ulm university universität
uulm

Universität Ulm
Fakultät für Mathematik und
Wirtschaftswissenschaften

Vorlesungsmanuskript

Numerik von Partiellen Differentialgleichungen II

Dozent

Prof. Dr. Karsten Urban

Universität Ulm, Institut für Numerische Mathematik
2012

Inhaltsverzeichnis

Einleitung

Inhalt: Reduzierte Basismethoden (RBM)

Idee: PDE hängt von „Parametern“ ab, z.B.

- rechte Seite (Kraft)
- Leitfähigkeit/Porosität $\rightarrow$ Koeffizienten
- Geometrie (Länge, Form, ...)

$\mu \in \mathcal{D} \subset \mathbb{R}^P$ Parameter, $u = u(\mu)$ Lösung.

Ziel: häufige oder extrem schnelle Approximation von $u(\mu)$:

- häufig z.B. bei Optimierung: $J(\mu) = \ell(u(\mu)) \overset{\mu \in \mathcal{D}}{\longrightarrow} \text{Max}$
  „multi-query-context“ (Alinghi, Bypass, Voith-Schneider-Propeller, ...)
- extrem schnell bei Echtzeit-Berechnungen (Ferrari, Smartphone, ...).

Methode: Trennung von offline- und online-Berechnungen

- offline:
  - Auswertung verschiedener $u(\mu_i)$, $1 \leq i \leq N$, $\mu_i \in \mathcal{D}$
  - Bestimmung eines niedrig-dimensionalen Raumes samt Basis („Reduzierte Basis“)
  - Vorabberechnung möglichst vieler Terme (Matrizen, Vektoren, ...)
- online: für neuen Parameter $\mu \notin \{\mu_1, \ldots, \mu_N\}$ approximiere $u(\mu)$ durch reduzierte Approximation $u_N(\mu)$.

Mathematische Fragestellungen:

- Größe von N? $\rightarrow$ Für welche Probleme „geht“ RBM?
- Bestimmung der μ_i?
- Approximation $\|u(\mu) - u_N(\mu)\| \leq ?$, $|\ell(u(\mu)) - \ell(u_N(\mu))| \leq ?$
- Effizienz, Implementierung
- Berechnung von Fehlerschranken
- Warum ist das besser als Interpolation?
- ...

1 Parametrische PDEs (PPDEs)

Seien X, Y zwei Hilbert-Räume (Funktionenräume) und $\mu = \{\mu_1, \dots, \mu_P\}^T \in \mathcal{D} \subset \mathbb{R}^P$ ein Parameter.

Definition:
(a) Eine Form $g : Y \times \mathcal{D} \to \mathbb{R}$ mit $g(\cdot, \mu) \in Y'$ für alle $\mu \in \mathcal{D}$ heißt parametrische Linearform.
(b) Eine Form $b : X \times Y \times \mathcal{D} \to \mathbb{R}$ heißt parametrische Bilinearform falls $b(\cdot, \cdot\,; \mu)$ für alle $\mu \in \mathcal{D}$ eine Bilinearform ist.

Ein parametrisches Variationsproblem lautet dann

$$\text{(1.1)} \qquad \textit{bestimme } \; u(\mu) \in X : \qquad b(u(\mu), v; \mu) = g(v, \mu) \; \forall v \in Y.$$

Beispiel 1.1:
Sei $\Omega = \bigcup_{i=1}^P \Omega_i$ mit Ω_i polygonal.

- $\mathcal{D} = [\mu_{min}, \mu_{max}]^P \subset \mathbb{R}^P$ *mit* $0 < \mu_{min} < \mu_{max}$ *(Wärmeleitungskoeffizienten)*
- $\mu = (\mu_1, \dots, \mu_P)^T \in \mathbb{R}^P$, $\alpha(x; \mu) := \sum_{i=1}^P \mu_i \mathbb{1}_{\Omega_i}(x)$
- *Wärmeleitung ohne Quellterm:* $-\nabla \cdot (\alpha(x; \mu) \nabla u(x)) = 0$ *in* Ω
- *Randbedingungen:* $\Gamma := \partial\Omega = \Gamma_D \,\dot\cup\, \Gamma_{N,0} \,\dot\cup\, \Gamma_{N,1}$
  - *Kühlung auf Null:* $u = 0$ *auf* Γ_D
  - *Isolation:* $(\alpha(x; \mu) \nabla u(x)) \cdot n(x) = 0$, $x \in \Gamma_{N,0}$
  - *Ausstrom:* $(\alpha(x; \mu) \nabla u(x)) \cdot n(x) = 1$, $x \in \Gamma_{N,1}$
- $X = H^1_{\Gamma_D}(\Omega) := \{u \in H^1(\Omega) : u|_{\Gamma_D} \equiv 0\} = Y$,

$$b(v, w, \mu) = \int_\Omega \alpha(x; \mu) \nabla v(x) \nabla w(x)\, dx, \qquad g(v; \mu) \equiv g(v) = \int_{\Gamma_{N,1}} v(x)\, dx.$$

Einfachste Form:

- $\Omega = (0, 1)^2$
- $\Omega_1 = (0, 1) \times (0, \frac{1}{2})$, $\Omega_2 = (0, 1) \times (\frac{1}{2}, 1)$
- $\Gamma_D = [0, 1] \times \{0\}$, $\Gamma_{N,1} = [0, 1] \times \{1\}$
- $\mu_1 = 1$ (Normierung) $\rightsquigarrow \mathcal{D} = [\mu_{\min}, \mu_{\max}] \subset \mathbb{R}^1$.

Formal können wir $u(\mu) = L^{-1}(\mu) g(\mu)$ mit $L(\mu) : X \to Y'$ und $g(\mu) \equiv g(\cdot; \mu) \in Y'$ schreiben. Der Satz von *Babuška* und *Aziz* gibt Auskunft über Korrekt-Gestelltheit von (1.1).

Satz 1.2 (*Babuška, Aziz*, 1972)**:**
Die Abbildung $L(\mu)$ ist genau dann ein Isomorphismus, falls

$$\gamma(\mu) := \sup_{w\in X}\sup_{v\in Y}\frac{b(w,v;\mu)}{\|w\|_X\cdot\|v\|_Y} < \infty \qquad \textit{(Stetigkeit)} \tag{1.2}$$

$$\beta(\mu) := \inf_{w\in X}\sup_{v\in Y}\frac{b(w,v;\mu)}{\|w\|_X\cdot\|v\|_Y} > 0 \qquad \textit{(Inf-Sup)} \tag{1.3}$$

$$\forall v\in Y\ \exists u\in X \text{ mit } b(u,v;\mu)\neq 0 \qquad \textit{(Nicht-Trivialität).} \tag{1.4}$$

Nehmen wir an, dass man (1.4) für alle $\mu\in\mathcal{D}$ zeigen kann, dann sind $\beta(\mu)$ und $\gamma(\mu)$ Indikatoren für die Korrekt-Gestelltheit von (1.1) für gegebenes $\mu\in\mathcal{D}$. Weiterhin gilt

$$\beta(\mu)\,\|u(\mu)\|_X \overset{(1.3)}{\leq} \sup_{v\in Y}\frac{b(u(\mu),v;\mu)}{\|v\|_Y} \overset{(1.1)}{=} \sup_{v\in Y}\frac{g(v;\mu)}{\|v\|_Y} = \|g(\mu)\|_{Y'},$$

also ist $\beta(\mu)$ Stabilitäts-Parameter. Die $\mu-$Abhängigkeit ist ein Problem, vor allem online. Daher:

$$\text{Annahme :}\qquad \beta_0\leq\beta(\mu)\ \forall\mu\in\mathcal{D},\qquad \gamma_\infty\geq\gamma(\mu)\ \forall\mu\in\mathcal{D}. \tag{1.5}$$

Bemerkung 1.3:
Ist $X = Y$, dann heißt b koerziv, falls

$$\alpha_0 := \min_{\mu\in\mathcal{D}}\alpha(\mu) > 0,\quad \alpha(\mu) := \inf_{w\in X}\frac{b(w,w,\mu)}{\|w\|_X^2} \tag{1.6}$$

Ziel: Effiziente online-Berechnung von $\beta(\mu), \beta_0$ als

- Stabilitäts-Indikator
- Fehlerschranke (später)

Bemerkung 1.4:
Sei $X = Y$. Dann heißt die Bilinearform b symmetrisch, falls

$$b(w,v;\mu) = b(v,w;\mu)\ \forall\, v,w\in X\quad\forall\mu\in\mathcal{D}.$$

Der symmetrische bzw. anti-symmetrische Teil von b ist definiert als

$$\begin{aligned} b_S(w,v;\mu) &:= \tfrac{1}{2}\left\{b(w,v;\mu)+b(v,w;\mu)\right\},\\ b_A(w,v;\mu) &:= \tfrac{1}{2}\left\{b(w,v;\mu)-b(v,w;\mu)\right\},\end{aligned}$$

also $b_A(w,v;\mu) = -b_A(v,w;\mu)$ und $b(w,v;\mu) = b_S(w,v;\mu) + b_A(w,v;\mu)$. Wegen $b_A(w,w;\mu) = 0$ gilt:

$$\alpha(\mu) = \inf_{w\in X}\frac{b_S(w,w;\mu)}{\|w\|_X^2},$$

also ist $\alpha(\mu)$ ein Rayleigh-Quotient, d.h. ein minimaler Eigenwert. Ist die Dimension von X tatsächlich endlich, dann kann dieser approximiert werden.

Definition 1.5:
Der supremierende Operator $T_\mu : X \to Y,\ \mu \in \mathcal{D}$ *ist definiert durch*

$$T_\mu w := \arg\sup_{v\in Y} \frac{b(w,v;\mu)}{\|v\|_Y}, \qquad w \in X. \tag{1.7}$$

Die Existenz von T_μ *folgt wegen der Stetigkeit von* b *und wegen* $b(\mu) \geq 0$*: Angenommen, es gelte* $\beta(\mu) < 0$*, dann sei* $(v_n)_{n\in\mathbb{N}} \subset Y$ *mit*

$$\sup_{v\in Y} \frac{b(w,v;\mu)}{\|w\|_X \cdot \|v\|_Y} = \lim_{n\to\infty} \frac{b(w,v_n;\mu)}{\|w\|_X \cdot \|v_n\|_Y} < 0.$$

Für $-v_n$ *ist dieser Term positiv.* $\lightning$

Bemerkung 1.6:
Der Operator T_μ *ist explizit gegeben durch*

$$(T_\mu w\,, v)_Y = b(w,v;\mu) \qquad \forall v \in Y. \tag{1.8}$$

Beweis. Es gilt $b(w,v;\mu) = (T_\mu w\,, v)_Y \overset{\substack{Cauchy-\\ Schwartz}}{\leq} \|T_\mu w\|_Y \cdot \|v\|_Y$. □

Für festes $w \in$ sei $\ell_w \in Y'$ definiert als $\ell_w(v) := b(w,v;\mu),\ v \in Y$.

$$\begin{aligned} \Rightarrow \|\ell_w\|_{Y'} &= \sup_{v\in Y} \frac{\ell_w(v)}{\|v\|_Y} = \sup_{v\in Y} \frac{b(w,v;\mu)}{\|v\|_Y} = \frac{b(w,T_\mu w;\mu)}{\|T_\mu w\|_Y} \\ &\overset{(1.8)}{=} \frac{(T_\mu w\,, T_\mu w)_Y}{\|T_\mu w\|_Y} = \|T_\mu w\|_Y. \end{aligned}$$

Also: Supremierender Operator $\sim$ Dualnorm des Funktionals!

Lemma 1.7:
Es gilt

$$\beta(\mu) = \inf_{w\in X} \frac{\|T_\mu w\|_Y}{\|w\|_X}.$$

Beweis.

$$\begin{aligned} \beta(\mu) &= \inf_{w\in X} \sup_{v\in Y} \frac{b(w,v;\mu)}{\|w\|_X \cdot \|v\|_Y} \overset{(1.7)}{=} \inf_{w\in X} \frac{b(w,T_\mu w;\mu)}{\|w\|_X \cdot \|T_\mu w\|_Y} \\ &\overset{(1.8)}{=} \inf_{w\in X} \frac{(T_\mu w\,, T_\mu)_Y}{\|w\|_X \cdot \|T_\mu w\|_Y} = \inf_{w\in X} \frac{\|T_\mu w\|_Y}{\|w\|_X} \end{aligned}$$

□ □

Also kann $\beta(\mu)$ als minimaler Rayleigh-Quotient ($\sim$ Eigenwert) aufgefasst werden:

$$\widetilde{r}(A, w) = r(A^T A, w) = \frac{w^T A^T A w}{w^T w} = \frac{\|Aw\|^2}{\|w\|^2},$$

also

$$\beta(\mu) = \sqrt{\inf_{w \in X} \widetilde{r}(T_\mu, w)}.$$

Bemerkung 1.8:
Analog zu Lemma 1.7 gilt $\gamma(\mu) = \|T_\mu w\| = \sup_{w \in X} \frac{\|T_\mu w\|_Y}{\|w\|_X}$, *also*

$$\gamma(\mu) = \sqrt{\sup_{w \in X} \widetilde{r}(T_\mu, w)}.$$

Damit sind also $\beta(\mu)$ und $\gamma(\mu)$ mit dem gleichen (verallgemeinerten) Eigenwert-Problem (bzgl. T_μ) assoziiert:

Bestimme $\chi_i(\mu) \in X,\ \nu_i(\mu) \in \mathbb{R},\ i = 1, \ldots, \dim(X)$ mit

$$(T_\mu \chi_i(\mu), T_\mu w)_Y = \nu_i(\mu) (\chi_i(\mu), w)_X \qquad \forall w \in X \tag{1.9}$$

und $\|\chi_i(\mu)\|_X = 1,\ (\chi_i(\mu), \chi_j(\mu))_X = \delta_{ij}$. Also

$$\beta(\mu) = \sqrt{\nu_1(\mu)}, \qquad \gamma(\mu) = \lim_{i \to \infty} \sqrt{\nu_i(\mu)}, \tag{1.10}$$

wenn (o.B.d.A. für koerzives b) $0 \leq \nu_1(\mu) \leq \ldots \leq \nu_{\dim(X)}(\mu)$.

2 Die „wahre“ Approximation

Die Hilbert-Räume bei PDEs (X, Y) sind in der Regel $\infty-$dimensional. In der offline-Phase verwendet man daher eine „hinreichend feine“ Diskretisierung, die „truth“. Mit „hinreichend“ ist gemeint, dass exakte und numerische Lösung im Rahmen der gewünschten Genauigkeit nicht unterscheidbar sind ($\forall \mu \in \mathcal{D}$!).

Seien also

$$X^{\mathcal{N}} \subset X, \quad Y^{\mathcal{N}} \subset Y, \quad \dim(X^{\mathcal{N}}) = \dim(Y^{\mathcal{N}}) = \mathcal{N} \tag{2.1}$$

und

$$u^{\mathcal{N}}(\mu) \in X^{\mathcal{N}} : \quad b(u^{\mathcal{N}}(\mu), v^{\mathcal{N}}; \mu) = g(v^{\mathcal{N}}; \mu) \qquad \forall v^{\mathcal{N}} \in Y^{\mathcal{N}}. \tag{2.2}$$

Mit $\Phi^{\mathcal{N}} := \left\{\varphi_1^{\mathcal{N}}, \ldots, \varphi_{\mathcal{N}}^{\mathcal{N}}\right\}$, $X^{\mathcal{N}} = \operatorname{span} \Phi^{\mathcal{N}}$ bezeichnen wir eine Basis von $X^{\mathcal{N}}$ sowie $\Psi^{\mathcal{N}} := \left\{\psi_1^{\mathcal{N}}, \ldots, \psi_{\mathcal{N}}^{\mathcal{N}}\right\}$, $Y^{\mathcal{N}} = \operatorname{span} \Psi^{\mathcal{N}}$.

Wir erhalten analog zu (1.2), (1.3) und Bemerkung 1.4:

$$\begin{aligned}
\alpha^{\mathcal{N}}(\mu) &:= \inf_{w \in X^{\mathcal{N}}} \frac{b_S(w, w; \mu)}{\|w\|_{X^{\mathcal{N}}}^2} && (X^{\mathcal{N}} = Y^{\mathcal{N}}), \\
\beta^{\mathcal{N}}(\mu) &:= \inf_{w \in X^{\mathcal{N}}} \sup_{v \in Y^{\mathcal{N}}} \frac{b(w, v; \mu)}{\|w\|_{X^{\mathcal{N}}} \cdot \|v\|_{Y^{\mathcal{N}}}}, \\
\gamma^{\mathcal{N}}(\mu) &:= \sup_{w \in X^{\mathcal{N}}} \sup_{v \in Y^{\mathcal{N}}} \frac{b(w, v; \mu)}{\|w\|_{X^{\mathcal{N}}} \cdot \|v\|_{Y^{\mathcal{N}}}},
\end{aligned}$$

wobei die Normen $\|\cdot\|_{X^{\mathcal{N}}}$, $\|\cdot\|_{Y^{\mathcal{N}}}$ noch gewählt werden können.

