{d}a \\
&\geq \left(\frac{1}{\triangle t} + C_p^{-2}\right)\|\phi_h\|_\Omega^2
\end{aligned}$$

Der erste und der letzte Term sind jeweils ≥ 0, da $\phi_h > 0$. Bleibt zu zeigen, dass

$$(A_h\phi_h,\phi_h)_\Omega > 0:$$

$$\begin{aligned}
(A_h\phi_h,\phi_h)_\Omega &= \int_\Omega \kappa_\mathrm{c}\nabla\phi_h\cdot\nabla\phi_h\,\mathrm{d}x + \int_\Omega \nabla\phi_h\cdot q\phi_h\,\mathrm{d}x + \int_{\Gamma_\mathrm{R}} \alpha\phi_h^2\,\mathrm{d}a \\
&= \int_\Omega \kappa_\mathrm{c}\nabla\phi_h\cdot\nabla\phi_h\,\mathrm{d}x + \frac{1}{2}\int_{\Gamma_\mathrm{N}} \phi_h^2 q\cdot n\,\mathrm{d}a + \int_{\Gamma_\mathrm{R}} \alpha\phi_h^2\,\mathrm{d}a \geq 0
\end{aligned}$$

$\underline{J}^{n,k}$ ist folglich positiv definit und somit ist das gedämpfte Newton-Verfahren global konvergent. $\square$

Bemerkung

A) Die ODE $\dot{c} = F(c)$ ist dissipartiv und jedes B-stabile Runge-Kutta-Verfahren konvergiert.

B) Wenn R' beschränkt ist, dann gilt das Euler-Verfahren für hinreichend kleine $\triangle t > 0$.

FE-Approximation

Sei $c_h^n \in V_h$ mit $c_h^n(z) = c_\mathrm{D}(z,t_n), z \in \mathcal{N}_h \cap \Gamma_\mathrm{D}$ und

$$\begin{aligned}
&\int_\Omega \left(\frac{1}{\triangle t}(c_h^n - c_h^{n-1})\cdot\phi_h + \kappa_c\nabla c_h^n\cdot\nabla\phi_h + q\cdot\nabla c_h^n\cdot\phi_h\right)\mathrm{d}x + \int_{\Gamma_\mathrm{R}} \alpha c_h^n\phi_h\,\mathrm{d}a \\
&= \int_\Omega r(t_n, c_h^n)\phi_h\,\mathrm{d}x + \int_{\Gamma_\mathrm{N}} g_N\cdot\phi_h\,\mathrm{d}a + \int_{\Gamma_\mathrm{R}} g_R\cdot\phi_h\,\mathrm{d}a
\end{aligned}$$

$\forall\ \phi_h \in V_h(0)$, d.h. $\phi_h(z) = 0$ für $z \in \mathcal{N}_h \cap \Gamma_\mathrm{D}$.
Sei $V_h(0) = span\{\lambda_1, ..., \lambda_N\}, \mathcal{N}_h \setminus \Gamma_\mathrm{D} = \{z_1, ..., z_N\}$ mit $\lambda_j(z_k) = \delta_{jk}$. Zu vorgegebenem c_h^{n-1} definiere

$$\begin{aligned}
\underline{G}^n(c_h) = &\int_\Omega \left(\frac{1}{\triangle t}(c_h^n - c_h^{n-1})\cdot\lambda_j + \kappa_c\nabla c_h\cdot\lambda_j + q\cdot\nabla c_h\cdot\lambda_j\right)\mathrm{d}x + \int_{\Gamma_\mathrm{R}} \alpha c_h\lambda_h\,\mathrm{d}a \\
&- \int_\Omega r(t_n, c_h)\lambda_h\,\mathrm{d}x - \int_{\Gamma_\mathrm{N}} g_N\cdot\lambda_h\,\mathrm{d}a - \int_{\Gamma_\mathrm{R}} g_R\cdot\lambda_h\,\mathrm{d}a
\end{aligned}$$

mit $j = 1, ..., N \in \mathbb{R}^N$ und

$$\underline{J}^n(c_h) = \int\limits_\Omega \left(\frac{1}{\triangle t} \lambda_j \lambda_k + \kappa_c \nabla \lambda_k \cdot \nabla \lambda_j + q \cdot \nabla \lambda_k \lambda_j \right) \mathrm{d}x + \int\limits_{\Gamma_\mathrm{R}} \alpha \lambda_k \lambda_j \, \mathrm{d}x$$
$$- \int\limits_\Omega \partial r(t_n, c_h) \lambda_k \lambda_j \, \mathrm{d}x$$

für $k, j = 1, ..., N \in \mathbb{R}^{N \times N}$. Zu $\underline{G} \in \mathbb{R}^N$ definiere $G_n = \sum \underline{G}[j] \lambda_j \in V_h(0)$ und zu $\underline{J} \in \mathbb{R}^{N \times N}$ definiere $J_h : V_h(0) \to V_h(0)$ mit $J_h \lambda_k = \sum\limits_{j=1}^N \underline{J}[j, k] \lambda_j \in V_h(0)$.

Implizites Euler-Verfahren mit Newton-Iteration in jedem Zeitschritt

S0) Wähle $\triangle t > 0,\ \varepsilon > 0,\ \theta > 0,\ l_{max} \in \mathbb{N}$ und setze $c_h^{-1} = 0,\ c_h^0(z) = c_0(z)\ \forall z \in \mathcal{N}_h,\ t_0 = 0,\ n = 1$

S1) $t_n = t_{n-1}$, setze $c_h^{n,0}(z) = \begin{cases} c_D(t_n, z) & z \in \mathcal{N}_h \cap \Gamma_\mathrm{D} \\ 2c_h^{n-1} - c_h^{n-2} & z \in \mathcal{N}_h \setminus \Gamma_\mathrm{D} \end{cases}$,
d.h. $c_h^{n,0} \in V(c_D(t_n))$, setze $k = 0$

S2) Berechne $\underline{G}^{n,0} = \underline{G}^n(c_h^{n,0})$ und setze $\varepsilon_n = \max\{\varepsilon, \theta|\underline{G}^{n,0}\}$

S3) Falls $|\underline{G}^{n,k}| \leq \varepsilon$ dann setze $c_h^n = c_h^{n,k}$,
falls $t_n < T$ setze $t_{n+1} = t_n + \triangle t,\ n := n + 1$ gehe zu S1)

S4) Berechne $\underline{J}^{n,k} = \underline{J}^n(c_h^{n,k})$ und löse $\underline{J}^{n,k} d^{n,k} = -\underline{G}^{n,k}$ (iterativ),
$d_h^{n,k} = \sum \underline{d}^{n,k}[j] \lambda_j \in V_h(0)$

S5) Setz $l = 0,\ s_{n,k} = 1$

S6) Setze $c_h = c_h^{n,k} + s_{n,k} d_h^{n,k}$ und berechne $\underline{G} = \underline{G}^n(c_h)$

S7) Falls $|\underline{G}| < |\underline{G}^{n,k}|$, dann setze $c_h^{n,k+1} = c_h,\ \underline{G}^{n,k+1} = \underline{G},\ k = k+1$ und gehe zu S3)

S8) Falls $l = l_{max}$ setze $\triangle t := \frac{1}{2} \triangle t$ und gehe zu S1)

S9) Setze $l := l + 1,\ s_{n,k} := \frac{1}{2} s_{n,k}$ und gehe zu S6)

Betrachte

$$\dot{u} = F(u) \text{ in } \mathbb{R}^N \text{ mit } DF(u) y \cdot y \leq -\alpha |y|^2 \text{ und } \alpha \geq 0\ \forall u \in \mathbb{R}^N.$$

Dann gilt

$$(F(u) - F(v)) \cdot (u - v) = \int\limits_0^1 DF(v + t(u - v))(u - v) \cdot (u - v) \, \mathrm{d}t$$
$$\leq -\alpha |u - v|^2$$

$\Rightarrow$ DGL ist dissipativ
$\Rightarrow$ $|u(t) - v(t)| \leq \exp(-\alpha t)|u(0) - v(0)|$ für $\dot{u} = F(u)$ und $\dot{v} = F(v)$

Implizites Eulerverfahren:

$u^n = u^{n-1} + \triangle t F(u^n)$
Zu gegebenem u^{n-1} bestimme u^n als Nullstelle von $G^n(u) = u - \triangle t F(u) - u^{n-1}$ d.h. $G^n(u^n) = 0$. u^n ist Fixpunkt $G^n(u) = u - (\mathcal{I} - \triangle t DF(u))^r G^n(u)$ für $\triangle t$ klein.

Analyse für verschwindende Diffusion $\kappa \to 0$

Betrachte $\kappa = \kappa_0 \mathcal{I},\ \kappa_0 > 0$ und

$$\begin{aligned} \alpha &= q \cdot n \\ g_{\mathrm{R}} &= 0 \\ r &= 0. \end{aligned}$$

Wir starten mit $c(0) > 0$, für $t > 0$: $\partial_t c + \operatorname{div}(cq - \kappa \nabla c) = 0$ mit

$$\begin{aligned} c = c_{\mathrm{D}} &\text{ auf } \varGamma_{\mathrm{D}} = \varGamma_{\mathrm{in}} \\ \kappa_0 \nabla c + \alpha c = 0 &\text{ auf } \varGamma_{\mathrm{R}} = \varGamma_{\mathrm{out}} \end{aligned}$$

1D- Reduktion (stationär):

Mit $\varOmega = (0,1),\ q \equiv 1,\ \varGamma_{\mathrm{in}} = \{0\},\ \varGamma_{\mathrm{out}} = \{1\}$ gilt:

$$\begin{aligned} -\kappa_0 c'' + c' &= 0 \qquad \text{in } (0,1), \\ c(0) &= 1, \\ \kappa_0 c'(1) + c(1) &= 0 \end{aligned}$$

$\exp(\lambda x)$ Ansatz liefert:

$$-\kappa_0 \lambda^2 + \lambda = 0 \Rightarrow \lambda = 0 \text{ oder } \lambda = \frac{1}{\kappa_0}$$
$$\Rightarrow c(x) = a_1 + a_2 \exp(\frac{x}{\kappa_0}) \text{ mit } c(0) = 1 = a_1 + a_1 \text{ und}$$
$$\kappa_0 \frac{a_2}{\kappa_0} \exp\left(\frac{1}{\kappa_0}\right) + a_1 + a_2 \exp\left(\frac{1}{\kappa_0}\right) = 0$$

Grenzwert $\rho' = 0,\ \rho(0) = 1$ auf $\varGamma_{\mathrm{in}} \Rightarrow \rho \equiv 1$.

(6.3) Lemma
Es gilt

$$\int_{\varOmega} |c(T)|^2 \,\mathrm{d}x \leq \int_{\varOmega} |c_0|^2 \,\mathrm{d}x + \int_0^T \left[\int_{\varGamma_{\mathrm{D}}} c^2 |q \cdot n| \,\mathrm{d}a + 2 \int_{\varGamma_{\mathrm{D}}} \kappa_0 \nabla c \cdot n c_{\mathrm{D}} \,\mathrm{d}a \right] \mathrm{d}t$$

Beweis.

$$\begin{aligned}
\int_\Omega \left(|c(T)|^2 - |c(0)|^2\right) \,\mathrm{d}x &= \int_\Omega \int_0^T \partial_t |c(t)|^2 \,\mathrm{d}t \,\mathrm{d}x = 2\int_0^T \int_\Omega \partial_t c \cdot c \,\mathrm{d}x \,\mathrm{d}t \\
&= 2\int_0^T \int_\Omega \mathrm{div}(\kappa_0 \nabla c - c\cdot q) c \,\mathrm{d}x \,\mathrm{d}t \\
&= 2\int_0^T \left(-\int_\Omega \kappa_0 \nabla c \cdot \nabla c \,\mathrm{d}x + \int_{\partial\Omega} \kappa_0 \nabla c \cdot n c \,\mathrm{d}a - \frac{1}{2}\int_\Omega \mathrm{div}(c^2 q) \,\mathrm{d}x \right) \,\mathrm{d}t \\
&\leq \int_0^T 2 \left[\int_{\Gamma_\mathrm{D}} \kappa_0 \nabla c \cdot c_\mathrm{D} \,\mathrm{d}a - 2\int_{\Gamma_\mathrm{R}} q\cdot n c^2 \,\mathrm{d}a + \int_{\Gamma_\mathrm{D}} |q\cdot n| c_D^2 \mathrm{d}a \right] \,\mathrm{d}t
\end{aligned}$$

□

Vergleiche mit Transportgleichung:

$$\partial_t \rho + \mathrm{div}(\rho q) = 0$$

mit $\rho = \rho_0 = c_\mathrm{D}$ auf Γ_D und $\rho(0) = \rho_0 = c_0$.

(6.4) Lemma
Es gilt:

$$\int_\Omega |c(T) - \rho(T)|^2 \leq 2\kappa_0 \int_0^T \|\nabla c\|_\Omega \|\nabla(\rho - c)\|_\Omega \,\mathrm{d}t,$$

falls $\nabla c, \nabla\rho \in L_2(\Omega)^\mathrm{D}$.

Folgerung: Falls $\|\nabla c\|$ und $\|\nabla \rho\|$ beschränkt für $\kappa_0 \to 0$, dann gilt $c \to \rho$ für $\kappa_0 \to 0$.