Kern-Annahme:

$$\max_{\mu \in \mathcal{D}} \inf_{w \in X^{\mathcal{N}}} \|u(\mu) - w^{\mathcal{N}}\|_X \to 0 \quad \text{mit } \mathcal{N} \to \infty \tag{2.3}$$

Dies kann z.B. mittels einer feinen FE-Diskretisierung sichergestellt werden. Wegen (2.2) ist $u^{\mathcal{N}}(\mu)$ die (Petrov-)Galerkin-Projektion von $u(\mu)$ auf $X^{\mathcal{N}}$. Konvergenztheorie wie in Numerik von PDEs I. Das diskrete System zu (2.2) lautet dann

$$\underline{B}^{\mathcal{N}}(\mu)\underline{u}^{\mathcal{N}}(\mu) = \underline{g}^{\mathcal{N}}(\mu) \tag{2.4}$$

mit

$$\underline{B}^{\mathcal{N}}(\mu) = \left[b(\varphi_i^{\mathcal{N}}, \psi_j^{\mathcal{N}}; \mu)\right]_{i,j=1,\dots,\mathcal{N}}, \qquad \underline{g}^{\mathcal{N}}(\mu) = \left[g(\psi_j^{\mathcal{N}}; \mu)\right]_{j=1,\dots,\mathcal{N}}$$

und $u^{\mathcal{N}}(\mu) = \sum_{i=1}^{\mathcal{N}} u_i^{\mathcal{N}}(\mu)\varphi_i^{\mathcal{N}}$, $\underline{u}^{\mathcal{N}}(\mu) = \left[u_1^{\mathcal{N}}(\mu), \dots, u_{\mathcal{N}}^{\mathcal{N}}(\mu)\right]^T$. Numerische Lösung mittels Standard-Techniken, Aufwand ist bestenfalls $\mathcal{O}(\mathcal{N})$ (optimal), aber $\mathcal{N}$ sehr groß!

Probleme:

- $\mathcal{N} >> 1$, (2.4) kann auch nichtlinear sein
- multi-query: Aufstellung und Lösung von (2.4) für viele μ
- realtime: nicht möglich

Eine entscheidende Annahme ist daher die folgende:

Definition 2.1:

(a) Eine parametrische Linearform heißt affin im Parameter, falls es Funktionen $\theta_g^q : \mathcal{D} \to \mathbb{R}$, $g^q : Y \to \mathbb{R}$, $1 \le q \le Q_g$ gibt mit

$$g(v; \mu) = \sum_{q=1}^{Q_g} \theta_g^q(\mu) g^q(v), \qquad v \in Y,\ \mu \in \mathcal{D}. \tag{2.5}$$

(b) Eine Bilinearform b heißt affin im Parameter, falls es Funktionen $\theta_b^q : \mathcal{D} \to \mathbb{R}$, $b^q : X \times Y \to \mathbb{R}$, $1 \le q \le Q_b$ gibt mit

$$b(v, w; \mu) = \sum_{q=1}^{Q_b} \theta_b^q(\mu) b^q(v, w). \tag{2.6}$$

Beispiel 2.2:

$$\begin{aligned}
&(a) \quad g(v; \mu) = \mu(f\,, v)_{L_2(\Omega)} \\
&(b) \quad b(v, w; \mu) = (\nabla v\,, \nabla w)_{L_2(\Omega)} + \mu_1(\underline{\beta} \cdot \nabla v\,, w)_{L_2(\Omega)} + \mu_2(v\,, w)_{L_2(\Omega)} \\
&(c) \quad b(v, w; \mu) = \int_0^T (\dot{v}(t)\,, w(t))_{L_2(\Omega)} dt + \mu \int_0^T (\nabla v(t)\,, \nabla w(t))_{L_2(\Omega)} dt \\
&(d) \quad b(v, w; \mu) = \int_\Omega \alpha(x; \mu) \nabla v(x) \nabla w(x) dx = \sum_{i=1}^{P} \mu_i \int_{\Omega_i} \nabla v(x) \cdot \nabla w(x) dx \\
&(e) \quad b(v, w; \mu) = \int_\Omega \sin(2\pi\mu x) \cdot \nabla v(x) \cdot \nabla w(x) dx
\end{aligned}$$

Beispiele (a) - (d) sind affin im Parameter, (e) nicht.

Unter den Voraussetzungen von Definition 2.1 gilt

$$\underline{B}^{\mathcal{N}}(\mu) = \sum_{q=1}^{Q_b} \theta_b^q(\mu)\mathbb{B}^{\mathcal{N},q}, \qquad \mathbb{B}^{\mathcal{N},q} = \left[b^q(\varphi_i^{\mathcal{N}}, \psi_i^{\mathcal{N}})\right]_{i,j=1,\dots,\mathcal{N}}$$

$$\underline{g}^{\mathcal{N}}(\mu) = \sum_{q=1}^{Q_g} \theta_g^q(\mu)\overline{g}^{\mathcal{N},q}, \qquad \overline{g}^{\mathcal{N},q} = \left[g^q(\psi_j^{\mathcal{N}})\right]_{j=1,\dots,\mathcal{N}}$$

Als nächstes betrachten wir die Auswirkungen auf die Berechnung von $\alpha^{\mathcal{N}}, \beta^{\mathcal{N}}, \gamma^{\mathcal{N}}$:

Bemerkung 2.3:
Seien $\Phi^{\mathcal{N}}, \Psi^{\mathcal{N}}$ Orthonormalbasen für $X^{\mathcal{N}}, Y^{\mathcal{N}}$, $\underline{B}^{\mathcal{N}}(\mu) = [b_{ij}(\mu)]_{ij} = \left[b(\varphi_i^{\mathcal{N}}, \psi_j^{\mathcal{N}}; \mu)\right]_{ij}$, dann gilt für

$$v^{\mathcal{N}} = \sum_{i=1}^{\mathcal{N}} v_i \varphi_i^{\mathcal{N}} \in X^{\mathcal{N}}, \qquad w^{\mathcal{N}} = \sum_{j=1}^{\mathcal{N}} w_j \psi_j^{\mathcal{N}} \in Y^{\mathcal{N}}$$

offenbar $\|v^{\mathcal{N}}\|_X^2 = \sum_{i=1}^{\mathcal{N}} v_i^2$, $\|w^{\mathcal{N}}\|_Y^2 = \sum_{i=1}^{\mathcal{N}} w_i^2$ sowie

$$b(v^{\mathcal{N}}, w^{\mathcal{N}}; \mu) = \sum_{i,j=1}^{\mathcal{N}} v_i b_{ij}(\mu) w_j = \underline{v}^T \underline{B}^{\mathcal{N}}(\mu)\underline{w}$$

mit den Koeffizienten-Vektoren $\underline{v} = (v_i)_i, \underline{w} = (w_j)_j \in \mathbb{R}^{\mathcal{N}}$ und damit

$$\beta(\mu) = \inf_{\underline{w}\in\mathbb{R}^{\mathcal{N}}} \sup_{\underline{v}\in\mathbb{R}^{\mathcal{N}}} \frac{\underline{v}^T \underline{B}^{\mathcal{N}}(\mu)\underline{w}}{\|\underline{v}\| \cdot \|\underline{w}\|},$$

also der kleinste Singulärwert von $\underline{B}^{\mathcal{N}}(\mu)$.
Beachte: $\underline{B}^{\mathcal{N}}(\mu) = \sum_{q=1}^{Q_b} \theta_b^q(\mu)\underline{\mathbb{B}}^{\mathcal{N},q}$.

3 RB-Approximation

- Suche $u^{\mathcal{N}}(\mu)$ für viele $\mu \in \mathcal{D} \Rightarrow \mathcal{O}(|\mu|\mathcal{N})$ mindestens: viel zu teuer

- Beobachtung: Wir suchen $\{u(\mu) : \mu \in \mathcal{D}\} =: \mathcal{M}$, dazu brauchen wir nicht $X^{\mathcal{N}}$, der alle möglichen Funktionen aus X gut approximiert, sondern wir müssen die Manigfaltigkeit $\mathcal{M}$ gut approximieren.

  - $\dim(\mathcal{M}) = \dim(\mathcal{D}) \leq P$
  - Idee für Approximation: Bestimme „snapshots“ (auch „Moden“ genannt) $\xi_1 = u^{\mathcal{N}}(\mu_1), \ldots, \xi_N = u^{\mathcal{N}}(\mu_N)$ offline und definiere $X_N := \operatorname{span}\{\xi_1, \ldots, \xi_N\}$, $N << \mathcal{N},\ X_N \subset X^{\mathcal{N}}$. Suche dann für $\mu \notin \{\mu_1, \ldots, \mu_N\}$ online eine Approximation $u_N(\mu)$ von $u^{\mathcal{N}}(\mu)$ in X_N.

- Fagen:

  - Wahl der μ_i?
  - Wie groß muss N gewählt werden?
  - Fehler-Kontrolle (a priori / a posteriori)
  - online-Aufwand, Ziel $\mathcal{O}(N^*)$ unabhängig von $\mathcal{N}$!
  - wieso wählt man nicht die Interpolation $P(u^{\mathcal{N}}|\xi_1, \ldots, \xi_N)(\mu)$?

Einige Bezeichnungen/Bemerkungen:

- $S_N := \{\mu_1, \ldots, \mu_N\}$ Sample-Menge

- oft fixiert man ein $N_{\max}$ und hat $S_1 \subset S_2 \subset \ldots \subset S_{N_{\max}}$, dann erhält man eine hierarchische Approximation $X_1 \subset X_2 \subset \ldots \subset X_{N_{\max}}$

- oft werden die snapshots (offline) noch orthonormalisiert, um eine stabile Basis $\xi_1, \ldots, \xi_N$ von X_N zu erhalten

Man erhält dann das RB-Galerkin-Problem

$$\text{(3.1)} \qquad \text{Suche} \quad u_N(\mu) \in X_N : \ b(u_N(\mu), v_N; \mu) = g(v_N; \mu) \qquad \forall v_N \in Y_N,$$

wobei $Y_N \subset Y^{\mathcal{N}},\ \dim(Y_N) = N$ geeignet zu wählen ist.

Lemma 3.1:
Sei $b(\cdot,\cdot;\mu)$ s.p.d., $X = Y$, $X_N = Y_N$, dann gilt

$$\|u^{\mathcal{N}}(\mu) - u_N(\mu)\|_X \leq \frac{\gamma(\mu)}{\alpha(\mu)} \inf_{w_N \in X_N} \|u^{\mathcal{N}}(\mu) - w_N\|_X. \tag{3.2}$$

Beweis. Wie in NumPDE I:

$$\begin{aligned} \alpha(\mu)\|u^{\mathcal{N}}(\mu) - u_N(\mu)\|_X^2 &\overset{(koerz.)}{\leq} b(u^{\mathcal{N}}(\mu) - u_N(\mu), u^{\mathcal{N}}(\mu) - u_N(\mu); \mu) \\ &\overset{Gal.-Orth.}{=}_{w_N \in X_N} b(u^{\mathcal{N}}(\mu) - u_N(\mu), u^{\mathcal{N}}(\mu) - w_N; \mu) \\ &\overset{Stetigk.}{\leq} \gamma(\mu)\|u^{\mathcal{N}}(\mu) - u_N(\mu)\|_X \cdot \|u^{\mathcal{N}}(\mu) - w_N\|_X. \end{aligned}$$

□

Bemerkung 3.2:
Wegen (3.1) ist $u_N(\mu)$ die (Petrov-)Galerkin-Projektion von $u^{\mathcal{N}}(\mu)$ auf X_N - deswegen gilt die Galerkin-Orthogonalität und deswegen ist $u_N(\mu)$ viel besser als der Interpolant, vgl. Lemma 3.1.

Nun zum online System: Gesucht ist $u_N(\mu) = \sum_{i=1}^N u_{N,i}(\mu)\xi_i$ mit

$$\underline{B}_N(\mu)\,\underline{u}_N(\mu) = \underline{g}_N(\mu) \quad mit \tag{3.3}$$

$\underline{B}_N(\mu) := [b(\xi_i, \eta_j; \mu)]_{ij}$, $\underline{g}_N(\mu) = [g(\eta_j;\mu)]_j$, $\underline{u}_N(\mu) = \left[\underline{u}_{N,i}(\mu)\right]_i$ der Dimension N mit $Y_N = \text{span}\{\eta_1, \ldots, \eta_N\}$. Offenbar ist $\underline{B}_N(\mu) \in \mathbb{R}^{N\times N}$, aber wegen

$$\xi_i = \sum_{k=1}^{\mathcal{N}} c_{i,k}\varphi_k^{\mathcal{N}}, \qquad \eta_j = \sum_{\ell=1}^{\mathcal{N}} d_{j,\ell}\psi_\ell^{\mathcal{N}}$$

gilt:

$$b(\xi_i, \eta_j; \mu) = \sum_{k=1}^{\mathcal{N}}\sum_{\ell=1}^{\mathcal{N}} c_{i,k}\, d_{j,\ell}\, b(\varphi_k^{\mathcal{N}}, \psi_\ell^{\mathcal{N}}; \mu) \sim \mathcal{O}(\mathcal{N}) \;-\; \text{zu teuer!}$$

Ist nun die Bilinearform affin im Parameter, so gilt

$$b(\xi_i, \eta_j; \mu) = \sum_{q=1}^{Q_b} \theta_b^q(\mu) \underbrace{\sum_{k,\ell=1}^{\mathcal{N}} c_{i,k}\, d_{j,\ell}\, b^q(\varphi_k^{\mathcal{N}}, \psi_\ell^{\mathcal{N}})}_{=: (\mathbb{B}_N^q)_{i,j} \text{ unabh. vom Parameter, kann offline berechnet werden}}$$

(online-offline-Zerlegung), also $\underline{B}_N(\mu) = \sum_{q=1}^{Q_b} \theta_b^q(\mu)\mathbb{B}_N^q$ mit

- $\mathbb{B}_N^q \in \mathbb{R}^{N\times N}$ (voll besetzt), aber unabhängig von μ
- Berechnung von $\mathbb{B}_N^q$: $\mathcal{O}(N^2\mathcal{N})$, wenn $\Phi^{\mathcal{N}}, \Psi^{\mathcal{N}}$ lokal sind - offline

Aufwand:	**offline**	**online**
• $\mathbb{B}_N^q$	$\mathcal{O}(N^2\mathcal{N}Q_b)$	$\mathcal{O}(Q_bN^2)$
• g_N^q	$\mathcal{O}(N\mathcal{N}Q_g)$	$\mathcal{O}(Q_gN)$

- Lösen online: $\mathcal{O}(N^3)$
- Speicher: $\mathcal{O}(N^2 \cdot Q_b + N \cdot Q_q)$

Fazit:

- online-Aufwand ist unabhängig von $\mathcal{N}$!
- online: $\mathcal{O}(N^3) \Rightarrow$ wähle N möglichst klein
- offline: $\mathcal{O}(N^2\mathcal{N}Q_b) \Rightarrow \mathcal{N}, Q_b$ klein

Bemerkung 3.3:
Die affine Zerlegung kann man auch zur effizienten Berechnung der Konstanten $\alpha_N(\mu), \beta_N(\mu), \gamma_N(\mu)$ verwenden, z.B.

$$\begin{aligned}\beta_N(\mu) &= \inf_{\underline{w}\in\mathbb{R}^N} \sup_{\underline{v}\in\mathbb{R}^N} \frac{\underline{v}^T \underline{B}_N(\mu)\underline{w}}{\|\underline{v}\| \cdot \|\underline{w}\|} \qquad \text{und}\\ \underline{B}_N(\mu) &= \sum_{q=1}^{Q_b} \theta_b^q(\mu)\mathbb{B}_N^q.\end{aligned}$$

Mehr dazu später.

Definition 3.4:
(a) Sei $S_N = \{\mu_1, \dots, \mu_N\} \subset \mathcal{D}$ und $\{u(\mu_i)\}_{i=1,\dots,N}$ linear unabhängig. Dann heißt $X_N := \operatorname{span}\{u(\mu_i) : 1 \le i \le N\}$ N-dimensionaler Lagrange-RB-Raum.
(b) Sei $\mu_0 \in \mathcal{D}$ und $u(\mu)$ bzgl. μ k-fach differenzierbar in $B(\mu_0, \delta)$, $\delta > 0$. Dann heißt $X_{k,\mu_0} := \operatorname{span}\left\{\partial_\mu^\sigma u(\mu_0) : \sigma \in \mathbb{N}_0^P, |\sigma| \le k\right\}$ Taylor-RB-Raum.
(c) Eine Basis $\Phi_N = \{\xi_1, \dots, \xi_N\}$ eines RB-Raums heißt reduzierte Basis.