Beweis. Für $\phi = c - \rho$ gilt:

$$\begin{aligned}
\int_\Omega |\phi(T)|^2 \,\mathrm{d}x &= \int_\Omega \left(|\phi(T)|^2 - |\phi(0)|^2\right) \,\mathrm{d}x \\
&= 2\int_0^T \int_\Omega \partial_t \phi \; \phi \,\mathrm{d}x \,\mathrm{d}t \\
&= 2\int_0^T \left[-\int_\Omega \mathrm{div}(\phi q)\phi \,\mathrm{d}x + \int_\Omega \mathrm{div}\, \kappa_0 \nabla c \cdot \phi \,\mathrm{d}x \right] \,\mathrm{d}t \\
&\leq \int_0^T \left[\int_{\partial\Omega} q\cdot n\phi^2 \,\mathrm{d}a - 2\int_\Omega \kappa_0 \nabla c \cdot \nabla\phi \,\mathrm{d}x + 2\int_{\partial\Omega} \kappa_0 \nabla c \cdot n\phi \,\mathrm{d}a \right] \,\mathrm{d}t \\
&\leq 2\kappa_0 \int_0^T \|\nabla c\| \, \|\nabla\phi\| \,\mathrm{d}t - \int_0^T \int_{\partial\Gamma_\mathrm{out}} q\cdot n(3c^2 + \rho^2) \,\mathrm{d}a \,.
\end{aligned}$$

□

Streamline-Upwind-Petrov-Galerkin

Betrachte 1D-Reduktion: $-\kappa_0 c'' + qc' = 0$ in $\Omega = (0,1)$

$$\begin{aligned}\underline{A}[j,k] &= \int\limits_{\Omega} \left(\kappa_c \lambda_k' \lambda_j' + q \cdot \lambda_k^2\, \lambda_j\right)\,\mathrm{d}x \\ &= \begin{cases} \frac{\kappa_0}{h^2} 2h & j=k \\ -\frac{\kappa_0}{h^2}h - \frac{q}{2h}h & k=j-1 \\ -\frac{\kappa_0}{h^2}h + \frac{q}{2h}h & k=j+1 \end{cases}\end{aligned}$$

Für $\kappa_0 \ll h$ gilt:

$$A \approx \frac{q}{2}\begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & -1 & 0 & 1 & \\ & & & & \dots \end{pmatrix}$$

Für κ_0 und Finite Volumen gilt:

$$A_0 = q\begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & & & \dots \end{pmatrix}$$

Beobachtung:

1) Die Lösung mit zentralen Differenzen oszilliert.

2) Die Lösung mit upwind ist monoton, d.h. die Matrix A_0 ist eine M-Matrix

3) Die Differenz $A_0 - A = \frac{q}{2}(-1,2,-1) =$ “künstliche” Diffusion

Definiere

$$\begin{aligned}a(c,\phi) &= \int\limits_{\Omega} (\kappa\nabla c \cdot \nabla\phi + q \cdot \nabla c\phi + r\; c \cdot \phi)\,\mathrm{d}x + \int\limits_{\Gamma_\mathrm{R}} \alpha c \cdot \phi\,\mathrm{d}a \\ b(\phi) &= \int\limits_{\Gamma_\mathrm{N}} g_\mathrm{N}\phi\,\mathrm{d}a + \int\limits_{\Gamma_\mathrm{R}} g_\mathrm{R}\phi\,\mathrm{d}a + \int\limits_{\Omega} f\;\phi\,\mathrm{d}x\end{aligned}$$

Dann gilt für ein Galerkin-Verfahren Ansatz- und Testfunktionen:

$$c_h, \phi_h \in V_h$$

Petrov-Galerkin Ansatzfunktion $c_h \in V_h$, Testfunktion $\phi_h|_K = \phi_h + \delta_K q \cdot \nabla\phi_h$

$$\begin{aligned}a_h(c,\phi) &= a(c,\phi) + \sum_K \delta_K\left(-\operatorname{div}\kappa\nabla c + q\cdot\nabla c + rc, q\cdot\nabla\phi_h\right)_K \\ b_h(\phi) &= b(\phi) + \sum_K \delta_K(f, q\cdot\nabla\phi_h)_K\end{aligned}$$

Quadratische Elemente:

$$V_h = span\{\lambda_z : z \in \mathcal{Z}_h\} = \{v_h \in C(\overline{\Omega}) : v_h|_K \in V_K\}$$

mit

$$\begin{aligned}
&V_K = \{v_K : v_k \circ \varphi_K \in \hat{V}\} \\
&\varphi_K : \hat{K} \to \overline{K} \\
&\mathcal{Z}_h = \mathcal{N}_h + \{\frac{1}{2}(z+y) : z, y \in \mathcal{V}_E, E \in \mathcal{E}_h\}
\end{aligned}$$

Lagrangeelemente:

$$v_h(x) = \sum_{z \in \mathcal{Z}_h} v_h(z)\lambda_z \;\text{ mit }\; \lambda_z(y) = \begin{cases} 1 & , z = y \\ 0 & , z \neq y. \end{cases}$$

Für ein Dreieck mit $D = 0$ und $\hat{V} = \mathbb{P}_2(\hat{K})$ gilt:

$$\begin{aligned}
\hat{\phi}_1(\xi) &= (1 - \xi_1 - \xi_2)\,(1 - 2\xi_1 - 2\xi_2) \\
\hat{\phi}_2(\xi) &= \xi_1\,(2\xi_1 - 1) \\
\hat{\phi}_3(\xi) &= \xi_2\,(2\xi_2 - 1) \\
\hat{\phi}_4(\xi) &= 4\xi_1\,(1 - \xi_1 - \xi_2) \\
\hat{\phi}_5(\xi) &= 4\xi_1\xi_2 \\
\hat{\phi}_6(\xi) &= 4\xi_2\,(1 - \xi_1 - \xi_2)
\end{aligned}$$

Für ein Viereck mit $\hat{V} \subset \mathbb{P}_2 \otimes \mathbb{P}_2$ gilt für den Mittelpunkt:

$$\hat{\phi}_g(\xi) = 16\xi_1(1-\xi_1)\xi_2(1-\xi_2) \;\in \mathbb{P}_4(\hat{K})$$

$$\begin{aligned}
\frac{1}{\Delta t}(c^n - c^{n-1}) - \operatorname{div}(\kappa_c \nabla c^n - q c^n) &= r'(c^n)c^n + f \\
\Rightarrow -\operatorname{div}(\kappa_c \nabla c^n) + q \cdot \nabla c^n + \left(\frac{1}{\Delta t} - r'(c^n)\right) c^n &= f + \frac{1}{\Delta t} c^{n-1}
\end{aligned}$$

mit der Approximation von c^n durch $c_h^n \in V_h$.
Stationär gilt

$$-\left(\operatorname{div} \kappa_c \nabla c\right) + q \cdot \nabla c + rc = f$$

$$\int_\Omega \kappa_c \nabla c \cdot \phi \,\mathrm{d}x + \int_\Omega q \cdot \nabla c\phi \,\mathrm{d}x + \int_\Omega rc\phi \,\mathrm{d}x + \int_{\Gamma_\mathrm{R}} q \cdot nc\phi \,\mathrm{d}a = \int_\Omega f\phi \,\mathrm{d}x$$