Bemerkung 3.5:
(a) Es gibt auch andere Arten von RB-Räumen, z.B. PoD, vgl. Kapitel 5. Viele werden über „snapshots“ $\partial_\mu^\sigma(\mu_0)$ bestimmt.
(b) Oft wird Φ_N orthonormalisiert.

Lemma 3.6:
Sei Φ_N eine Orthonormalbasis von $X_N = Y_N$ $(X = Y)$ und b koerziv. Dann gilt

$$\kappa_2(\underline{B}_N(\mu)) \le \frac{\gamma(\mu)}{\alpha(\mu)}.$$

Beweis. $\underline{B}_N(\mu)$ s.p.d. und koerziv $\Rightarrow \forall$ Eigenwerte $\lambda \in \sigma(\underline{B}_N(\mu))$ gilt $0 < \lambda_{\min} \leq \lambda_{\max}$ und $\kappa_2(\underline{B}_N(\mu)) \leq \frac{\lambda_{\max}}{\lambda_{\min}}$. Für den Eigenvektor

$$u_{\max} := \sum_{i=1}^{N} u_i \xi_i \in X_N, \qquad \underline{u}_{\max} = (u_i)_{i=1,\dots,N} \in \mathbb{R}^N$$

zu $\lambda_{\max}$ gilt

$$\begin{aligned} \lambda_{\max} \|\underline{u}_{\max}\|^2 &= \lambda_{\max} \underline{u}_{\max}^T \underline{u}_{\max} = \underline{u}_{\max}^T \underline{B}_N(\mu) \underline{u}_{\max} \\ &= b(u_{\max}, u_{\max}; \mu) \leq \gamma(\mu) \|u_{\max}\|_X^2. \end{aligned}$$

Mit $\mathbb{X} := [(\xi_i\, \xi_j)_X]_{ij} \overset{ONB}{=} I$ gilt $\|u_{\max}\|_X^2 = \underline{u}_{\max}^T \mathbb{X} u_{\max} = \|\underline{u}_{\max}\|^2$, also $\lambda_{\max} \leq \gamma(\mu)$. Analog seien $\underline{u}_{\min} \in \mathbb{R}^N$, $u_{\min} \in X_N$ definiert. Dann gilt:

$$\alpha(\mu) \cdot \|u_{\min}\|_X^2 \overset{koerz.}{\leq} b(u_{\min}, u_{\min}; \mu) = \lambda_{\min} \|\underline{u}_{\min}\|^2 = \lambda_{\min} \|u_{\min}\|_X^2.$$

□

4 Sampling-Strategien

Ziel: Möglichst gute Wahl von X_N, also S_N.

Wunsch: Mit dem Fehler der besten Approximation

$$\sigma_X(u(\mu), X_N) := \inf_{w_N \in X_N} \|u(\mu) - w_N\|_X$$

wäre

$$\sup_{\mu \in \mathcal{D}} \sigma_X(u(\mu), X_N) \leq \varepsilon \ (= \varepsilon_{\text{tol,min}})$$

optimal, wobei $\varepsilon > 0$ eine gegebene Toleranz und $N = N(\varepsilon)$ so klein wie möglich ($N \leq N_{\max}$).

Theoretisch könnten wir versuchen, X_N durch lösen eines Optimierungsproblems zu bestimmen. Dies wird in Einzelfällen möglich sein, im Allgemeinen aber schwierig. Für $P = 1$ ($\mathcal{D} \subset \mathbb{R}^P$) und koerzive Probleme kann man sehr gute S_N a priori bestimmen, im Wesentlichen Gauß-Punkte. Für $P > 1$ ist dies nicht bekannt und man unterliegt dem Fluch der Dimension (exponentieller Aufwand). Daher verwendet man „ad hoc"-Strategien, basierend auf einer „Trainings-Menge"

$$\Xi^{\text{train}} := \left\{\mu_1^{\text{train}}, \ldots, \mu_{n_{\text{train}}}^{\text{train}}\right\} \subset \mathcal{D}. \tag{4.1}$$

- möglichst „repräsentativ" für $\mathcal{D}$
- möglichst „klein" (Aufwand)
- in realistischen Anwendungen: $n_{\text{train}} \sim 10^6$

Frage: Was ist ein guter Benchmark?

Definition 4.1:
(a) Der optimale Kolmogorov-Raum X_N^{Kol} lautet

$$X_N^{\text{Kol}} := \arg \inf_{\substack{X_N \subset X \\ \dim(X_N) = N}} \left\{ \sup_{\mu \in \Xi^{\text{train}}} \sigma_X(u(\mu), X_N) \right\}. \tag{4.2}$$

(b) Die Kolmogorov N-Breite (engl.: „N-width") lautet

$$\bar{\varepsilon}_N^{\text{Kol}} := \sup_{\mu \in \Xi^{\text{train}}} \sigma_X(u(\mu), X_N^{\text{Kol}}). \tag{4.3}$$

Bemerkung 4.2:
(a) Abhängig von der Wahl von Ξ^{train} zeigt $\overline{\varepsilon}_N^{\text{Kol}}$ den „besten Fehler“ an, wenn wir X_N frei wählen könnten. Dies ist also eine theoretische untere Schranke.
(b) Die Räume X_N^{Kol} sind nicht hierarchisch. Es ist ein nichtlineares (NP-hartes) kombinatorisches Problem. Wesentliche Frage:
- *Wie verhält sich $\overline{\varepsilon}_N^{\text{Kol}}$ für $N \to \infty$? Z.B. $\overline{\varepsilon}_N^{\text{Kol}} \lesssim N^{-s}$, $e^{-\alpha N}, \ldots$*
- *Ist ein numerisches Verfahren so gut, dass es das gleiche Abklingverhalten hat? „Asymptotisch optimal“: Z.B. wenn*

$$\overline{\varepsilon}_N^{\text{Kol}} \leq c_1\, e^{-\alpha N}, \text{ dann } \quad \varepsilon_N^{\text{Numerisch}} \leq c_2\, e^{-\alpha N}.$$

(c) Wenn die Abhängigkeit von $u(\mu)$ bzgl. μ „glatt“ ist, dann kann man $\overline{\varepsilon}_N^{\text{Kol}} \lesssim e^{-\alpha N}$ hoffen, dies begründet die Effizienz der RBM.
(d) Wegen des Supremums in (4.3) ist dies eine L_∞-Betrachtung (worst case).

Frage:

- Kann man a priori etwas über das Abklingverhalten von $\overline{\varepsilon}_N^{\text{Kol}}$ für $N \to \infty$ sagen?
- Kann man daraus „optimale“ Samples S_N gewinnen?

Beispiel 4.3:
Heizblock mit $P = 2$ wie im einführenden Beispiel. Für

$$X_N := \operatorname{span}\left\{u\left((\mu_{\min}, \mu_{\min})^T\right), u\left((\mu_{\max}, \mu_{\min})^T\right)\right\}$$

gilt $\sigma_X(u^{\mathcal{N}}(\mu), X_N) = 0\ \forall \mu \in \mathcal{D}$. (Übung)

Man hat hier also Exaktheit. Eine analoge Aussage gilt auch, wenn nur die rechte Seite von μ abhängt, *nicht* die Bilinearform ($a(u(\mu), v) = f(v; \mu)\ \forall v$). Dies ist aber die Ausnahme.

Satz 4.4 (Fink, Rheinboldt, 1983)**:**
Sei $\mu_0 \in \mathcal{U} \subset \mathcal{D} \subset \mathbb{R}$, $\mathcal{U}$ eine Umgebung und $u(\mu)$ sei analytisch in $\mathcal{U}$. Dann existiert ein $\sigma > 0$ und ein $c < \infty$ mit

$$\sigma_X(u(\mu), X_{k,\mu_0}) \leq c|\mu - \mu_0|^{k+1} \qquad \forall \mu \in \mathcal{B}(\mu_0, \delta).$$

Beweis. Mit Taylor gilt

$$\begin{aligned} u(\mu) &= \sum_{i=0}^{\infty} \frac{\partial^i}{\partial \mu^i} u(\mu_0) \frac{(\mu - \mu_0)^i}{i!} \\ &= \underbrace{\sum_{i=0}^{k} \frac{\partial^i}{\partial \mu^i} u(\mu_0) \frac{(\mu - \mu_0)^i}{i!}}_{=: v_k(\mu) \in X_{k,\mu_0}} + (\mu - \mu_0)^{k+1} \underbrace{\sum_{i=k+1}^{\infty} \frac{1}{i!} \frac{\partial^i}{\partial \mu^i} u(\mu_0)(\mu - \mu_0)^{i-k-1}}_{=: w_k(\mu)} \end{aligned}$$

Wähle $\delta < 1$ so klein, dass $\mathcal{B}(\mu_0, \delta) \subset \mathcal{U}$ und wegen der Analytizität gilt $\tilde{c} := \sup_{i \in \mathbb{N}} \|\frac{1}{i!}\frac{\partial^i}{\partial \mu^i} u(\mu_0)\|_X < \infty$. Dann gilt für $\mu \in \mathcal{B}(\mu_0, \delta)$:

$$\begin{aligned} \|w_k(\mu)\| &\leq \sum_{i=k+1}^{\infty} \underbrace{\left\| \frac{1}{i!}\frac{\partial^i}{\partial \mu^i} u(\mu_0) \right\|_X}_{\leq \tilde{c}} |\mu - \mu_0|^{i-k-1} \\ &\overset{\substack{geom. \\ Reihe}}{\leq} \tilde{c}\,\frac{1}{1-|\mu-\mu_0|} \leq \frac{\tilde{c}}{1-\delta} =: c. \end{aligned}$$

Da $v_k(\mu) \in X_{k,\mu_0}$ folgt

$$\sigma_X(u(\mu), X_{k,\mu_0}) \leq \|u(\mu) - v_k(\mu)\|_X = \|(\mu-\mu_0)^{k+1} w_k(\mu)\| \leq c|\mu-\mu_0|^{k+1}.$$

□

Also: analytische Parameter-Abhänigkeit ⇒ lokal polynomiale Konvergenz

Satz 4.5 (Maday, Patera, Turicini, 2002)**:**
Sei $\mathcal{D} = [\mu_{\min}, \mu_{\max}] \subset \mathbb{R}^+$, $0 < \mu_{\min} < 1$, $\mu_{\max} := \frac{1}{\mu_{\min}}$, $b(v, w; \mu) := b_0(u, v) + \mu b_1(u, v)$, $g(v; \mu) \equiv g(v)$, $X = Y$, b_0, b_1 *s.p.d. Weiter sei* $a := \ln \frac{\mu_{\max}}{\mu_{\min}} > \frac{1}{2e}$ *und setze* $N_0 := 1 + \lfloor 2ea + 1 \rfloor$. *Wähle* $\mu_{\min} = \mu_1 < \ldots < \mu_N = \mu_{\max}$, $N \geq N_0$, *logarithmisch äquidistant, d.h.*

$$\ln(\mu_{i+1}) - \ln(\mu_i) = \frac{\ln(\mu_{\max}) - \ln(\mu_{\min})}{N-1} \tag{4.4}$$

und $X_N := \operatorname{span}\{u(\mu_i) : i = 1, \ldots, N\}$. *Dann gilt*

$$\frac{\|u(\mu) - u_N(\mu)\|_X}{\|u(\mu)\|_X} \leq e^{-\frac{N-1}{N_0-1}} \qquad \forall \mu \in \mathcal{D}\ \forall N \geq N_0. \tag{4.5}$$

Beweis. (länglich), [MPT 2002], vgl. auch [PR, Proposition 3d]. □

Korollar 4.6:
Unter den Voraussetzungen von Satz 4.5 gilt

$$\sigma_X(u(\mu), X_N) \leq c \cdot e^{-\frac{N-1}{N_0-1}} \qquad \forall \mu \in \mathcal{D}\ \forall N \geq N_0.$$

Beweis. Es gilt $\|u(\mu)\|_X \leq \frac{1}{\alpha(\mu)} \|g(\mu)\|_{X'}$ und damit

$$\begin{aligned} \sigma_X(u(\mu), X_N) &\leq \frac{\|u(\mu) - u_N(\mu)\|_X}{\|u(\mu)\|_X} \cdot \|u(\mu)\|_X \\ &\leq \frac{\|u(\mu) - u_N(\mu)\|_X}{\|u(\mu)\|_X} \cdot \underbrace{\frac{1}{\alpha(\mu)} \|g(\mu)\|_{X'}}_{=:c} \end{aligned}$$

□

Bemerkung 4.7:

(a) In dieser speziellen Situation hat man also globale exponentielle Konvergenz!

(b) Die μ_i in (4.4) heißen „magic points“.

(c) Bislang kennt man solche Punkte im Wesentlichen in 1D ($P = 1$). Im Allgemeinen kennt man „gute“ μ_i nicht *a-priori, man bestimmt sie z.B. über Sampling-Strategien. Wir werden zwei sehen:*

- *Proper orthogonal Decomposition (PoD)*
- *Greedy*

5 Proper orthogonal Decomposition

- auch bekannt als Korhunen-Loève-Zerlegung (KL)
  oder Hauptkomponenten-Analyse (Principle Component Analysis - PCA)
- Verwendung seit ca. 1960er Jahre bei
  - turbulenten Strömungen
    (Analyse, „kohärente Strukturen“ - reduzierte Simulation)
  - Strömungs-Struktur-Interaktion
  - Struktur-Mechanik
- Einsatz ist verbreitet bei Zeitabhängigen Problemen und Optimierung
- Idee: Ersetzung von L_∞ (worst case) durch L_2 (average) - Skalarprodukt

Definition 5.1:
(a) Der PoD-Raum X_N^{PoD} ist definiert als

$$X_N^{\mathrm{PoD}} := \arg \inf_{\substack{X_N \subset X^{\mathrm{train}} \\ \dim(X_N)=N}} \frac{1}{n_{\mathrm{train}}} \sum_{\mu \in \Xi^{\mathrm{train}}} \sigma_X(u(\mu), X_N)^2, \tag{5.1}$$

mit $X^{\mathrm{train}} := \operatorname{span}\left\{u(\mu_i^{\mathrm{train}}) : 1 \le i \le n_{\mathrm{train}}\right\}$.
(b) Der PoD-Fehler lautet

$$\overline{\varepsilon}_N^{\mathrm{PoD}} := \left(\frac{1}{n_{\mathrm{train}}} \sum_{\mu \in \Xi^{\mathrm{train}}} \sigma_X(u(\mu), X_N)^2 \right)^{\frac{1}{2}} \tag{5.2}$$

Bemerkung 5.2:
(a) Es gilt $X_N \subset X^{\mathrm{train}}$ und die Räume sind hierarchisch.
(b) Man nennt ein RB-Raum vom Lagrange-Typ, falls $X_N = \operatorname{span}\left\{u^{\mathcal{N}}(\mu) : \mu \in S_N\right\}$. In (5.1) werden Linearkombinationen aus X^{train} gebildet, also sind X_N^{PoD} nicht vom Lagrange-Typ.