Suche diskrete Funktion:

$$c_h \in V_h(c_0) = \{v_h \in V_h : c_h(z) = c_\mathrm{D}(z) \;\forall z \in \mathcal{Z}_h\}$$

$$\begin{aligned}
a(c_h, \phi_h) = b(\phi_h) \;\Leftrightarrow\; &\underline{Ac} = \underline{b} \text{ mit} \\
\underline{A}[z, y] &= a(\lambda_z, \lambda_y) \\
\underline{b}[z] &= b(\lambda_z) \\
c_h &= \sum \underline{c}[z]\lambda_z
\end{aligned}$$

$$a(\phi,\phi) = \int_\Omega \kappa \nabla\phi\cdot\nabla\phi\,\mathrm{d}x + \int_\Omega q\cdot\nabla\phi\phi\,\mathrm{d}x + \int_\Omega r|\phi|^2\,\mathrm{d}x + \int_{\Gamma_\mathrm{R}} q\cdot n\phi^2\,\mathrm{d}a$$
$$\geq \kappa_0\|\nabla\phi\|^2 + r_0\|\phi\|^2_\Omega$$

$$\begin{aligned}
a_h(\phi_h,\phi_h) &\geq \kappa_0\|\nabla\phi\|^2 + r_0\|\phi\|^2_\Omega + \sum \delta_K\left[(-\operatorname{div}\kappa\nabla\phi + r\phi, q\cdot\nabla\phi)_K + \|q\cdot\nabla\phi\|^2_K\right]\\
&\geq \kappa_0\|\nabla\phi\|^2 + r_0\|\phi\|^2_\Omega + \sum \delta_K\left[(\|\operatorname{div}\kappa\nabla\phi\|_K + \|r\phi\|_K)\,\|q\cdot\nabla\phi\|_K + \|q\cdot\nabla\phi\|^2_K\right]\\
&\geq \kappa_0\|\nabla\phi\|^2 + r_0\|\phi\|^2_\Omega + \sum \delta_K\left[\|\operatorname{div}\kappa\nabla\phi\|^2_K + \|r\phi\|^2_K + \frac{1}{2}\|q\cdot\nabla\phi\|^2_K + \|q\cdot\nabla\phi\|^2_K\right]\\
&\geq \sum_K \left(\kappa_0 - \delta_K\|\kappa\|^2_\infty c_h^{-2}\right)\|\nabla\phi\|^2_K + \left(r_0 - \delta_K\|r\|_{\infty,K}\right)\|\phi\|^2_K + \frac{\delta_K}{2}\|q\cdot\nabla\phi\|^2_K
\end{aligned}$$

für $\kappa \ll h,\ \delta_K = O(h)$ und $q = O(1)$.
Zusammenfassend gilt

$$\begin{aligned}
a_h(\phi_h,\phi_h) &\geq \|\phi_h\|^2_{SD} \qquad (\star)\\
&= \frac{\kappa}{2}\|\nabla\phi_h\|^2_\Omega + \frac{r_0}{2}\|\phi_h\|^2_\Omega + \sum_K \frac{\delta_K}{2}\|q\cdot\nabla\phi\|^2_K
\end{aligned}$$

mit $\phi_h \in V_h(0)$ und $\mathbb{P}_1 \setminus \mathbb{P}_2$
Wofür brauchen wir diese Abschätzung?
Betrachte

$$c_h \in V(c_\mathrm{D}):\ a_h(c_h,\phi_h) = b_h(\phi_h)$$

und definiere die Finite Elemente Funktion $u_h = c_h - c_\mathrm{D}\ \in V_h(0)$ mit

$$c_\mathrm{D} = \sum_{z\in\mathcal{Z}_h\cap\Gamma_\mathrm{D}} c_\mathrm{D}(z)\,\phi_z + \sum_{z\in\mathcal{Z}_h\setminus\Gamma_\mathrm{D}} \left(2c_h^{n-1}(z) - c_h^{n-2}(z)\right)\phi_z$$

$$\begin{aligned}
&\Rightarrow a_h(u_h,\phi_h) = b_h(\phi_h) + a_h(c_\mathrm{D},\phi_h)\\
&\Rightarrow a_h(u_h,u_h) \overset{\phi_h=u_h}{=} b_h(u_h + a(c_\mathrm{D}),u_h)\\
&\overset{(\star)}{\Rightarrow} \|u_h\|^2_{SD} \leq \|f\|_\Omega\|u_h\|_\Omega + 2\|f\|_\Omega\|u_h\|_{SD} + 2\|c_\mathrm{D}\|_{SD}\|u_h\|_{SD}\\
&\Rightarrow \|u_h\|^2_{SD} \leq c\left(\|f\|_\Omega + \|c_\mathrm{D}\|_{SD}\right)
\end{aligned}$$

Bemerkung:
Für $\kappa \ll h,\ q = O(1),\ \delta_K = O(h)$ gilt für die Steifigkeitsmatrix

$$\underline{A}[j,k] = \left(a_h(\phi_j,\phi_k)\right)_{j,k}$$

positiv definit und wir haben Stabilität für $\kappa \to 0$ bzgl. $\|\cdot\|_{SD}$.

6.2 Linear-implizite Verfahren (Rosenbrock-Typ)

Aufgabe: Approximiere steife, nichtlineare gewöhnliche Differentialgleichungen der Form

$$\begin{aligned}
\dot{u} &= F(u) \quad \text{in } [0,T]\\
u(0) &= u_0 \quad\ \ \text{in } \mathbb{R}^N
\end{aligned}$$

mit $N = \dim V_h(0) = O\left(h^{-2}\right)$.
Anwendung:

$$u = c_h(z) - c_\mathrm{D}(z)_{z\in\mathcal{Z}}$$

$$F(u) = \underline{M}^{-1}\left(\int_\Omega -\kappa_c \nabla c \cdot \nabla\phi_k - q\cdot\nabla c_h\phi_k + r(c_h)\phi_k - \int_{\Gamma_\mathrm{R}} \alpha c_h\phi_k \,\mathrm{d}a\right)_{k=1,\dots,N}$$

mit $\underline{M}^{-1} = \int_\Omega \phi_j\phi_k \,\mathrm{d}x$.
Beobachtung: Im impliziten Eulerverfahren löst man

$$\frac{1}{\triangle t}\left(u^n - u^{n-1}\right) = F(u^n)$$

mit dem Newton-Verfahren. Wähle daher:

1) $\triangle t$ abhängig von h, so dass Fehler in Ort und Zeit vergleichbar groß sind.