Nun zu einer expliziten Berechnungsmethode für X_N^{PoD}:
Sei $\underline{C}^{\mathrm{PoD}} := \left[C_{i,j}^{\mathrm{PoD}}\right]_{1 \le i,j \le n_{\mathrm{train}}}$ mit

$$C_{i,j}^{\mathrm{PoD}} := \frac{1}{n_{\mathrm{train}}} \left(u(\mu_i^{\mathrm{train}}), u(\mu_j^{\mathrm{train}})\right)_X \tag{5.3}$$

die Korelationsmatrix (Gram-Matrix).
Mit der Gram-Matrix $\mathbb{X} := \left[(\varphi_k^{\mathcal{N}}, \varphi_\ell^{\mathcal{N}})_X\right]_{1\le k,\ell\le\mathcal{N}}$ von $\Phi^{\mathcal{N}}$ in X und

$$u(\mu_i^{\text{train}}) = \sum_{k=1}^{\mathcal{N}} u_k(\mu_i^{\text{train}})\varphi_i^{\mathcal{N}}, \qquad \underline{u}(\mu_i^{\text{train}}) := (u_k(\mu_i^{\text{train}}))_{1\le k\le\mathcal{N}}$$

erhalten wir die Darstellung

$$\begin{aligned} C_{i,j}^{\text{PoD}} &= \frac{1}{n_{\text{train}}} \sum_{k=1}^{\mathcal{N}}\sum_{\ell=1}^{\mathcal{N}} u_k(\mu_i^{\text{train}})u_\ell(\mu_j^{\text{train}})(\varphi_k^{\mathcal{N}}, \varphi_\ell^{\mathcal{N}})_X \\ &= \frac{1}{n_{\text{train}}}\, \underline{u}(\mu_i^{\text{train}})^T\, \mathbb{X}\, \underline{u}(\mu_j^{\text{train}}). \end{aligned} \tag{5.4}$$

Offenbar ist nach Annahme (Wahl der Samples) $\underline{C}^{\text{PoD}}$ s.p.d. und wir suchen Eigenpaare $(\underline{\chi}_k^{\text{PoD}}, \lambda_k^{\text{PoD}}) \in \mathbb{R}^{n_{\text{train}}} \times \mathbb{R}^+,\ 1 \le k \le n_{\text{train}}$ mit

$$\underline{C}^{\text{PoD}}\underline{\chi}_k^{\text{PoD}} = \lambda_k^{\text{PoD}}\underline{\chi}_k^{\text{PoD}}, \qquad (\underline{\chi}_k^{\text{PoD}})^T\underline{\chi}_\ell^{\text{PoD}} = \delta_{k,\ell} \tag{5.5}$$

sowie (o.B.d.A.) $\lambda_1^{\text{PoD}} \ge \lambda_2^{\text{PoD}} \ge \ldots \ge \lambda_{n_{\text{train}}}^{\text{PoD}} > 0$. Die entsprechenden Funktionen in X^{train} seien dann

$$\chi_k^{\text{PoD}} := (\lambda_k^{\text{PoD}})^{-\frac{1}{2}} \sum_{m=1}^{n_{\text{train}}} \chi_{k,m}^{\text{PoD}}\, u(\mu_m^{\text{train}}), \qquad \underline{\chi}_{k,m}^{\text{PoD}} = (\chi_{k,m}^{\text{PoD}})_{1\le m\le n_{\text{train}}}.$$

Wegen

$$\begin{aligned} (\chi_k^{\text{PoD}}, \chi_\ell^{\text{PoD}})_X &= (\lambda_k^{\text{PoD}} \cdot \lambda_\ell^{\text{PoD}})^{-\frac{1}{2}} \sum_{m,n=1}^{n_{\text{train}}} \chi_{k,m}^{\text{PoD}}\chi_{\ell,n}^{\text{PoD}}(u(\mu_m^{\text{train}}), u(\mu_n^{\text{train}}))_X \\ &= n_{\text{train}}(\lambda_k^{\text{PoD}} \cdot \lambda_\ell^{\text{PoD}})^{-\frac{1}{2}}(\underline{\chi}_k^{\text{PoD}})^T \underbrace{\underline{C}^{\text{PoD}}\underline{\chi}_\ell^{\text{PoD}}}_{\lambda_\ell^{\text{PoD}}\cdot\underline{\chi}_\ell^{\text{PoD}}} \\ &= n_{\text{train}} \left(\frac{\lambda_\ell^{\text{PoD}}}{\lambda_k^{\text{PoD}}}\right)^{\frac{1}{2}} \underbrace{(\underline{\chi}_k^{\text{PoD}}, \underline{\chi}_\ell^{\text{PoD}})}_{=\delta_{k,\ell}} = \delta_{k,\ell}\cdot n_{\text{train}} \end{aligned} \tag{5.6}$$

ist das eine orthogonale Basis für X^{train}.

Lemma 5.3:
Für $1 \le N \le n_{\text{train}}$ gilt: $\overline{\varepsilon}_N^{\text{PoD}} = (\sum_{k=N+1}^{n_{\text{train}}} \lambda_k^{\text{PoD}})^{\frac{1}{2}}$

Bemerkung 5.4:
Wenn λ_k^{PoD} für $k \to n_{\text{train}}$ schnell abklingt wird $\overline{\varepsilon}_N^{\text{PoD}}$ bzgl. N schnell klein.

Beweis. Jedes $u(\mu) \in X^{\text{train}},\ \mu \in \Xi^{\text{train}}$, besitzt eine eindeutige Entwicklung in der Basis χ_k^{PoD}, $k = 1, \ldots, n_{\text{train}}$, als $u(\mu) = \sum_{m=1}^{n_{\text{train}}} u_m(\mu)\chi_m^{\text{PoD}}$, ebenso jedes $w_N \in X_N =$

span $\{\chi_1^{\mathrm{PoD}}, \dots, \chi_N^{\mathrm{PoD}}\}$ als $w_N = \sum_{m=1}^N w_{N,m}\chi_m^{\mathrm{PoD}}$. Wegen (5.6) gilt

$$\begin{aligned}\|u(\mu) - w_N\|_X^2 &\overset{(5.6)}{\underset{orth.}{=}} \left\|\sum_{m=1}^N (u_m(\mu) - w_{N,m})\chi_m^{\mathrm{PoD}}\right\|_X^2 + \left\|\sum_{m=N+1}^{n_{\mathrm{train}}} u_m(\mu)\chi_m^{\mathrm{PoD}}\right\|_X^2 \\ &= n_{\mathrm{train}}\left\{\sum_{m=1}^N (u_m(\mu) - w_{N,m})^2 + \sum_{m=N+1}^{n_{\mathrm{train}}} u_m(\mu)^2\right\},\end{aligned}$$

also $\sigma_X(u(\mu), X_N)^2 = n_{\mathrm{train}} \sum_{m=N+1}^{n_{\mathrm{train}}} u_m(\mu)^2$. Weiter gilt

$$\begin{aligned}\lambda_k^{\mathrm{PoD}}\chi_{k,m}^{\mathrm{PoD}} &= \lambda_k^{\mathrm{PoD}}(\underline{\chi}_k^{\mathrm{PoD}})_m = \sum_{\ell=1}^{n_{\mathrm{train}}} C_{m,\ell}^{\mathrm{PoD}}\chi_{k,\ell}^{\mathrm{PoD}} \\ &\overset{\text{Def. von } \underline{C}^{\mathrm{PoD}}}{=} \frac{1}{n_{\mathrm{train}}}\sum_{\ell=1}^{n_{\mathrm{train}}} (u(\mu_m^{\mathrm{train}}), u(\mu_\ell^{\mathrm{train}}))_X \cdot \chi_{k,\ell}^{\mathrm{PoD}} \\ &= \frac{1}{n_{\mathrm{train}}}(u(\mu_m^{\mathrm{train}}), \underbrace{\sum_{\ell=1}^{n_{\mathrm{train}}} \chi_{k,\ell}^{\mathrm{PoD}} u(\mu_\ell^{\mathrm{train}})}_{=(\lambda_k^{\mathrm{PoD}})^{\frac{1}{2}}\chi_k^{\mathrm{PoD}}})_X \\ &= \frac{(\lambda_k^{\mathrm{PoD}})^{\frac{1}{2}}}{n_{\mathrm{train}}}(u(\mu_m^{\mathrm{train}}), \chi_k^{\mathrm{PoD}})_X \\ &\overset{(5.6)}{=} (\lambda_k^{\mathrm{PoD}})^{\frac{1}{2}}\frac{(u(\mu_m^{\mathrm{train}}), \chi_k^{\mathrm{PoD}})_X}{\|\chi_k^{\mathrm{PoD}}\|_X^2} = (\lambda_k^{\mathrm{PoD}})^{\frac{1}{2}} u_k(\mu_m^{\mathrm{train}}),\end{aligned}$$

also

$$u_k(\mu_m^{\mathrm{train}}) = (\lambda_k^{\mathrm{PoD}})^{\frac{1}{2}}\chi_{k,m}^{\mathrm{PoD}}. \tag{5.7}$$

Schließlich gilt wegen $\|\underline{\chi}_k^{\mathrm{PoD}}\| = 1$ (ONB)

$$\begin{aligned}(\bar{\varepsilon}_N^{\mathrm{PoD}})^2 &= \frac{1}{n_{\mathrm{train}}}\sum_{\mu\in\Xi^{\mathrm{train}}} \sigma_X(u(\mu), X_N)^2 = \sum_{i=1}^{n_{\mathrm{train}}}\sum_{m=N+1}^{n_{\mathrm{train}}} (u_m(\mu_i^{\mathrm{train}}))^2 \\ &\overset{(5.7)}{=} \sum_{i=1}^{n_{\mathrm{train}}}\sum_{m=N+1}^{n_{\mathrm{train}}} \lambda_m^{\mathrm{PoD}}(\chi_{m,i}^{\mathrm{PoD}})^2 = \sum_{m=N+1}^{n_{\mathrm{train}}} \lambda_m^{\mathrm{PoD}} \underbrace{\sum_{i=1}^{n_{\mathrm{train}}} (\chi_{m,i}^{\mathrm{PoD}})^2}_{=\|\underline{\chi}_m^{\mathrm{PoD}}\|^2 = 1} \\ &\overset{ONB}{=} \sum_{m=N+1}^{n_{\mathrm{train}}} \lambda_m^{\mathrm{PoD}}.\end{aligned}$$

□

Will man also eine Zielgenauigkeit (im quadratischen Mittel) $\varepsilon_{\mathrm{tol,min}} > 0$ erreichen und dies bei minimalem N, also $N_{\max} := \min\{N \in \mathbb{N} : \bar{\varepsilon}_N^{\mathrm{PoD}} \le \varepsilon_{\mathrm{tol,min}}\}$, dann bestimmt man das minimale $N \in \mathbb{N}$ mit $\sum_{k=N+1}^{n_{\mathrm{train}}} \lambda_k^{\mathrm{PoD}} \le \varepsilon_{\mathrm{tol,min}}^2$.
Weiter gilt für die PoD-Räume $X_n^{\mathrm{PoD}} = \operatorname{span}\{\chi_k^{\mathrm{PoD}} : k = 1, \dots, N\}$ und diese Funktionen sind orthogonal.

Bemerkung 5.5:
(a) Der kombinatorische Aufwand verschwindet zu Lasten der schwächeren L_2-Norm.
(b) Der offline-Aufwand ist:
- $\mathcal{O}(n_{\text{train}} \cdot \mathcal{N})$ *für die snapshots*
- $\mathcal{O}(n_{\text{train}}^3)$ *für das Eigenwert-Problem*

$\Rightarrow$ n_{train} *darf nicht zu groß sein!*

(c) Vorteil: a priori Fehlerinformation

6 A posteriori-Fehleranalysis

Die Alternative zu PoD - Greedy - beruht auf einem Greedy-Optimierungsverfahren bezüglich Ξ^{train}, mehr dazu später. Um diesen Greedy-Schritt ausführen zu können, muss der Fehler

$$e_N(\mu) := u^{\mathcal{N}}(\mu) - u_N(\mu) \quad \text{bzw.} \quad \|e_N(\mu)\|_X$$

berechnet werden. Dies würde für jedes $\mu \in \Xi^{\text{train}}$ bedeuten, dass wir die truth $u^{\mathcal{N}}(\mu)$ berechnen müssen. $\lightning$

Idee: Ersetze den Fehler durch Fehlerschätzer $\Delta_N(\mu)$.

Definition 6.1:
Der Ausdruck

$$r_N(v;\mu) := g(v;\mu) - b(u_N(\mu), v;\mu), \ v \in Y, \qquad \mu \in \mathcal{D} \tag{6.1}$$

heißt Residuum von (3.1).

Bemerkung 6.2:
(a) Offenbar gilt $r_N(\mu) := r_N(\cdot\,;\mu) \in Y'$, $r_N(v_N;\mu) = 0$ für alle $v_N \in Y_N$.
(b) Es gilt

$$r_N(v;\mu) = b(e_N(\mu), v;\mu) \qquad \forall v \in Y \tag{6.2}$$

denn

$$\begin{aligned} r_N(v;\mu) &= g(v;\mu) - b(u_N(\mu), v;\mu) = b(u(\mu), v;\mu) - b(u_N(\mu), v;\mu) \\ &= b(u(\mu) - u_N(\mu), v;\mu) = b(e_N(\mu), v;\mu) \end{aligned}$$

Beachte: Das Residuum ist berechenbar ohne $u^{\mathcal{N}}(\mu)$ *zu berechnen!*

Lemma 6.3:
Es gilt

$$\|e_N(\mu)\|_X \le \Delta_N(\mu) := \frac{\|r_N(\mu)\|_{Y'}}{\beta_{LB}(\mu)} \tag{6.3}$$

mit $0 < \beta_{LB}(\mu) \le \beta^{\mathcal{N}}(\mu)$.

Beweis. Es gilt per Definition $\beta^{\mathcal{N}}(\mu)\,\|w\|_X \le \sup_{v\in Y^{\mathcal{N}}} \frac{b(w,v;\mu)}{\|v\|_Y}$ für alle $w \in X^{\mathcal{N}}$, also für $w = e_N(\mu) = u^{\mathcal{N}}(\mu) - u_N(\mu) \in X^{\mathcal{N}}$:

$$\begin{aligned}\beta^{\mathcal{N}}(\mu)\|e_N(\mu)\|_X &\le \sup_{v\in Y^{\mathcal{N}}} \frac{b(e_N(\mu),v;\mu)}{\|v\|_Y} \\ &\overset{(6.2)}{=} \sup_{v\in Y^{\mathcal{N}}} \frac{r_N(v;\mu)}{\|v\|_Y} \le \|r_N(\mu)\|_{Y'}.\end{aligned}$$

□

Bemerkung 6.4:

(a) *Analog kann man andere Normen betrachten, z.B. die* Energienorm *für ein $\overline{\mu} \in \mathcal{D}$ (im Fall $X = Y$)*

$$\|w\|_{\overline{\mu}} := \sqrt{(w\,,w)_{\overline{\mu}}}, \qquad (w\,,v)_{\overline{\mu}} := b_S(w,v,\overline{\mu}).$$

Es gilt

$$\sqrt{\alpha(\mu)}\,\|w\|_X \le \|w\|_{\overline{\mu}} \le \sqrt{\gamma(\mu)}\,\|w\|_X,\ w \in X.$$

Ebenso kann man die relativen Fehler

$$E_N^{\overline{\mu}} := \frac{\|e_N(\mu)\|_{\overline{\mu}}}{\|u^{\mathcal{N}}(\mu)\|_{\overline{\mu}}}, \qquad E_N^X := \frac{\|e_N(\mu)\|_X}{\|u^{\mathcal{N}}(\mu)\|_X}$$

abschätzen (Übung).

(b) *Die Abschätzung im Beweis bleibt gültig, wenn wir $\beta^{\mathcal{N}}(\mu)$ durch eine* berechenbare untere Schranke *$\beta_{LB}(\mu)$ ersetzen.*

(c) *Damit $\Delta_N(\mu)$ berechenbar ist (unabhängig von $\mathcal{N}$) brauchen wir effiziente Methoden für*

- *$\beta_{LB}(\mu)$*
- *$\|r_N(\mu)\|_{Y'}$ (Dualnorm des Residuums)*

Zunächst zur Frage: Wie gut ist $\Delta_N(\mu)$?

Definition 6.5:
Die Effektivität von $\Delta_N(\mu)$ ist definiert als

$$\eta_N(\mu) := \frac{\Delta_N(\mu)}{\|e_N(\mu)\|_X}, \tag{6.4}$$

und analog für andere Normen.

Satz 6.6:
Sei $\infty > \gamma_{UB}(\mu) \ge \gamma^{\mathcal{N}}(\mu)$ eine obere Schranke für $\gamma^{\mathcal{N}}(\mu)$. Dann gilt

$$\eta_N(\mu) \le \frac{\gamma_{UB}(\mu)}{\beta_{LB}(\mu)}. \tag{6.5}$$

Beweis. Es gilt

$$\begin{aligned}\|r_N(v;\mu)\|_Y &= |b(e_N(\mu),v;\mu)| \le \gamma^{\mathcal{N}}(\mu)\,\|e_N(\mu)\|_X\,\|v\|_Y \\ &\le \gamma_{UB}(\mu)\,\|e_N(\mu)\|_X\,\|v\|_Y,\end{aligned}$$

also $\|r_N(\mu)\|_{Y'} \le \gamma_{UB}(\mu)\,\|e_N(\mu)\|_X$ und damit

$$\eta_N(\mu) = \frac{\Delta_N(\mu)}{\|e_N(\mu)\|_X} = \frac{1}{\beta_{LB}(\mu)}\,\frac{\|r_N(\mu)\|_{Y'}}{\|e_N(\mu)\|_X} \le \frac{\gamma_{UB}(\mu)}{\beta_{LB}(\mu)}.$$

□

Ebenso kann man Effektivitäten für andere Normen abschätzen. Nun zur Berechnung der Schranken für $\alpha(\mu), \beta(\mu), \gamma(\mu)$.

Wir beginnen mit der Berechnung der Schranken und beschränken uns auf den koerziven Fall, also $\alpha_{UB}(\mu)$ und $\gamma_{UB}(\mu)$.