2) Abbruch des Newton-Verfahrens, wenn der Abbruchfehler kleiner als der Diskretisierungsfehler ist. Oft genügt hierfür ein Newton-Schritt!

Idee: Linearisiere ODE in jedem einzelnen Schritt, approximiere die steife, lineare ODE implizit und die nichtlineare explizit.
Sei $F(u^n) \approx F(u^{n-1}) + J_{n-1}\left(u^n - u^{n-1}\right)$ mit $J_{n-1} = \mathrm{D}F(u^{n-1})$. Berechne

$$\frac{1}{\triangle t}\left(u^n - u^{n-1}\right) = F(u^{n-1}) + J_{n-1}\left(u^n - u^{n-1}\right),$$

d. h. es ergibt sich mit dem Newton-Verfahren

$$u^n = u^{n-1} + \triangle t\left(I_N - \triangle t J_{n-1}\right)^{-1} F(u^{n-1}).$$

Anwendung

$$J_{n-1} = \underline{M}_h^{-1}\left(\int_\Omega \left(-\kappa_c\nabla\phi_j\cdot\nabla\phi_k - q\cdot\nabla\phi_j\phi_k + r'(c_h^{n-1})\phi_j\phi_k\right)\,\mathrm{d}x - \int_{\Gamma_\mathrm{R}}\alpha\phi_j\phi_k\,\mathrm{d}a\right)_{j,k}$$
$$\in \mathbb{R}^{N\times N}.$$

Beispiel: Das Verfahren ode23s ergibt sich durch: Löse

$$\begin{aligned}\left(I_N - \alpha\triangle t J_{n-1}\right)k_1 &= F(u^{n-1}), \qquad \alpha = \frac{1}{2+\sqrt{2}}\\ \left(I_N - \alpha\triangle t J_{n-1}\right)k_2 &= F(u^{n-1} + \frac{\triangle t}{2}k_1) - \alpha\triangle t J_{n-1}k_1.\end{aligned}$$

Setze $u^n = u^{n-1} + \triangle t k_2$.

(6.5) Satz
Das Rosenbrock-Typ-Verfahren ode23s ist ein Verfahren zweiter Ordnung.

Beweis. i) Sei $\triangle t < \dfrac{1}{2\alpha\|J_{n-1}\|}$. Dann gilt $\|\triangle t \alpha J_{n-1}\| < 1$ und folglich ist die Matrix $I - \alpha \triangle t J_{n-1}$ invertierbar mit

$$(I_N - \alpha \triangle t J_{n-1})^{-1} = \sum_{l \geq 0} (\alpha \triangle t J_{n-1})^l$$

Das Verfahren ode23s ist also wohldefiniert für hinreichend kleine $\triangle t$.

ii) Setze $F = F(u^{n-1})$, $J = \mathrm{D}F(u^{n-1})$.
Es gilt

$$\begin{aligned} k_1 &= F + \alpha \triangle t J k_1 \\ &= F + \alpha \triangle t J \left(F + \alpha \triangle t J k_1\right) \\ &= F + \alpha \triangle t J F + \mathcal{O}(\triangle t^2) \end{aligned}$$

sowie

$$\begin{aligned} k_2 &= F(u^{n-1} + \frac{\Delta t}{2} k_1) - \alpha \Delta t J k_1 + \alpha \Delta t J k_2 \\ &= F + J\frac{\Delta t}{2} k_1 - \alpha \Delta t J k_1 + \alpha \Delta t J k_2 + \mathcal{O}(\Delta t^2) \\ &= F + J\frac{\Delta t}{2} \left(F + \mathcal{O}(\Delta t)\right) - \alpha \Delta t J (F + \mathcal{O}(\Delta t)) + \alpha \Delta t J (F + \mathcal{O}(\Delta t)) + \mathcal{O}(\Delta t^2) \\ &= F + \frac{\Delta t}{2} JF - \alpha \Delta t JF + \alpha \Delta t JF + \mathcal{O}(\Delta t^2) \\ &= F + \frac{\Delta t}{2} JF + \mathcal{O}(\Delta t^2). \end{aligned}$$

Damit folgt

$$u^n = u^{n-1} + \triangle t k_2 = u^{n-1} + \triangle t F + \frac{\Delta t^2}{2} JF + \mathcal{O}(\Delta t^3). \qquad \square$$

iii) Für den lokalen Diskretisierungsfehler gilt (mit $u^{n-1} = u(t_{n-1})$)

$$\begin{aligned} g^n &= \frac{1}{\Delta t} \left(u(t_n) - u(t_{n-1})\right) - k_2 \\ &= \frac{1}{\Delta t} \int_{t_{n-1}}^{t_n} F(u(t_n))\,\mathrm{d}t - k_2 \\ &= \frac{1}{\Delta t} \int_{t_{n-1}}^{t_n} \left(F + (t - t_n)\, JF\right)\,\mathrm{d}t - F - \frac{\Delta t}{2} JF + \mathcal{O}(\Delta t^2) = \mathcal{O}(\Delta t^2) \end{aligned}$$

Stabilitätsanalyse im linearen Fall $F(u) = Ju$

Das Verfahren ode23s ist ein Runge-Kutta-Verfahren mit

$$\begin{aligned} k_1 &= J(u^{n-1} + \alpha \triangle t k_1) \\ k_2 &= J(u^{n-1} + \frac{\Delta t}{2} k_1 + \alpha \triangle t k_2 - \alpha \triangle t k_1). \end{aligned}$$

Weiter gilt:

$$\begin{aligned} u^n &= u^{n-1} + \triangle t k_2 \\ &= u^{n-1} + (I_N - \alpha \triangle t J)^{-1} J \left(u^{n-1} + \frac{\Delta t}{2} k_1 - \alpha \triangle t k_1 \right) \\ &= R(\triangle t J) u^{n-1} \end{aligned}$$

mit

$$R(\xi) = \frac{1 + (1-2\alpha)\,\xi}{(1-\alpha\xi)^2}, \ |R(\xi)| \to 0 \ \text{ für } \ |\xi| \to \infty.$$

Mit $\alpha = \dfrac{1}{2+\sqrt{2}}$ folgt: $\alpha^2 - 2\alpha + \frac{1}{2} = 0$.

(6.6) Lemma
ode23s ist A-und L-stabil.