Lemma 6.7 (min-θ für α)**:**
Sei b affin im Parameter und koerziv mit

$$b^q(v,v) \ge 0 \quad \forall v \in X, \qquad \theta_b^q(\mu) > 0 \quad \forall \mu \in \mathcal{D}. \tag{6.6}$$

Sei $\overline{\mu} \in \mathcal{D}$ fest mit bekanntem $\alpha(\overline{\mu})$ und setze

$$\alpha_{LB}(\mu) := \alpha(\overline{\mu}) \min_{q=1,\dots,Q_b} \frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}. \tag{6.7}$$

Dann gilt: $0 < \alpha_{LB}(\mu) \le \alpha(\mu) \quad \forall \mu \in \mathcal{D}$.

Beweis. Für $v \in X$ gilt:

$$\begin{aligned}b(v,v;\mu) &= \sum_{q=1}^{Q_b} \theta_b^q(\mu) b^q(v,v) = \sum_{q=1}^{Q_b} \frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})} \theta_b^q(\overline{\mu}) b^q(v,v) \\ &\ge \left(\min_{q=1,\dots,Q_b} \frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}\right) \underbrace{\sum_{q=1}^{Q_b} \theta_b^q(\overline{\mu}) b^q(v,v)}_{= b(v,v,\overline{\mu}) \ge \alpha(\overline{\mu})\|v\|_X^2} \\ &\ge \alpha_{LB}(\mu)\|v\|_X^2\end{aligned}$$

□

Korollar 6.8 (min-θ für β)**:**
Sei b affin im Parameter mit $\theta_b^q(\mu) > 0\ \forall \mu \in \mathcal{D}$; $\overline{\mu} \in \mathcal{D}$ sei fest mit bekanntem $\beta(\overline{\mu}) > 0$. Mit

$$\beta_{LB}(\mu) := \beta(\overline{\mu}) \min_{q=1,\dots,Q_b} \frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}$$

gilt $0 < \beta_{LB}(\mu) \le \beta(\mu)\ \forall \mu \in \mathcal{D}$.

Beweis. Für $w \in X$ gilt wie oben

$$\begin{aligned}\sup_{v\in Y}\frac{b(w,v;\mu)}{\|v\|_Y} &\geq \left(\min_{q=1,\dots,Q_b}\frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}\right)\sup_{v\in Y}\frac{\sum_{q=1}^{Q_b}\theta_b^q(\overline{\mu})b^q(w,v)}{\|v\|_Y}\\ &= \left(\min_{q=1,\dots,Q_b}\frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}\right)\sup_{v\in Y}\frac{b(w,v,\overline{\mu})}{\|v\|_Y}\\ &\geq \left(\min_{q=1,\dots,Q_b}\frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}\right)\beta(\overline{\mu})\|w\|_X = \beta_{LB}(\mu)\|w\|_X.\end{aligned}$$

□

Lemma 6.9 (min-θ für γ)**:**
Sei b affin im Parameter mit $\theta_b^q(\mu) > 0\ \forall\mu\in\mathcal{D}$. Sei $\overline{\mu}\in\mathcal{D}$ fest und $\gamma(\overline{\mu})$ bekannt. Mit

$$\gamma_{UB}(\mu) := \gamma(\overline{\mu})\max_{q=1,\dots,Q_b}\frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}$$

gilt $\gamma(\mu)\leq\gamma_{UB}(\mu)\ \forall\mu\in\mathcal{D}$.

Beweis. Seien $w\in X$, $v\in Y$ beliebig. Dann gilt

$$\begin{aligned}b(w,v;\mu) &= \sum_{q=1}^{Q_b}\frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}\theta_b^q(\overline{\mu})b(w,v;\mu)\\ &\leq \left(\max_{q=1,\dots,Q_b}\frac{\theta_b^q(\mu)}{\theta_b^q(\overline{\mu})}\right)b(w,v,\overline{\mu}) \leq \gamma_{UB}(\mu)\,\|w\|_X\,\|v\|_Y.\end{aligned}$$

□

Wir werden später noch andere Methoden zur Berechnung der Schranken kennenlernen (SCM).

Nun zur Dualnorm des Residuums. Aufgrund des Riesz'schen Darstellungssatzes existiert zu $r_N(\mu)\in Y'$ genau ein $\hat{r}_N(\mu)\in Y$ (der Riesz-Repräsentant) mit

$$(\hat{r}_N(\mu),v)_Y = r_N(v;\mu) \qquad \forall v\in Y \qquad \text{und} \tag{6.8}$$

$$\|\hat{r}_N(\mu)\|_Y = \|r_N(\mu)\|_{Y'} \qquad \forall\mu\in\mathcal{D}. \tag{6.9}$$

Wegen (6.2) gilt

$$(\hat{r}_N(\mu),v)_Y = b(e_N(\mu),v;\mu) \qquad \forall\mu\in\mathcal{D}\ \forall v\in Y. \tag{6.10}$$

Wir nehmen an, dass b und g affin im Parameter sind und bezeichnen für jedes $1\leq q\leq Q_g$ mit $v_g^q\in Y$ die Lösung von

$$(v_g^q,v)_Y = g^q(v), \qquad v\in Y, \tag{6.11}$$

sowie für $1 \leq k \leq N$ mit $v_{b,k}^q \in Y$ die Lösung von

$$(v_{b,k}^q, v)_Y = b^q(\xi_k, v), \qquad v \in Y. \tag{6.12}$$

Dies sind offenbar $Q_g + N\,Q_b =: Q_r$ Gleichungssysteme der Dimension $\mathcal{N}$, die offline gelöst werden können (unabh. von μ). Dann gilt

$$\begin{aligned} r_N(v;\mu) &= g(v;\mu) - b(u_N(\mu), v;\mu) = \sum_{q=1}^{Q_g} \theta_g^q(\mu) g^q(v) - \sum_{q=1}^{Q_b} \theta_b^q(\mu) b^q(u_N(\mu), v) \\ &= \sum_{q=1}^{Q_g} \theta_g^q(\mu) \underbrace{g^q(v)}_{\overset{(6.11)}{=} (v_g^q, v)_Y} - \sum_{q=1}^{Q_b} \sum_{k=1}^{N} \theta_b^q(\mu) u_{N,k}(\mu) \underbrace{b^q(\xi_k, v)}_{\overset{(6.12)}{=} (v_{b,k}^q, v)_Y} \\ &=: \sum_{q=1}^{Q_r} \theta_r^q(\mu) (v_r^q, v)_Y, \end{aligned}$$

also eine affine Zerlegung des Residuums. Also:

$$\hat{r}_N(\mu) = \sum_{q=1}^{Q_r} \theta_r^q(\mu) v_r^q \in Y.$$

Mit (6.9) gilt also:

$$\begin{aligned} \|r_N(\mu)\|_{Y'}^2 &= \|\hat{r}_N(\mu)\|_Y^2 = (\hat{r}_N(\mu), \hat{r}_N(\mu))_Y \\ &= \sum_{q=1}^{Q_r} \sum_{q'=1}^{Q_r} \theta_r^q(\mu) \theta_r^{q'}(\mu) \underbrace{(v_r^q, v_r^{q'})_Y}_{=: (\underline{\underline{G}})_{q,q'}} =: \underline{\theta}_r(\mu)^T \, \underline{\underline{G}} \, \underline{\theta}_r(\mu), \end{aligned}$$

mit $\underline{\underline{G}} \in \mathbb{R}^{Q_r \times Q_r}$ offline bekannt und unabhängig von $\mathcal{N}$. Wir erhalten also

$$\|r_N(\mu)\|_{Y'} = \sqrt{\underline{\theta}_r(\mu)^T \, \underline{\underline{G}} \, \underline{\theta}_r(\mu)}. \tag{6.13}$$

Bemerkung 6.10:
Durch das Wurzelzeichen in (6.13) verliert man i.d.R. die Hälfte der Stellen bzgl. der Genauigkeit. Dieser „square-roof-effect“ ist für die Genauigkeit des Fehlerschätzers zu beachten.

7 Greedy-Algorithmus

Wir kennen nun - im Prinzip - einen effizienten Fehlerschätzer $\Delta_N(\mu)$. Diesen werden wir nun benutzen, um einen „optimalen“ Lagrange-RB-Raum X_N^* bzgl. hierarchischer Samples

$$S_N^* := \{\mu_1^*, \dots, \mu_N^*\}, \qquad \textit{also} \tag{7.1}$$

$$X_N^* := \operatorname{span}\{u(\mu_i^*) : 1 \le i \le N\}. \tag{7.2}$$

Wie bei PoD wählen wir $\varepsilon_{\text{tol,min}}$ und $\Xi^{\text{train}} \subset \mathcal{D}$.

Algorithmus 7.1 (GREEDY ($\Xi^{\text{train}}, \varepsilon_{\text{tol,min}}$))**:**

for $N = 1 : \overline{N}_{\max}$ **do**
  $\mu_N^* := \arg\max_{\mu \in \Xi^{\text{train}}} \Delta_{N-1}(\mu)$
  $\varepsilon_N^* := \Delta_{N-1}(\mu_N^*)$
  if $\varepsilon_N^* \le \varepsilon_{\text{tol,min}}$ **then**
    $N_{\max} := N - 1$
    exit
  end
  $S_N^* := S_{N-1}^* \cup \{\mu_N^*\}$
  $X_N^* := X_{N-1}^* \cup \operatorname{span}\{u(\mu_N^*)\}$
end

Bemerkung 7.2:

a) Greedy ermöglicht hierarchische Räume X_N als auch eine L_∞-Approximation.

b) Falls $\overline{\varepsilon}_N^{\text{Kol}} := \mathcal{O}(e^{-\alpha N})$, dann gilt für den maximalen „wahren“ Fehler

$$\overline{\varepsilon}_N^* := \max_{\mu \in \Xi^{\text{train}}} \|u^{\mathcal{N}}(\mu) - u_N(\mu)\|_X = \mathcal{O}(e^{-\beta N})$$

(Buffa, Maday, Patera, Prud'homme, Turicini, 2007).

c) Im Gegensatz zu PoD werden hier keine *Linearkombinationen gebildet, man erhält eine Lagrange-RB-Basis.*

d) Greedy ist offline deutlich effizienter als PoD, also kann man n_{train} größer wählen.

e) Es werden nur die snapshots für die ausgewählten μ_i^* mit dem Truth-Löser berechnet, bei PoD müssen alle *berechnet werden.*

f) PoD ist in der Regel für Zeit-artige Probleme und solche mit L_2-Struktur bzgl. des Parameters besser. PoD erlaubt a-priori-Abschätzung bzgl. N. Vergleich zwischen PoD und Greedy für Optimal-Steuerungs-Probleme siehe (Tonn, Ku, Volkmann, 2011).

Satz 7.3 (Binev, Cohen, Dahmen, DeVore, 2011)**:**
Seien $\mathcal{M} := \{u(\mu) : \mu \in \mathcal{D}\}$, $\mathcal{D}$ *kompakt und*

$$d_n(\mathcal{M}) := \inf_{\substack{X_n \subset X \\ dim(X_n)=n}} \sup_{\mu \in \mathcal{D}} \sigma_X(u(\mu), X_n),$$

(Kolmogorov n-Breite mit $\mathcal{D}$ *anstelle von* Ξ^{train}*) sowie* $\sigma_N := \max_{\mu \in \mathcal{D}} \Delta_N(\mu)$*. Dann gilt:*

(a) Falls $d_n(\mathcal{M}) \leq M \cdot n^{-\alpha}, \quad d_0(\mathcal{M}) \leq M, \quad \alpha, M > 0, \ n \in \mathbb{N}$*, dann gilt:*

$$\sigma_n = \mathcal{O}(M\, n^{-\alpha}), \ n > 0 \qquad \textit{(algebraische Konvergenz).}$$

(b) Falls $d_n(\mathcal{M}) \leq M \cdot e^{-a\, n^{\alpha}}, \quad a, M, \alpha > 0, \ n \in \mathbb{N}$*, dann gilt mit* $\beta := \frac{\alpha}{\alpha+1}$*:*

$$\sigma_n = \mathcal{O}(M \cdot e^{-n^{\beta}}) \qquad \textit{(exponentielle Konvergenz).}$$

Bemerkung 7.4:
Numerisch ist $\Xi^{\text{train}} = \mathcal{D}$ *oft unmöglich. In dem zitierten Paper wird die Optimalität aber auch für realisierbare Varianten gezeigt.*

8 Singulärwert-Zerlegung

Dies ist eine PoD-Variante insbesondere bekannt aus der Analyse von Strömungsdaten („coherent structures“). Sei $\Phi^{\mathcal{N}} := \left\{\varphi_k^{\mathcal{N}} : 1 \le k \le \mathcal{N}\right\}$ eine Basis von $X^{\mathcal{N}}$ (truth-Raum) und

$$u(\mu_i) = \sum_{i=1}^{\mathcal{N}} u_k(\mu_i)\varphi_k^{\mathcal{N}}, \qquad 1 \le i \le n_{\text{train}}$$

die Darstellung der snapshots bzgl. $\Phi^{\mathcal{N}}$. Man nennt

$$S := (u_k(\mu_i))_{1\le k\le \mathcal{N}, 1\le i\le n_{\text{train}}} \in \mathbb{R}^{\mathcal{N}\times n_{\text{train}}} \tag{8.1}$$

auch *snapshot-Matrix*.

Lemma 8.1:
Sei $\Phi^{\mathcal{N}}$ eine Orthonormalbasis und sei $\operatorname{rank}(U) = n' \le n_{\text{train}}$*, wobei $S = U\Sigma V^T$ die SVD von S mit $\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_{n'}, 0, \dots, 0)^T$ und $\sigma_1 > \sigma_2 > \dots > \sigma_{n'} > 0$ (echt fallend und strikt positiv). Dann gilt $(U^1, \dots, U^{n'}) = (\underline{\chi}_1^{\text{PoD}}, \dots, \underline{\chi}_{n'}^{\text{PoD}})$.*

Beweis. Für $\underline{u} \in \mathbb{R}^{n_{\text{train}}}$ gilt für $1 \le k \le n_{\text{train}}$

$$\begin{aligned}
(\underline{C}^{\text{PoD}}\underline{u})_k &= \sum_{\ell=1}^{n_{\text{train}}} \frac{1}{n_{\text{train}}} (u(\mu_k^{\text{train}}), u(\mu_\ell^{\text{train}}))_X \, u_\ell \\
&= \frac{1}{n_{\text{train}}} \sum_{\ell=1}^{n_{\text{train}}} \sum_{i,j=1}^{\mathcal{N}} u_i(\mu_k^{\text{train}}) u_j(\mu_\ell^{\text{train}}) \underbrace{(\varphi_i^{\mathcal{N}}, \varphi_j^{\mathcal{N}})_X}_{=\delta_{i,j}} \, u_\ell \\
&= \frac{1}{n_{\text{train}}} (SS^T\underline{u})_k.
\end{aligned}$$

Damit gilt:

$$\begin{aligned}
\underline{C}^{\text{PoD}} U^i &= \frac{1}{n_{\text{train}}} SS^T U^i = \frac{1}{n_{\text{train}}} U\Sigma \underbrace{V^T V}_{=Id} \Sigma \underbrace{U^T U^i}_{=e^i} \\
&= \frac{1}{n_{\text{train}}} \sigma_i^2 \, U^i,
\end{aligned}$$

also ist U^i ein Eigenvektor von $\underline{C}^{\text{PoD}}$. Die Eigenwerte $\frac{1}{n_{\text{train}}}\sigma_i^2$ sind monoton fallend, also genau so sortiert wie die Eigenwerte λ_i^{PoD} von $\underline{C}^{\text{PoD}}$ und damit $U^i = \underline{\chi}_i^{\text{PoD}}$ oder $U^i = -\underline{\chi}_i^{\text{PoD}}$. □

9 Output-Funktionale

Oft ist man nicht, oder zumindest nicht nur, an einer Approximation des *Zustands* $u(\mu)$ interessiert.

Definition 9.1:
(a) Sei $\ell : X \times \mathcal{D} \to \mathbb{R},\ \ell(\mu) := \ell(\cdot\,;\mu) \in X'$ *(oder* $\ell(\cdot\,;\mu) : X \to \mathbb{R}$*). Dann heißt*

$$s(\mu) := \ell(u(\mu);\mu) \tag{9.1}$$

Output-Funktional bzgl. (1.1). Analog bezeichnen wir

$$s^{\mathcal{N}}(\mu) := \ell(u^{\mathcal{N}};\mu), \qquad s_N(\mu) := \ell(u_N(\mu);\mu).$$

(b) Wir nennen das Problem

$$\text{Suche } u(\mu) \in X \text{ mit } \quad b(u(\mu), v;\mu) = g(v;\mu) \qquad \forall v \in Y. \tag{9.2}$$

$$\text{Bestimme } \quad s(\mu) := \ell(u(\mu);\mu) \tag{9.3}$$

*verträglich („*compliant*"), falls*
(i) $X = Y$ *und* $b(\cdot,\cdot;\mu)$ *symmetrisch* $\forall \mu \in \mathcal{D}$,
(ii) $\ell(\mu) = g(\cdot;\mu) \quad \forall \mu \in \mathcal{D}$.