Beweis. Noch zu zeigen:

$$\{\xi \in \mathbb{C} \colon |R(\xi)| \leq 1\} \subset \mathbb{C}_- = \{\xi \in \mathbb{C} \colon \operatorname{Re} \xi \leq 0\}.$$

Sei $y \in \mathbb{R}$. Es gilt

$$\begin{aligned} |R(\mathrm{i}y)|^2 &= \left| \frac{1 + \mathrm{i}\,(1-2\alpha)\,y}{|1 - \mathrm{i}\alpha y|^2} \right|^2 = \frac{1 + (1-2\alpha)^2\,y^2}{(1+\alpha^2 y^2)^2} \\ &= \frac{1 + (1 - 4\alpha + 4\alpha^2)\,y^2}{1 + 2\alpha^2 y^2 + \alpha^4 y^4} = \frac{1 + 2\alpha^2 y^2}{1 + 2\alpha^2 y^2 + \alpha^4 y^4} \leq 1\,. \end{aligned}$$

Außerdem gilt:

- $R(\cdot)$ ist meromorph mit Polstelle $\xi = \frac{1}{\alpha}$ mit $\operatorname{Re} \xi > 0$.
- $R(\cdot)$ ist holomorph in $\mathbb{C}_-$
- $|R(\xi)| \leq 1$ auf $\partial \mathbb{C}_-$
- $\lim\limits_{|\xi| \to \infty} |R(\xi)| = 0$.

Mit dem Maximumsprinzip folgt

$$|R(\xi)| \leq 1 \ \text{ für } z \in \mathbb{C}_-.$$

Somit ist das Verfahren A- und L-stabil. □

7 Discontinuous Galerkin-Methode für die Diffusionsgleichung

Methode: SIP (Symmetric Inner Penalty)
Ziel: Übertragung des DG-Diskretisierung für die Transportgleichung auf elliptische (und parabolische Gleichungen)
Betrachte dafür

$$\begin{aligned} \operatorname{div} q &= 0 \\ q &= -\kappa \nabla u && \text{in } \Omega \\ \kappa \nabla u \cdot n &= g_{\mathrm{N}} && \text{auf } \Gamma_{\mathrm{N}} \\ u &= u_{\mathrm{D}} && \text{auf } \Gamma_{\mathrm{D}}. \end{aligned}$$

In $\overline{\Omega} = \overline{\bigcup K}$ für $\phi_h \in Q_h = \prod \mathbb{P}_P(K)$ gilt:

$$\begin{aligned} 0 &= \int\limits_K \operatorname{div} q \phi_h \, \mathrm{d}x \\ &= -\int\limits_K q \cdot \nabla \phi_h \, \mathrm{d}x + \int\limits_{\partial K} q \cdot n_K \phi_h \, \mathrm{d}a \end{aligned}$$

(7.1) Lemma (DG-Formel)
Sei $F = \partial K \cap \partial K_F$ und $(q|_K - q|_{K_F}) \cdot n_F = 0$. Dann gilt

$$q_K \cdot n_K \phi_K + q_{K_F} \cdot n_{K_F} \phi_{K_F} = \{q\}_F \cdot [\phi_h]_F$$

mit

$$\begin{aligned} \{q\}_F &= \frac{1}{2}(q_K + q_{K_F}) \ \textit{ und} \\ [\phi_h]_F &= \phi_K \cdot n_K + \phi_{K_F} \cdot n_{K_F} \\ &= (\phi_K - \phi_{K_F}) \cdot n_K \end{aligned}$$

Beweis.

$$\begin{aligned} 2\{q\}_F \cdot [\phi_h]_F &= q_K \cdot n_K \phi_K - q_K \cdot n_K \phi_{K_F} + q_{K_F} \cdot n_K \phi_K - q_{K_F} \cdot n_K \phi_{K_F} \\ &= 2q \cdot n_K(\phi_K - \phi_{K_F}) \end{aligned}$$ □

Daraus folgt $\forall \phi_h \in Q_h$

$$\begin{aligned} &\sum_K -\int\limits_K q \cdot \nabla \phi_h \, \mathrm{d}x + \int\limits_{\partial K \setminus \Gamma_{\mathrm{N}}} q \cdot n_{\mathrm{N}} \phi_h \, \mathrm{d}x = \int\limits_K \kappa \nabla u_h \cdot \nabla \phi_h \, \mathrm{d}x \\ \Leftrightarrow &\sum_K \int\limits_K \kappa \nabla u_h \cdot \nabla \phi_h \, \mathrm{d}x - \sum_{F \in \mathcal{F}_h \cap \Omega} \int\limits_F \{\kappa \nabla u\}_F \cdot [\phi_h] \, \mathrm{d}a - \sum_{F \in \mathcal{F}_h \cap \Gamma_{\mathrm{D}}} \int\limits_F \kappa \nabla u \cdot \phi_h \, \mathrm{d}a = \int\limits_{\Gamma_{\mathrm{N}}} g_N \phi_h \, \mathrm{d}a \end{aligned}$$

Bisher sind wir von Stetigkeit ausgegangen, was ändert sich, wenn Unstetigkeit vorliegt?
Sei $u|_K = u|_{K_F}$ auf F und $u = u_\mathrm{D}$ auf Γ_D

$$0 = \int_F [u]_F \cdot \{\kappa \nabla \phi_h\}_F \,\mathrm{d}a = \int_F [u]_F \cdot [\phi_h]_F \,\mathrm{d}a \quad \text{auf } F = \partial K \cap \partial K_F$$
$$0 = \int_F (u - u_\mathrm{D}) \phi_h \,\mathrm{d}a = \int_F (u - u_\mathrm{D}) \kappa \nabla \phi_h \cdot n \,\mathrm{d}a \quad \text{auf } F \subset \Gamma_\mathrm{D}.$$

Zusammen ergibt sich

$$a_h^\mathrm{SIP}(u_h, \phi_h) = b_h^\mathrm{SIP}(\phi_h)$$

mit

$$\begin{aligned} a_h^\mathrm{SIP}(u_h, \phi_h) &= \sum_K \int_K \kappa \nabla u_h \cdot \nabla \phi_h \,\mathrm{d}x \\ &\quad - \sum_{F \subset \Omega} \left[\int_F \{\kappa \nabla u_h\} \cdot [\phi_h]_F \,\mathrm{d}x + \int_F [u_h]_F \cdot \{\kappa \nabla \phi_h\}_F \,\mathrm{d}x \right] \\ &\quad - \sum_{F \subset \Gamma_\mathrm{D}} \left[\int_F \kappa \nabla u_h \cdot n \phi_h \,\mathrm{d}a + \int_F u_h \kappa \nabla \phi_h \cdot n \,\mathrm{d}a \right] \\ &\quad + \sum_{F \subset \Omega} \gamma_F \int_F [u_h]_F \cdot [\phi_h]_F \,\mathrm{d}a + \sum_{F \subset \Gamma_\mathrm{D}} \gamma_F \int_F u_h \phi_h \,\mathrm{d}a\,, \\ b_h^\mathrm{SIP}(\phi_h) &= \int_{\Gamma_\mathrm{N}} g_N \phi_h \,\mathrm{d}x \qquad - \sum_{F \subset \Gamma_\mathrm{D}} \int_F u_\mathrm{D} \kappa \nabla \phi_h \cdot n \,\mathrm{d}a + \sum_{F \subset \Gamma_\mathrm{D}} \gamma_F \int_F u_\mathrm{D} \phi_h \,\mathrm{d}a\,. \end{aligned}$$