Spezialfall: $\ell(u(\mu);\mu) \equiv \ell(u(\mu))$ (keine explizite Parameter-Abhängigkeit)

Wir nehmen wie für b und g an, dass ℓ im Parameter μ affin ist, also

$$\ell(w;\mu) = \sum_{q=1}^{Q_l} \theta_\ell^q(\mu)\, \ell^q(w), \qquad w \in X,\ \mu \in \mathcal{D} \tag{9.4}$$

Dies führt wieder zu einer effizienten offline-online-Berechnungsmethode unabhängig von $\mathcal{N}$ (online).

Beispiel (Fortsetzung)**:**
Im einführenden Beispiel des Komposit-Blocks sei

$$s(\mu) := \int_{\Gamma_1} u(x;\mu)\, dx, \tag{9.5}$$

also die mittlere Temperatur über Γ_1*. Wegen* $\ell(v;\mu) \equiv \ell(v)$ *ist dieser Output affin im Parameter und wegen* $\ell(v) = g(v)$ *ist das Problem verträglich.*

Lemma 9.2:
Für alle $\mu \in \mathcal{D}$ gilt im verträglichen s.p.d. Fall:

$$s^{\mathcal{N}}(\mu) - s_N(\mu) = \|e_N(\mu)\|_\mu^2, \tag{9.6}$$

$$\alpha(\mu) \cdot \|e_N(\mu)\|_X^2 \leq s^{\mathcal{N}}(\mu) - s_N(\mu) \leq \gamma(\mu) \cdot \|e_N(\mu)\|_X^2, \tag{9.7}$$

$$s^{\mathcal{N}}(\mu) - s_N(\mu) \leq \gamma(\mu)\sigma_X(u^{\mathcal{N}}(\mu), X_N). \tag{9.8}$$

Beweis. Es gilt

$$\begin{aligned} s^{\mathcal{N}}(\mu) - s_N(\mu) &= g(u^{\mathcal{N}}(\mu) - u_N(\mu); \mu) = b(u^{\mathcal{N}}(\mu), \underbrace{u^{\mathcal{N}}(\mu) - u_N(\mu)}_{= e_N(\mu)}; \mu) \\ &\overset{\substack{Gal.- \\ Orth.}}{=} b(e_N(\mu), e_N(\mu); \mu) = \|e_N(\mu)\|_\mu^2, \end{aligned}$$

also (9.6). Weiter gilt

$$\begin{aligned} \alpha(\mu) \cdot \|e_N(\mu)\|_X^2 &\leq b(e_N(\mu), e_N(\mu); \mu) \overset{s.o.}{=} g(e_N(\mu); \mu) \overset{s.o.}{=} s^{\mathcal{N}}(\mu) - s_N(\mu) \\ &= b(e_N(\mu), e_N(\mu); \mu) \leq \gamma(\mu) \cdot \|e_N(\mu)\|_X^2 \end{aligned}$$

und wegen

$$\|u^{\mathcal{N}}(\mu) - u_N(\mu)\|_\mu = \inf_{v_N \in X_N} \|u^{\mathcal{N}}(\mu) - v_N\|_\mu = \sigma_\mu(u^{\mathcal{N}}(\mu), X_N) \text{ (Galerkin-Projektion)}$$

sowie $\|w\|_\mu^2 = b(w, w; \mu) \leq \gamma(\mu) \cdot \|w\|_X^2$, also $\|w\|_\mu \leq \sqrt{\gamma(\mu)} \cdot \|w\|_X$ gilt zuletzt

$$\begin{aligned} s^{\mathcal{N}}(\mu) - s_N(\mu) &= \|u^{\mathcal{N}}(\mu) - u_N(\mu)\|_\mu^2 = (\inf_{v_N \in X_N} \|u^{\mathcal{N}} - v_N\|_\mu)^2 \\ &\leq \gamma(\mu) \cdot (\inf_{v_N \in X_N} \|u^{\mathcal{N}} - v_N\|_X)^2 = \gamma(\mu) \cdot \sigma_X(u^{\mathcal{N}}(\mu), X_N)^2 \end{aligned}$$

□

Bemerkung 9.3:
Es gilt $0 \leq s_N(\mu) < s^{\mathcal{N}}(\mu)$ im compliant-Fall.

Bemerkung 9.4:
Nach (9.7) ist der Output-Fehler äquivalent zum Quadrat *des Zustands-Fehlers. Es genügen also etwa die* Hälfte *der Genauigkeits-Stellen!*

Nun zur Fehlerschätzung. Ganz einfach wäre

$$s^{\mathcal{N}}(\mu) - s_N(\mu) = \ell(e_N(\mu); \mu) \leq \|\ell\|_{X'} \cdot \|e_N(\mu)\|_X \leq \Delta_N(\mu) \cdot \|\ell\|_{X'}$$

mit (6.3). Wegen Bemerkung 9.4 ist diese Abschätzung aber *nicht* gut:

- quadratischer Effekt geht verloren
- Berechnung von $\|\ell\|_{X'}$ pessimistisch und aufwändig

Lemma 9.5:
Es gilt für $X = Y$ und im compliant-s.p.d.-Fall

$$
\begin{aligned}
&(a) \quad \|e_N(\mu)\|_\mu \le \Delta_N^{\mathrm{En}}(\mu) := \frac{\|r_N(\mu)\|_{X'}}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}},\\
&(b) \quad s^{\mathcal{N}}(\mu) - s_N(\mu) \le \Delta_N^{s}(\mu) := \frac{\|r_N(\mu)\|_{X'}^2}{\alpha_{\mathrm{LB}}(\mu)} = (\Delta_N^{\mathrm{En}}(\mu))^2,\\
&(c) \quad \frac{s^{\mathcal{N}}(\mu) - s_N(\mu)}{s_N(\mu)} \le \Delta_N^{s,\mathrm{rel}}(\mu) := \frac{\|r_N(\mu)\|_{X'}^2}{\alpha_{\mathrm{LB}}(\mu)\cdot s_N(\mu)}.
\end{aligned}
$$

Beweis. (a) Es gilt

$$
\begin{aligned}
\|e_N(\mu)\|_\mu^2 &= b(e_N(\mu), e_N(\mu);\mu) = r_N(e_N(\mu);\mu) \le \|r_N(\mu)\|_{X'} \cdot \|e_N(\mu)\|_X\\
&\overset{(6.3)}{\le} \frac{\|r_N(\mu)\|_{X'}^2}{\alpha_{\mathrm{LB}}(\mu)}.
\end{aligned}
$$

(b) $s^{\mathcal{N}} - s_N(\mu) = b(e_N(\mu), e_N(\mu);\mu) = \|e_N(\mu)\|_\mu^2$, Rest mit (a).
(c) Es gilt $s^{\mathcal{N}}(\mu) = g(u^{\mathcal{N}}(\mu);\mu)$ sowie

$$
\begin{aligned}
s_N(\mu) &= g(u_N(\mu);\mu) = b(u_N(\mu), u_N(\mu);\mu)\\
&\ge \alpha(\mu)\cdot\|u_N(\mu)\|_X^2 > 0
\end{aligned}
$$

und damit folgt (c) aus (b).

□

- Nun kann man $\Delta_N^s(\mu)$ im Greedy-Algorithmus verwenden, um einen speziellen RB-Raum für einen RB-Approximation des Output-Funktionals im compliant-Fall zu bestimmen. Man kann auch den „Zustands-RB-Raum“ verwenden.
- Der compliant-Fall macht nur für $X = Y$ Sinn, da sonst $\ell(u(\mu))$ nicht definiert sein muss.
- Außerdem brauchten wir die s.p.d.-Annahme.

Nun zum allgemeinen Fall.

Definition 9.6:
Das Problem

$$
\begin{aligned}
&(9.9) \qquad \text{Suche } u(\mu)\in X: \quad b(u(\mu), v;\mu) = g(v;\mu) \qquad \forall v\in Y\\
&(9.10) \qquad \text{Bestimme } z(\mu)\in Y: \quad b(w, z(\mu);\mu) = -\ell(w;\mu) \qquad \forall w\in X
\end{aligned}
$$

mit der Ausgabe

$$s(\mu) := \ell(u(\mu);\mu)$$

heißt primal-duales Problem.

Bemerkung 9.7:
Zur Berechnung von $s(\mu)$ braucht man das duale Problem (9.10) offenbar nicht, wir werden es aber zur Fehlerabschätzung benutzen. Falls b symmetrisch und $\ell = q$ (compliant), dann gilt $u(\mu) = -z(\mu)$, (9.10) ist also unnötig.

Bemerkung 9.8 (Zusammenhang mit der Optimierung)**:**
Betrachte das restringierte Optimierungs-Problem

$$u \in X: \ \ell(u) \to \min, \quad NB: \ b(u,v) = g(v) \qquad \forall v \in Y. \tag{9.11}$$

Die Lagrange-Funktion $\mathcal{L} : X \times Y \to \mathbb{R}$ lautet

$$\mathcal{L}(u,z) := \ell(u) + [b(u,z) - g(z)], \qquad u \in X, \ z \in Y.$$

Also ist $z \in Y$ Lagrange-Parameter.
$\rightsquigarrow$ KKT-Bedingungen:

$$\begin{aligned} 0 &\overset{!}{=} \langle \mathcal{L}_u(u,z), v\rangle_{X' \times X} = \ell(v) + b(v,z), \qquad v \in X, \\ 0 &\overset{!}{=} \langle \mathcal{L}_z(u,z), w\rangle_{Y' \times Y} = b(u,w) - g(w), \qquad w \in Y, \end{aligned}$$

also ist (u^*, z^*) kritischer Punkt genau dann wenn

$$\begin{aligned} &\text{(i)} \ b(u^*, w) = g(w) \qquad \forall w \in Y, \\ &\text{(ii)} \ b(v, z^*) = -\ell(v) \qquad \forall v \in X, \end{aligned}$$

also ist der optimale Lagrange-Parameter identisch mit der Lösung der dualen Lösung!

Definition 9.9:
Sei $\mu \in \mathcal{D}$, $X_N \subset X$, $\widetilde{Y}_{\widetilde{N}} \subset Y$ zwei RB-Räume mit $\dim(X_N) = N$, $\dim(\widetilde{Y}_{\widetilde{N}}) = \widetilde{N}$, $Y_N \subset Y$, $\widetilde{X}_{\widetilde{N}} \subset X$ „geeignet“.

(a) Die primale Lösung $u_N(\mu) \in X_N$ ist gegeben durch

$$b(u_N(\mu), v; \mu) = g(v; \mu) \qquad \forall v \in Y_N. \tag{9.12}$$

(b) Die duale Lösung $z_{\widetilde{N}}(\mu) \in \widetilde{Y}_{\widetilde{N}}$ ist gegeben durch

$$b(w, z_{\widetilde{N}}(\mu); \mu) = -\ell(w) \qquad \forall w \in \widetilde{X}_{\widetilde{N}}. \tag{9.13}$$

(c) Primales bzw. duales Residuum $r_N(\mu) \in Y'$, $\tilde{r}_{\widetilde{N}}(\mu) \in X'$:

$$\begin{aligned} r_N(v;\mu) &:= g(v;\mu) - b(u_N(\mu), v; \mu) = b(e_N(\mu), v; \mu), \qquad v \in Y, \quad (9.14) \\ \tilde{r}_{\widetilde{N}}(w;\mu) &:= -\ell(w) - b(w, z_{\widetilde{N}}(\mu); \mu) = b(w, \tilde{e}_{\widetilde{N}}(\mu); \mu), \qquad w \in X, \quad (9.15) \end{aligned}$$

mit dem dualen Fehler

$$\tilde{e}_{\widetilde{N}}(\mu) := z^{\mathcal{N}}(\mu) - z_{\widetilde{N}}(\mu).$$

(d) Die RB-Ausgabe ist definiert als

$$s_N(\mu) := \ell(u_N(\mu)) - r_N(z_{\widetilde{N}}(\mu); \mu). \tag{9.16}$$

Lemma 9.10:
Es gilt

$$\|\tilde{e}_{\widetilde{N}}(\mu)\|_Y \leq \widetilde{\Delta}_{\widetilde{N}}(\mu) := \frac{\|\tilde{r}_{\widetilde{N}}(\mu)\|_{X'}}{\beta^*_{\mathrm{LB}}}$$

mit

$$0 < \beta^*_{\mathrm{LB}} \leq \beta^*(\mu) = \inf_{v \in Y} \sup_{w \in X} \frac{b(w, v; \mu)}{\|w\|_X \cdot \|v\|_Y}.$$

Beweis. Es ist

$$\beta^*_{\mathrm{LB}} \cdot \|\tilde{e}_{\widetilde{N}}(\mu)\|_Y \leq \sup_{w \in X} \frac{b(w, \tilde{e}_{\widetilde{N}}(\mu); \mu)}{\|w\|_X} = \sup_{w \in X} \frac{\tilde{r}_{\widetilde{N}}(w; \mu)}{\|w\|_X} = \|\tilde{r}_{\widetilde{N}}(\mu)\|_{X'}$$

□

Frage: Wie kommt man an β^*_{LB}? Eine Option: min-θ, aber nicht gut, da sich auch hier der Aufwand verdoppelt. Alternative:

Satz 9.11 (*Nečes*, 1962)**:**
Sei $b : X \times Y \to \mathbb{R}$ eine stetige Bilinearform. Dann hat das Problem

$$\text{Finde } u \in X : \ b(u, w) = g(w) \qquad \forall w \in Y$$

für $g \in Y'$ eine eindeutige Lösung, die stetig von g abhängt, genau dann wenn eine der folgenden äquivalenten Bedingungen gilt:

(a) $\exists \alpha > 0 : \ \sup_{w \in Y} \frac{b(v,w)}{\|w\|_Y} \geq \alpha \cdot \|v\|_X$ *und für jedes* $0 \neq w \in Y \ \exists v \in X : \ b(v, w) \neq 0.$
(b) Es gilt $\beta > 0, \ \beta^* > 0.$
(c) $\exists \alpha > 0$ *mit* $\beta = \beta^* = \alpha.$

Falls das Problem korrekt gestellt ist, gilt $\beta = \beta^*$ (bzw. $\beta(\mu) = \beta^*(\mu)$) und wir können die gleichen unteren Schranken wählen

$$\beta^*_{\mathrm{LB}} = \beta_{\mathrm{LB}}. \tag{9.17}$$

Proposition 9.12:
Es gilt

$$|s^{\mathcal{N}}(\mu) - s_N(\mu)| \leq \Delta^s_N(\mu) := \frac{\gamma(\mu)}{\beta^2_{LB}} \|r_N(\mu)\|_{Y'} \cdot \|\tilde{r}_{\widetilde{N}}(\mu)\|_{X'}$$

für s_N aus (9.16).

Beweis. Es gilt

$$\begin{aligned}
|s^{\mathcal{N}}(\mu) - s_N(\mu)| &\overset{Def.}{=} |\ell(u^{\mathcal{N}}(\mu)) - \ell(u_N(\mu)) + r_N(z_{\widetilde{N}}(\mu);\mu)| \\
&= |\ell(e_N(\mu)) + g(z_{\widetilde{N}}(\mu);\mu) - b(u_N(\mu), z_{\widetilde{N}}(\mu);\mu)| \\
&= |-b(e_N(\mu), z^{\mathcal{N}}(\mu);\mu) + b(u^{\mathcal{N}}(\mu), z_{\widetilde{N}}(\mu);\mu) - b(u_N(\mu), z_{\widetilde{N}}(\mu);\mu)| \\
&= |-b(e_N(\mu), z^{\mathcal{N}};\mu) + b(e_N(\mu), z_{\widetilde{N}}(\mu);\mu)| \\
&= |b(e_N(\mu), \tilde{e}_{\widetilde{N}}(\mu);\mu| \\
&\leq \gamma(\mu) \cdot \|e_N(\mu)\|_X \cdot \|\tilde{e}_{\widetilde{N}}(\mu)\|_Y.
\end{aligned}$$

□

- Wir erhalten also den „quadratischen Effekt“ - ein wesentlicher Vorteil.
- Greedy: $\Delta_N(\mu) \to X_N$
  $\widetilde{\Delta}_{\widetilde{N}}(\mu) \to \widetilde{Y}_{\widetilde{N}}$
- offline-online-Zerlegung ist wie beim primalen Problem
- im s.p.d.-compliant-Fall ist Proposition 9.12 gerade Lemma 9.2 bzw. 9.5.

10 Eine empirische Interpolationsmethode (EIM) für nichtaffine Probleme

Die Annahmen in Definition 2.1 (affine Abhängigkeit) sind für die Effizienz (online-offline-Zerlegung) von fundamentaler Bedeutung. Was tun, wenn dies nicht der Fall ist?

Wir betrachten zunächst nur die rechte Seite, also $g(v;\mu)$. Da $v : \Omega \to \mathbb{R}$ fassen wir g als Funktion $g : \Omega \times \mathcal{D} \to \mathbb{R}$ auf. Gesucht ist also eine Approximation

$$g(x;\mu) \approx I_Q(g(\cdot\,;\mu))(x) := \sum_{q=1}^{Q} \theta_q^q(\mu)\, g^q(x) \qquad \text{(also der Form (2.5))} \tag{10.1}$$

mit

- skalaren Funktionen $\theta_g^q : \mathcal{D} \to \mathbb{R}$,
- einer „kollateralen reduzierten Basis“ $G_Q = \{g^q\}_{q=1,\dots,Q}$.

Bemerkung 10.1:
Man könnte (ähnlich der Suche nach „optimalen“ snapshots) versuchen, (10.1) durch Optimierung zu lösen, dies führt aber zu einem hochdimensionalen Problem. Man könnte auch $g(\cdot\,;\mu)$ in eine Taylor-Reihe entwickeln, was aber die Berechnung von Ableitungen notwendig macht und nur eine lokale Approximation liefert. Wir verwenden deswegen auch hier einen snapshot-basierten Ansatz.

Die EIM stammt von

- Barrault, Maday, Nguyen, Patera (2004)
- Maday, Nguyen, Patera, Pau (2007)
- Drohmann, Haasdonk, Ohlberger (2012)

Definition 10.2:
Sei $G \subset \mathcal{C}(\overline{\Omega}, \mathbb{R})$, $\dim(\mathrm{span}(G)) < \infty$. Für $1 \le Q \le \dim(\mathrm{span}(G))$, $Q \in \mathbb{N}$, definiere

- *$T_Q \subset \overline{\Omega}$: Interpolationspunkte-Menge (Knotenmenge)*
- *$G_Q \subset \mathrm{span}(G)$: kollaterale reduzierte Basis*

durch

(a) $Q = 1$:

$$\begin{aligned} \tilde{q}_1 &:= \arg\max_{g\in G} \|g\|_{L_\infty(\Omega)}, \\ x_1 &:= \arg\max_{x\in\overline{\Omega}} |\tilde{q}_1(x)|, \qquad T_1 = \{x_1\}, \\ g_1 &:= \frac{\tilde{q}_1}{\tilde{q}_1(x_1)}, \qquad G_1 = \{g_1\} \end{aligned}$$

(b) $Q > 1$:

$$\begin{aligned} (10.2)\quad \tilde{q}_Q &:= \arg\max_{g\in G} \|g - I_{Q-1}g\|_{L_\infty(\Omega)} \qquad \textit{mit } I_Q g := P(g|x_1,\ldots,x_Q), \\ r_Q &:= \tilde{q}_Q - I_{Q-1}\tilde{q}_Q, \\ x_Q &:= \arg\max_{x\in\overline{\Omega}} |r_Q(x)|, \qquad T_Q = T_{Q-1}\cup\{x_Q\}, \\ g_Q &:= \frac{r_Q}{r_Q(x_Q)}, \qquad G_Q = G_{Q-1}\cup\{g_Q\}. \end{aligned}$$

Bemerkung 10.3:
(a) g_Q und x_Q müssen nicht eindeutig sein,
(b) G_Q ist nicht orthogonal, auch nicht nodal,
(c) G_Q ist hierarchisch: $G_{Q-1} \subset G_Q$,
(d) offenbar ist die EIM Greedy-artig. Der Fehler klingt i.A. aber nicht monoton ab.

Lemma 10.4:
Seien T_Q, Q_Q gemäß Definition 2.1. Dann ist

(a) Die Matrix $\underline{G}_Q := [g_j(x_i)]_{i,j=1,\ldots,Q} \in \mathbb{R}^{Q\times Q}$ ist eine untere Dreiecks-Matrix mit Diagonale 1*.*
(b) Seien $g \in C(\overline{\Omega})$, $\underline{g}_Q := (g(x_i))_{i=1,\ldots,Q}$ und $\underline{\alpha}_Q \in \mathbb{R}^Q$ Lösung von $\underline{G}_Q\underline{\alpha}_Q = \underline{g}_Q$, dann gilt:

$$(10.3)\qquad I_Q g = \sum_{i=1}^{Q} \alpha_i g_i.$$

Beweis. (a) Seien $1 \le i, j \le Q$, dann gilt nach Definition 2.1:

$$(\underline{G}_Q)_{i,i} = g_i(x_i) = \frac{r_i(x_i)}{r_i(x_i)} = 1$$

und für $j > 1$:

$$(\underline{G}_Q)_{i,j} = g_j(x_i) = \frac{1}{r_j(x_j)} r_j(x_i) = \frac{1}{r_j(x_j)} \underbrace{(\tilde{q}_j - I_{j-1}\tilde{q}_j)}_{=0,\ \text{da } i<j}(x_i) = 0$$

$\Rightarrow$ (a).

(b) Sei $1 \leq i \leq Q$. Wegen

$$\sum_{j=1}^{Q} \alpha_j g_j(x_i) = \sum_{j=1}^{Q} (\underline{G}_Q)_{i,j} \alpha_j = (\underline{G}_Q \underline{\alpha}_Q)_i \overset{\text{(LGS)}}{=} (g_Q)_i = g(x_i)$$

interpoliert $\sum_{j=1}^{Q} \alpha_j g_j$ die Funktion an den Stellen $x_1, \ldots, x_Q$, also ist

$$I_Q g = \sum_{j=1}^{Q} \alpha_j g_j.$$

□

Bemerkung 10.5:
Das obige Lemma sichert die Wohldefiniertheit des Interpolationsproblems und auch dessen effiziente Lösbarkeit.

Bemerkung 10.6:
Für das Gebiet $\Omega \subset \mathbb{R}^d$ ist das lineare Programm nicht realisierbar, ebenso gilt diese Aussage für $\dim(\text{span}(G)) = \infty$. *Man wählt also $\overline{\Omega}_h \subset \overline{\Omega}$ endlich,* $\dim(\text{span}(G)) < \infty$, *beide Größen beeinflussen die Komplexität des Problems aber maßgeblich!*

Beispiel 10.7:
$G := \{1, x, x^2\}$, $\overline{\Omega} = [-1, 1]$

(a) $Q = 1$:

$$\begin{aligned}
&\tilde{q}_1 \quad \textit{ist beliebig, wähle z.B. } \tilde{q}_1 \equiv 1 \\
&x_1 \quad \textit{ist beliebig, wähle } x_1 = -1,\ T_1 = \{-1\} \\
&g_1 \quad = \frac{\tilde{q}_1}{\tilde{q}_1(x_1)} = \tilde{q}_1,\ G_1 = \{g_1\}
\end{aligned}$$

(b) $Q = 2$:

$$\begin{aligned}
\tilde{q}_2 &= \arg\max_{g \in G} \|g - I_1 g\|_\infty \\
&= \arg\max_{g \in \{1,x,x^2\}} \|g - P(g|-1)\|_\infty \\
&= \arg\max \left\{0, \|x+1\|_\infty, \|x^2-1\|_\infty\right\} = x \\
r_2 &= \tilde{q}_2 - I_1 \tilde{q}_2 = x + 1 \\
x_2 &= \arg\max_{x \in [-1,1]} |r_2(x)| = 1, \quad T_2 = \{-1, 1\} \\
g_2 &= \frac{r_2}{r_2(x_2)} = \frac{x+1}{2} = \frac{1}{2}(x+1), \quad G_2 = \left\{1, \frac{1}{2}(x+1)\right\}
\end{aligned}$$

$Q = 3$:

$$\begin{aligned}
\tilde{q}_3 &= \arg\max_{g\in\{1,x,x^2\}} \|g - P(g|-1,1)\|_\infty \\
&= \arg\max \left\{0, 0, \|x^2-1\|_\infty\right\} = x^2 \\
r_3 &= \tilde{q}_3 - I_2\tilde{q}_3 = x^2 - 1 \\
x_3 &= \arg\max_{x\in[-1,1]} |r_3(x)| = 0,\ T_3 = \{-1, 1, 0\} \\
g_3 &= \frac{r_3}{r_3(x_3)} = \frac{x^2-1}{-1} = 1 - x^2,\ G_3 = \left\{1, \frac{1}{2}(x+1), 1-x^2\right\}
\end{aligned}$$

Bemerkung 10.8:
Wenn man dieses Beispiel für $G = \mathcal{P}_n$ durchführt kann man zeigen:

(a) $\|g_i\|_\infty = 1 \qquad \forall\ 1 \le i \le n$
(b) $g_j(x_i) = 0$ *für* $i < j$ *(Interpolations-Bedingung)*
(c) $g_j(x_j) = 1$
(d) T_M *approximiert die Chebyshev-Knoten, die bekanntermaßen für* $\mathcal{P}_n$ *optimal sind* $\rightsquigarrow$ *„magic points“*

Nun zur Fehleranalyse:

Definition 10.9:
Sei $I_M f = P(f|x_1, \ldots, x_M)$, $f \in \mathcal{C}(\overline{\Omega})$, $\{x_i\}_{i=1,\ldots,M} \subset \overline{\Omega}$ *und* $\{\xi_i\}_{i=1,\ldots,M}$ *sei eine nodale Basis, d.h.* $\xi_i(x_j) = \delta_{ij}$. *Die Zahl*

$$\Lambda_M := \max_{x\in\overline{\Omega}} \sum_{m=1}^{M} |\xi_m(x)| \tag{10.4}$$

heißt Lebesgue-Konstante.

Bemerkung 10.10:
Die Funktionen $\{g_1, \ldots, g_{M_G}\}$, $M_G := \dim(\mathrm{span}(G))$ *bilden eine Basis für* G, *da* $\underline{G}_{M_G}$ *invertierbar ist, also eine Basis-Transformation ist. („kollaterale Basis“)*

Satz 10.11:
Es gilt

$$\|u - I_M u\|_\infty \le (1 + \Lambda_M) \inf_{v_M\in X_M} \|u - v_M\|_\infty \tag{10.5}$$

für $u \in \mathcal{C}^0(\overline{\Omega})$ *mit* $X_M := \mathrm{span}\,\{\xi_i\}_{i=1,\ldots,M}$.

Beweis. Sei $u \in \mathcal{C}^0(\overline{\Omega})$, $x \in \overline{\Omega}$ und $v_M = \sum_{i=1}^M \alpha_i\xi_i \in X_M$

$$\begin{aligned}
\Rightarrow |u(x) - I_M u(x)| &= \left| u(x) - v_M(x) + \sum_{i=1}^M \alpha_i\xi_i(x) - \sum_{i=1}^M u(x_i)\xi_i(x) \right| \\
&\overset{\text{Dreiecks-Ungl.}}{\le} |u(x) - v_M(x)| + \left| \sum_{i=1}^M \alpha_i\xi_i(x) - \sum_{i=1}^M u(x_i)\xi_i(x) \right|.
\end{aligned}$$

Für den zweiten Term gilt:

$$\begin{aligned}\left|\sum_{i=1}^{M}(\alpha_i - u(x_i)\xi(x)\right| &= |\sum_{i=1}^{M}\xi_i(x)\Big\{\underbrace{\sum_{j=1}^{M}\alpha_j\overbrace{\xi_j(x_i)}^{=\delta_{ij}}}_{=v_M(x_i)}\overbrace{-u(x_i)}^{=\alpha_i}\Big\}| \\ &\overset{\text{Dreiecks-}}{\underset{\text{Ungl.}}{\leq}} \sum_{i=1}^{M}|\xi_i(x)|\,|v_M(x_i)-u(x_i)| \\ &\leq \|v_M-u\|_\infty \cdot \underbrace{\max_{x\in\overline{\Omega}}\sum_{i=1}^{M}|\xi(x_i)|}_{=\Lambda_M}.\end{aligned}$$

□

Proposition 10.12:
Für die EIM gilt $\Lambda_M \leq 2^M - 1$.

Beweis. Es gilt für $1 \leq j \leq M$ und $x \in \overline{\Omega}$

$$\xi_j(x) = g_j(x) - \sum_{\substack{i=1\\ i\neq j}}^{M} g_j(x_i)\xi_i(x)$$

(beide Seiten sind aus $\mathcal{P}_{M-1}$ und stimmen an den M Stellen $x_1, \ldots, x_M$ überein), also:

$$\begin{aligned}|\xi_j(x)| &= |g_j(x) - \sum_{\substack{i=1\\ i\neq j}}^{M} g_j(x_i)\xi_i(x)| = |g_j(x) - \sum_{i=j+1}^{M} g_j(x_i)\xi_i(x)| \\ (*) &\leq \underbrace{|g_j(x)|}_{\leq 1} + \sum_{i=j+1}^{M}\underbrace{|g_j(x_i)|}_{\leq 1}\,|\xi_i(x)| \text{ (da } x_Q = \arg\max_{x\in\overline{\Omega}}|r_Q(x)|,\ g_Q := \frac{r_Q}{r_Q(x_Q)}),\end{aligned}$$

sowie

$$|\xi_M(x)| = |g_M(x)| \leq 1.$$

Damit gilt

$$|\xi_{M-1}(x)| \leq 1 + |\xi_M(x)| \leq 2$$

und

$$|\xi_{M+1-j}(x)| \overset{(*)}{\leq} 1 + \sum_{i=M+2-j}^{M}|\xi_i(x)| \overset{\text{induktiv}}{\leq} 1+1+2+\ldots+2^{j-2} = 2^{j-1}, \qquad 2 \leq j \leq M$$

und damit

$$\Lambda_M = \sup_{x\in\overline{\Omega}}\sum_{j=1}^{M}|\xi_j(x)| \leq \sum_{j=1}^{M}2^{j-1} = 2^{M-1}.$$

□

Bemerkung 10.13:
Wir sehen im Beweis, dass obige Abschätzung sehr pessimistisch ist (Dreiecks-Ungl.). Oft sieht man in numerischen Experimenten deutlich bessere Schranken. Allerdings ist - wie bei der Polynominterpolation - die Abschätzung scharf in dem Sinne, dass es Beispiele gibt, in denen Gleichheit gilt. Es exisitieren aber quantitative a-priori Konvergenzaussagen, vgl. Maday, Nguyen, Patera, Pau, 2009:

Satz 10.14:
Seien $Z_1 \subset Z_2 \subset \ldots \subset \operatorname{span}(G),\ \dim(Z_M) = M$ *mit*

$$\inf_{v_M \in Z_M} \|u - v_M\|_\infty \leq c \cdot e^{-\alpha M} \qquad \forall u \in G,\ M \in \mathbb{N},\ c > 0, \alpha > \log(4). \tag{10.6}$$

Dann gilt für die EIM-Approximation

$$\|u - I_Q u\|\infty \leq c \cdot e^{-(\alpha - \log(4))M}.$$

Dieser Satz besagt, dass die EIM „quasi-optimal“ ist. Man kann also hoffen, mit wenigen Termen auskommen zu können. Dazu ist es für die Praxis noch wichtiger, einen a-posteriori Fehlerschätzer zu haben. Neue Arbeiten zu Term-Reduktion: Eftang (2011), Tonn (2012).