Konsistenz:
Für die exakte Lösung $u \in C(\overline{\Omega})$ mit $q \cdot n|_K = q \cdot n|_{K_F}$ ($\to q = -\kappa \nabla u$) und Testfuktionen $\phi_h \in Q_h$ mit $\phi_h = 0$ auf Γ_D gilt:

$$a_h^\mathrm{SIP}(u_h, \phi_h) = a(u, \phi_h) = \int_\Omega \kappa \nabla u \cdot \nabla \phi_h \,\mathrm{d}x$$
$$b_h^\mathrm{SIP} = b(\phi_h) = \int_{\Gamma_\mathrm{N}} g_\mathrm{N} \phi_h \,\mathrm{d}a$$

Stabilität:
Sei $u_h \in Q_h = \prod \mathbb{P}_k(K)$ Lösung von $a_h^\mathrm{SIP}(u_h, \phi_h) = b_h^\mathrm{SIP}(\phi_h)\ \forall \phi_h \in Q_h$. Dann gilt

$$\|u_h\|_\mathrm{DG} \le c$$
$$\|u_h\|_\mathrm{DG}^2 \le u_h^\mathrm{SIP}(u_h, u_h) = b_h^\mathrm{SIP}(u_h) \le c \|u_h\|_\mathrm{DG}$$

(7.2) Lemma
Sei $\|\kappa \nabla \phi_h\|_{\partial K}^2 \le \frac{\gamma_h}{4} \int_K \kappa \nabla \phi_h \cdot \nabla \phi_h \,\mathrm{d}x$ *mit* $\gamma_F \equiv \gamma_h\ \forall K \in \mathcal{K}$. *Dann gilt*

$$a_h^\mathrm{SIP}(u_h, \phi_h) \ge \frac{1}{2} \left(\sum_K \int_K \kappa \nabla \phi_h \cdot \nabla \phi_h \,\mathrm{d}x + \sum_{F \in \mathcal{F}_h \setminus \Gamma_\mathrm{N}} \gamma_h \int_F [\phi_h]_F^2 \,\mathrm{d}a \right) =: \|\phi_h\|_\mathrm{DG}^2$$

Beweis.

$$\int_F \{\kappa\nabla\phi_h\}[\phi_h]\,\mathrm{d}a \le \left(\gamma_h^{-1}\|\{\kappa\nabla\phi_h\}\|_F\right)\left(\gamma_h\|[\phi_h]\|_F\right)$$
$$\le \frac{1}{\gamma_h}\|\{\kappa\nabla\phi_h\}\|_F^2 + \frac{\gamma_h}{4}\|[\phi_h]\|_F^2$$

$$\begin{aligned}\Rightarrow\ a_h^{\text{SIP}}(\phi_h,\phi_h) &= \sum \int_K \kappa\nabla\phi_h\cdot\nabla\phi_h\,\mathrm{d}x + \sum_{F\not\subseteq\Gamma_{\mathrm{N}}} \gamma_h\|[\phi_h]\|_F^2 - 2\int_F \{\kappa\nabla\phi_h\}[\phi_h]\,\mathrm{d}a \\ &\ge \sum \int_K \kappa\nabla\phi_h\cdot\nabla\phi_h\,\mathrm{d}x + \sum_{F\not\subseteq\Gamma_{\mathrm{N}}} \gamma_h\|[\phi_h]\|_F^2 - \frac{2}{\gamma_h}\|\{\kappa\nabla\phi_h\}\|^2 \\ &\quad - \frac{\gamma_h}{2}\|[\phi_h]\|\end{aligned}$$

□

Folgerungen:

1) $\gamma_F = O\left(\frac{\|\kappa\|_{\infty,K}}{h_F}\right)$

2)

$$\begin{aligned} a_h^{\text{SIP}}(\phi_h,\phi_h) = 0 &\Rightarrow \|\phi_h\|_{\text{DG}} = 0 \\ &\Rightarrow \nabla\phi_h \equiv 0 \ \text{ auf } K \\ &\Rightarrow \phi_h \ \text{ konstant in } K \\ &\Rightarrow \phi_h \ \text{ konstant in } \Omega \end{aligned}$$

Falls $\Gamma_{\mathrm{D}} \neq \emptyset \ \Rightarrow\ \phi_h = 0$ auf Γ_{D}, d.h. $\phi_h \equiv 0$ in Ω.

$$\Rightarrow \underline{A}[n,k] = a_h^{\text{SIP}}(\phi_h,\phi_h) \text{ symmetrisch positiv definit.}$$

Ergebnis: Das SIP-DG-Verfahren ist mit $u_h \in Q_h$:

$$a_h^{\text{SIP}}(u_h,\phi_h) = b_h^{\text{SIP}}(\phi_h)\ \forall\phi_h \in Q_h \text{ wohldefiniert.}$$

Beachte: Die konforme Approximation $u_h \in V_h$ konvergiert für $h \longrightarrow 0$ als $u_h \xrightarrow{\text{Cauchy-Folge}} u$ bzgl. $\|\nabla\phi\|_1 = (\|\nabla\phi\|_0^2 + \|\phi\|_0^2)$.
Die nichtkonforme Approximation $u_h \in Q_h$ konfergiert $u_h \xrightarrow{\text{Cauchy-Folge}} u$ bzgl. $\|\cdot\|_{\text{DG}}$, aber i.A. gilt keine Konvergenz für $\kappa\nabla u_h\cdot n$.

Anwendung auf die Reaktions-Diffusionsgleichung

Betrachte stationären Fall:

$$\begin{aligned} -\operatorname{div}\kappa_c\nabla c + q\cdot\nabla c + rc &= f && \text{in } \Omega \\ c &= c_{\mathrm{D}} && \text{auf } \Gamma_{\mathrm{D}} = \Gamma_{\text{in}} \\ \kappa_c\nabla c + q\cdot n\, c &= 0 && \text{auf } \Gamma_{\mathrm{R}} = \Gamma_{\text{out}} \end{aligned}$$

Wir berechnen dann in jedem Zeitschritt die Bilinearform

$$c_h \in Q_h : a_h(c_h,\phi_h) = b_h(\phi_h)\ \forall\phi_h \in Q_h$$

mit

$$\begin{aligned} a_h(c_h,\phi_h) = a_h^{\text{SIP}}(c_h,\phi_h) + \sum_K \Bigg[(q\cdot\nabla c_h,\phi_h)_K - \sum_{F\not\subseteq\Gamma_{\mathrm{D}}} \int_F (\psi(c_h)-\psi^*(c_h))\cdot n_K\cdot\phi_h\,\mathrm{d}a \Bigg] \\ + \int_\Omega rc_h\phi_h\,\mathrm{d}x + \int_{\Gamma_{\mathrm{R}}} q\cdot nc_h\phi_h\,\mathrm{d}a \end{aligned}$$

und

$$b_h(\phi_h) = b_h^{\text{SIP}}(\phi_h) + \int_\Omega f\,\phi_h\,\mathrm{d}x + \int_{\Gamma_{\mathrm{D}}} q\cdot nc_{\mathrm{D}}\phi_h\,\mathrm{d}a.$$

8 UQ (Uncertainty-Quantification) für die Diffusionsgleichung

Aufgabe: Bei unsicheren Daten für die Permeabilität bestimme den Erwartungswert für den Fluss $q = -\kappa \nabla u$ bzw. der Messgröße $J(q) = \int_\Gamma q \cdot n \, \mathrm{d}a$.
Mathematisches Modell:

- $\mathcal{B} \subset \mathbb{R}^D$ Rechengebiet
- Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$
- $\mathbb{P}(A)$: Wahrscheinlichkeit für Ereignis $A \in \mathcal{A}$, $\mathbb{P}(\Omega) = 1$.
- für $x \in \mathcal{B}$ sei $\kappa(\cdot, x) : \Omega \to [\kappa_0, \kappa_1]$ Zufallsvariable
- zu einer Realisierung $\kappa(\omega, \cdot)$ mit $\omega \in \Omega$ bestimme $u_h(\omega) \in V_h(u_\mathrm{D})$ durch

$$\int_{\mathcal{B}} \kappa(\omega, x) \nabla u_h(\omega, x) \cdot \nabla \phi_h(x) \, \mathrm{d}x = \int_{\mathrm{N}} g_\mathrm{N} \phi_h \, \mathrm{d}a$$

Monte-Carlo-Approximation

Seien κ^m für $m = 1, ..., M$ gleichverteilte, stochastisch unabhängige Realisationen (i.i.d.). Berechne

$$\hat{u}_h^m \in V_h(u_\mathrm{D}) : \int_{\mathcal{B}} \kappa^m \nabla \hat{u}_h^m \cdot \nabla \phi_h \, \mathrm{d}x = \int_{\mathrm{N}} g_\mathrm{N} \phi_h \, \mathrm{d}a \,, \qquad \phi_h \in V_h(0)$$

und definiere

$$E_M(u_h) = \frac{1}{M} \sum_{m=1} \hat{u}_h^m \,.$$

Beobachtung: Keine Konvergenz bei Realisierungen κ^m unabhängig in $x \in \mathcal{B}$.
Lösung: Definiere den Erwartungswert

$$\overline{\kappa}(x) = \mathbb{E}[\kappa(\cdot, x)] = \int_{\mathcal{B}} \kappa(\omega, x) \mathbb{P}(\,\mathrm{d}\omega)$$

und Kovarianz

$$C(x, y) = \mathbb{E}[(\kappa(\cdot, x) - \overline{\kappa}(x))(\kappa(\cdot, y) - \overline{\kappa}(y))]$$

$$\Rightarrow \int_{\mathcal{B}} C(x, y) \varphi_k(y) \, \mathrm{d}y = \lambda_k \varphi_k(x)$$

$$\Rightarrow \kappa(\omega, x) = \overline{\kappa}(x) + \sum_{k=1}^{\infty} \xi_k(\omega) \varphi_k(x)$$

mit ξ_k stochastisch unabhängigen reellen Zufallsvariablen.

(8.1) Lemma
Sei $\|v\|_V = \sqrt{\|v\|_{\mathcal{B}}^2 + \|\nabla v\|_{\mathcal{B}}^2}$ *und zu* $u_h \in L_2(\Omega, V_h)$ *definiere*

$$\|u_h\|_\Omega^2 = \mathbb{E}[\|u(\cdot,\cdot)\|_V^2] = \int\limits_\Omega \|u(\omega,x)\|_V^2 \mathbb{P}(\,\mathrm{d}\omega)\,.$$

Dann gilt für den "statistischen Fehler"

$$\|\mathbb{E}[u_h] - E_M(u_h)\|_\Omega \leq \frac{1}{\sqrt{M}}\|u_h\|_\Omega\,.$$

Regularitätsvoraussetzung:

$$\mathbb{E}[u] = \lim_{h\to 0}\mathbb{E}[u_h] \text{ existiert mit } \|\hat{u} - \hat{u}_h\|_V \leq C\,h^\alpha\,, \qquad \alpha \in (0,1].$$

(8.2) Satz
$\|\mathbb{E}[u_h] - E_M(u_h)\|_\Omega = O\left(\frac{1}{\sqrt{M}} + h^\alpha\right)$

Abschätzung für $s = 1, N_h = O(h^{-2})$ mit $N_h = \dim V_h, V_h = \operatorname{span}\{\lambda_z\}$:
$O\left(\frac{1}{\sqrt{M}}\right) = O(h) = O\left(\frac{1}{\sqrt{N_h}}\right)$
$\implies$ Monte-Carlo benötigt so viele Realisierungen wie $\dim V_h$.

<u>Idee</u>: $u_L = u_0 + \sum\limits_{l=1}^{L} u_l - u_{l-1}$ mit $u_l \in V_l(u_\mathrm{D})$ und $V_l = V_{h_l}, h_l = 2^{-l}h_0$.

Multi-Level-Monte-Carlo-Approxomation

Berechne Realisierungen $\hat{\kappa}_l^m$ für $m = 1,...,M,\ l = 0,...,L$ und FE-Lösung $\hat{u}_l^m \in V_l(u_\mathrm{D})$.
Definiere außerdem

$$E^L(u) = \frac{1}{M_0}\sum_{m=1}^{M_0}\hat{u}_0^m + \sum_{l=1}^{L}\frac{1}{M_l}\sum_{m=1}^{M_l}(\hat{u}_l^m - \hat{u}_{l-1}^m) \text{ mit } M_0 \geq M_1 \geq M_2...$$

Regularitätsvoraussetzung:

$$\mathbb{E}[u_h - u_{2h}] \leq C\,h^\beta\,, \qquad \beta \in (0,1].$$

(8.3) Satz
$\|\mathbb{E}[u_h] - E^L(u)\|_\Omega = O\left(h_L^\alpha + \sum\limits_{l=1}^{L}\frac{h_l^\beta}{\sqrt{M_l}}\right)$

Bemerkung
Aufwand $O(N_L(\log N_L)^4)$ *mit* $N_L = \dim V_L$ *und optimaler Wahl von* M_l.