Satz 10.15:
Seien $I_Q,\ I_{Q'} : \mathcal{C}^0(\overline{\Omega}) \to \operatorname{span}(G)$ *EIM-Operatoren,* $Q' > Q, G_Q \subset G_{Q'},\ T_Q \subset T_{Q'}$. *Seien*

$$\begin{aligned} \underline{G}_{Q,Q'} &:= [g_j(x_i)]_{i,j=Q+1,\ldots,Q'},\ \underline{g}' := (g(x_i) - (I_Q g)(x_i))_{i=Q+1,\ldots,Q'}, \quad g \in G, \\ \underline{K}_{Q,Q'} &:= \left[(g_i, g_j)_{L_2(\Omega)}\right]_{i,j=Q+1,\ldots,Q'},\ \underline{\alpha}' = (\alpha'_i)_{i=Q+1,\ldots,Q'} := \underline{G}'_{Q,Q'}\underline{g}'. \end{aligned}$$

Falls $g \in \operatorname{span}(G_{Q'})$, *dann gelten*

$$\|g - I_Q(g)\|_\infty \leq \Delta_{Q,Q',\infty}(g) := \|\underline{\alpha}'\|_1 \tag{10.7}$$

$$\|g - I_Q(g)\|_{L_2} \leq \Delta_{Q,Q',0}(g) := ((\underline{\alpha}')^T \underline{K}_{Q,Q'}\underline{\alpha}')^{1/2} \tag{10.8}$$

Beweis. Nach (10.3) gilt

$$I_Q g = \sum_{i=1}^{Q} \alpha_i g_i, \qquad I_{Q'} g = \sum_{i=1}^{Q'} \alpha'_i g_i$$

mit

$$\underline{G}_Q \underline{\alpha}_Q = \underline{g}_Q, \qquad \underline{G}_{Q'} \underline{\alpha}_{Q'} = \underline{g}_{Q'}$$

sowie

$$\underline{G}_{Q'} := \left(\begin{array}{ccc|ccc} 1 & & 0 & & & \\ & \ddots & & & 0 & \\ * & & 1 & & & \\ \hline & & & 1 & & 0 \\ & * & & & \ddots & \\ & & & * & & 1 \end{array}\right)$$

$\Rightarrow \alpha_= \alpha'_i,\ 1 \leq i \leq Q$ und

$$g - I_Q(g) \overset{g \in \operatorname{span}(G_{Q'})}{=} I_{Q'}(g) - I_Q(g) = \sum_{i=Q+1}^{Q'} \alpha'_i g_i.$$

Wegen $\|g\|_\infty = 1$ (Bem. 10.8 (a)) ist damit (10.7) klar. Weiter folgt:

$$\|g - I_Q(g)\|_{L_2}^2 = \sum_{i,j=Q+1}^{Q'} \alpha'_i \alpha'_j (g_i\,, g_j)_{L_2} = (\Delta_{Q,Q',0}(g))^2.$$

□

Bemerkung 10.16:
Im Allgemeinen gilt $g \in \operatorname{span}(G_{Q'})$ *natürlich* nicht *notwendigerweise, also ist* $\Delta_{Q,Q',\cdot}$ *nicht unbedingt eine rigorose obere Schranke. Man untersucht die Effektivität daher meist in numerischen Experimenten.*

Nun zur Anwendung der EIM für die RBM: Wir können die EIM auch für Bilinearformen durchführen und erhalten

$$b(w, v; \mu) \approx b_Q(w, v; \mu) = \sum_{q=1}^{Q} \theta_b^q(\mu) b^q(w, v)$$

sowie für die rechte Seite

$$f(v; \mu) \approx f_Q(v; \mu) = \sum_{q=1}^{Q} \theta_f^q(\mu) \underbrace{(g^q\,, v)_0}_{=:\, f^q(v)} =: \sum_{q=1}^{Q} \theta_f^q(\mu) f^q(v),$$

wobei hier auch $Q^b \neq Q^f$ gewählt werden kann ($Q := \max\{Q^b, Q^f\}$).

<u>offline</u>:

- $\underline{B}^q = [b^q(\xi_i, \xi_j)]_{i,j=1,\ldots,N}$ werden offline berechnet: $\mathcal{O}(N^2\mathcal{N})$
- ebenso $\underline{f}^q = (f^q(\xi_i))_{i=1,\ldots,N}$: $\mathcal{O}(N\mathcal{N})$

$\Rightarrow$ Q Matrizen und Q Vektoren berechnen und speichern. $\mathcal{O}(Q(N^2 + N))$ Speicher, $\mathcal{O}(Q\mathcal{N}(N^2 + N))$ Berechnungen.

<u>online</u>:

- die Summationen werden online durchgeführt: $\mathcal{O}(Q(N^2 + N))$

$\Rightarrow$ Q beeinflusst maßgeblich den Speicheraufwand sowie die online-Komplexität.

Oft haben nicht-affine Abhängigkeiten folgende Struktur:

$$b(w, v, g(x; \mu)) \quad \text{bzw.} \quad f(w, g(x; \mu)).$$

Beispiel:

- $-\Delta u + \vec{\beta} u \cdot \nabla u = f \quad in \;\; \Omega_\mu$
- $u = 0 \quad auf \;\; \partial B_\mu$
- *Randbedingungen auf* ∂D

B_0 bzw. Ω_0 sei ein Refernzgebiet und $T_\mu : \Omega_\mu \to \Omega_0$.

$$\hat{u}(\hat{x}) := u(T_\mu^{-1}\hat{x}), \;\; \hat{x} \in \Omega_0$$

$$\begin{aligned} \Rightarrow \int_{\Omega_\mu} -\Delta u(x)\, v(x)\, dx &= \int_{\Omega_\mu} \nabla_x u(x) \cdot \nabla_x v(x)\, dx = \ldots \\ &= \int_{\Omega_0} \nabla_{\hat{x}} \hat{u}(\hat{x}) \cdot \nabla_{\hat{x}} \hat{v}(\hat{x})\, T_\mu(\hat{x})\, d\hat{x} \end{aligned}$$

→ *Parameter-Abhängigkeit vom Integrations-Gebiet auf Integranden verschoben*

$$\int_{\Omega_\mu} \vec{\beta}(x)\, u(x) \cdot \nabla u(x)\, v(x)\, dx = \ldots = \int_{\Omega_0} \vec{\beta}(\hat{x})\, \hat{u}(\hat{x}) \cdot \nabla_{\hat{x}} \hat{u}(\hat{x})\, \hat{v}(\hat{x})\, t_\mu(\hat{x})\, d\hat{x}$$

→ *Bilinearform lautet: $a_0(u, v, g_0(\mu)) + a_1(u, u, v, g_1(\mu))$, rechte Seite:*

$$\int_{\Omega_\mu} f(x)\, v(x)\, dx = \int_{\Omega_0} \hat{f}(\hat{x})\, \hat{v}(\hat{x})\, |JT_\mu^{-1}(\hat{x})|\, d\hat{x},$$

also $f(v, g_2(\mu))$. Man nutzt die EIM dann für $g(x;\mu)$ aus.

$$g(x;\mu) \approx g_Q(x;\mu) = \sum_{q=1}^{Q} \theta_q(\mu) g_q(x)$$

und die Bilinearform / rechte Seite hängt linear *von den Parameterfunktionen ab. Das EIM-RBM-Problem lautet dann: Bestimme $u_{N,Q}(\mu) \in X_N$ mit*

$$b(u_{N,Q}(\mu), v_N, g_Q(\cdot\,;\mu)) = f(v_N, g_Q(\cdot\,;\mu)) \qquad \forall v_N \in Y_N \tag{10.9}$$

$$s_{N,Q}(\mu) := \ell(u_{N,Q}(\mu)) \tag{10.10}$$

Frage: Welche Auswirkungen hat die EIM auf die RBM bezüglich des Fehlers?
Dazu einige Definitionen:

$$\varepsilon_Q(\mu) := \|g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)\|_\infty \tag{10.11}$$

$$\phi_1(\mu) := \frac{1}{\varepsilon_Q(\mu)} \|f(\cdot, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu))\|_{Y'} \tag{10.12}$$

$$\phi_2(\mu) := \frac{1}{\varepsilon_Q(\mu)} \sup_{v \in X} \|b(v, \cdot, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu))\|_{Y'} \tag{10.13}$$

$$\phi_3(\mu) := \|f(\cdot, g_Q(\cdot\,;\mu))\|_{Y'} \tag{10.14}$$

Satz 10.17 (EIM-RBM a-priori-Abschätzung)**:**
Die Bilinearform sei inf-sup-stabil und stetig. Falls

$$\varepsilon_Q(\mu) \leq \frac{1}{2}\,\frac{\beta_N(\mu)}{\phi_2(\mu)}, \tag{10.15}$$

dann gilt

$$\begin{aligned}\|u(\mu) - u_{N,Q}(\mu)\|_X \;\leq\; & \left(1 + \frac{\gamma_N(\mu)}{\beta_N(\mu)}\right) \inf_{v_N \in X_N} \|u(\mu) - v_N\|_X \\ & + \varepsilon_Q(\mu)\,\frac{\phi_1(\mu)\beta_N(\mu) + 2\phi_2(\mu)\phi_3(\mu)}{\beta_N^2(\mu)}\end{aligned} \tag{10.16}$$

Beweis. Sei $z_N \in X_N$ beliebig, dann gilt

$$\begin{aligned}\beta_N(\mu)\,\|z_N - u_{N,Q}(\mu)\|_X &\leq \sup_{w_N \in Y_N} \frac{b(z_N - u_{N,Q}(\mu), w_N, g(\cdot\,;\mu))}{\|w_N\|_Y} \\ &\leq \underbrace{\sup_{w_N \in X_N} \frac{b(z_N - u(\mu), w_N, g(\mu))}{\|w_N\|_Y}}_{\leq \gamma_N(\mu)\,\|z_N - u(\mu)\|_X} \underbrace{\sup_{w_N \in Y_N} \frac{b(u(\mu) - u_{N,Q}(\mu), w_N, g(\cdot\,;\mu))}{\|w_N\|_Y}}_{?}.\end{aligned}$$

Für den zweiten Term gilt:

$$\begin{aligned}b(u(\mu) - u_{N,Q}(\mu), w_N, g(\cdot\,;\mu)) &= f(w_N, g(\cdot\,;\mu)) - b(u_{N,Q}(\mu), w_N, g(\cdot\,;\mu)) \\ &= f(w_N, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)) - b(u_{N,Q}(\mu), w_N, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)) \\ &\leq \|w_N\|_Y \Big\{ \|f(\cdot, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu))\|_{Y'} \\ &\quad + \|w_N\|_Y \|u_{N,Q}(\mu)\|_X \cdot \sup_{v \in X} \|b(v, \cdot, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu))\|_{Y'} \Big\} \\ &\leq \|w_N\|_Y \left\{\varepsilon_Q(\mu)\phi_1(\mu) + \varepsilon_Q(\mu)\|u_{N,Q}(\mu)\|_X \phi_2(\mu)\right\} \qquad (10.17)\end{aligned}$$

Weiter gilt:

$$\begin{aligned}\beta_N(\mu)\,\|u_{N,Q}(\mu)\|_X \;&\leq\; \sup_{w_N \in Y_N} \frac{b(u_{N,Q}(\mu), w_N, g(\mu))}{\|w_N\|_Y} \\ &\leq\; \sup_{w_N \in Y_N} \frac{b(u_{N,Q}(\mu), w_N, g_Q(\mu))}{\|w_N\|_Y} \\ &\quad + \sup_{w_N \in Y_N} \frac{b(u_{N,Q}(\mu), w_N, g(\mu) - g_Q(\mu))}{\|w_N\|_Y} \\ &=\; \sup_{w_N \in Y_N} \frac{f(w_N, g_Q(\mu))}{\|w_N\|_Y} + \sup_{w_N \in Y_N} \frac{b(u_{N,Q}(\mu), w_N, g(\mu) - g_Q(\mu))}{\|w_N\|_Y} \\ &\leq\; \phi_3(\mu) + \varepsilon_Q(\mu)\|u_{N,Q}(\mu)\|_X\,\phi_2(\mu) \\ &\overset{(10.15)}{\leq}\; \phi_3(\mu) + \frac{1}{2}\beta_N(\mu)\|u_{N,Q}(\mu)\|_X,\end{aligned}$$

also

$$\|u_{N,Q}(\mu)\|_X \leq 2\frac{\phi_3(\mu)}{\beta_N(\mu)}. \tag{10.18}$$

Dies setzen wir in (10.17) ein und erhalten

$$\begin{aligned} b(u(\mu)-u_{N,Q}(\mu),w_N,g(\mu)) &\leq \|w_N\|_Y\varepsilon_Q(\mu)\left\{\phi_1(\mu)+2\frac{\phi_2(\mu)\phi_3(\mu)}{\beta_N(\mu)}\right\} \\ &= \|w_N\|_Y\frac{\varepsilon_Q(\mu)}{\beta_N(\mu)}(\phi_1(\mu)\beta_N(\mu)+2\phi_2(\mu)\phi_3(\mu)) \end{aligned}$$

und damit

$$\|z_N-u_{N,Q}(\mu)\|_X \leq \frac{\gamma_N(\mu)}{\beta_N(\mu)}\|z_N-u(\mu)\|_X+\varepsilon_Q(\mu)\frac{\phi_1(\mu)\beta_N(\mu)+2\phi_2(\mu)\phi_3(\mu)}{\beta_N^2(\mu)}.$$

Schließlich:

$$\begin{aligned} \|u(\mu)-u_{N,Q}(\mu)\|_X &\leq \|u(\mu)-z_N\|_X+\|z_N-u_{N,Q}(\mu)\|_X \\ &\leq \left(1+\frac{\gamma_N(\mu)}{\beta_N(\mu)}\right)\|z_N-u(\mu)\|_X \\ &\quad +\varepsilon_Q(\mu)\frac{\phi_1(\mu)\beta_N(\mu)+2\phi_2(\mu)\phi_3(\mu)}{\beta_N^2(\mu)} \end{aligned}$$

und Bildung des Infimums über $z_N \in X_N$ liefert die Behauptung. □

Bemerkung 10.18:
(a) Die Voraussetzung (10.15) erscheint wegen der EIM-Konvergenz realisitisch.
(b) Für den ersten Term in (10.16) erwartet man wegen der RB-Konvergenz (exponentielles) Abklingen. Da die ϕ_i Regularitätsterme sind, sollte man Q so wählen, dass beide Terme in (10.16) die gleiche Größenordnung besitzen.

Nun zur a-posteriori-Fehlerschätzung. Wir definieren den EIM-RBM-Fehlerschätzer als

$$\Delta_{N,Q}(\mu) := \frac{1}{\beta_{LB}(\mu)}\|r(\cdot,g_Q(\cdot\,;\mu))\|_{Y'}+\frac{\hat{\varepsilon}_Q(\mu)}{\beta_{LB}}\|r(\cdot,g_{Q+1}(\cdot\,;\mu)\|_{Y'}, \tag{10.19}$$

wobei

$$\hat{\varepsilon}_Q(\mu) := |g(x_{Q+1};\mu)-g_Q(x_{Q+1};\mu)|$$

ein Fehlerschätzer für $\varepsilon_Q(\mu)$ aus (10.11) ist.

Lemma 10.19:
Sei $g(\cdot\,;\mu) \in G_{Q+1}$. Dann gilt für ε_Q gemäß (10.11)

(i) $|g(x;\mu)-q_Q(x;\mu)| = |\hat{\varepsilon}_Q(\mu)\, g_{Q+1}(x)|$
(ii) $\varepsilon_Q(\mu) = \hat{\varepsilon}_Q(\mu)$

Beweis. Da $g(\mu) := g(\cdot\,;\mu) \in G_{Q+1}$ existieren $\kappa(\mu) \in \mathbb{R}^{Q+1}$ mit

$$g(x;\mu) - g_Q(x;\mu) = \sum_{q=1}^{Q+1} \kappa_q(\mu)\, g_q(x).$$

Setze hier $x = x_i$ (magic points) ein $(1 \le i \le Q+1)$

$$\Rightarrow \sum_{q=1}^{Q+1} \kappa_q(\mu)\, g_q(x_i) = g(x_i;\mu) - g_Q(x_i;\mu) = 0 \qquad 1 \le i \le Q,$$

wegen der Interpolationseigenschaft der EIM.

- da $g_m(x_i) = (\underline{G}_Q)_{i,m} = 0$ für $m > i$, ist obiges ein gestaffeltes Gleichungssystem und Rückwärts-Einsetzen liefert $\kappa_1(\mu) = \ldots = \kappa_Q(\mu) = 0$

- $\kappa_{Q+1}(\mu) = \kappa_{Q+1}(\mu) \underbrace{g_{Q+1}(x_{Q+1})}_{=1} = g(x_{Q+1};\mu) - g_Q(x_{Q+1},\mu)$, also

$$\begin{aligned} |g(x;\mu) - g_Q(x;\mu)| &= |\kappa_{Q+1}(\mu)\, g_{Q+1}(x)| \\ &= |(g(x_{Q+1};\mu) - g_Q(x_{Q+1};\mu))g_{Q+1}(x)| \\ &= |\hat{\varepsilon}_Q(\mu)\, g_{Q+1}(x)| \end{aligned}$$

  und damit (i).

Zu (ii):

$$\begin{aligned} \varepsilon_Q(\mu) &= \|g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)\|_\infty = \sup_x |g(x;\mu) - g_Q(x;\mu)| \\ &\overset{(i)}{=} \sup_x |\hat{\varepsilon}_Q(\mu)\, g_{Q+1}(x)| = \hat{\varepsilon}_Q(\mu) \underbrace{\|g_{Q+1}\|_\infty}_{=1} \\ &= \hat{\varepsilon}_Q(\mu). \end{aligned}$$

□

Bemerkung 10.20:
Aus (i) folgt

$$g(x;\mu) - g_Q(x;\mu) = \pm\hat{\varepsilon}_Q(\mu)\, g_{Q+1}(x). \tag{10.20}$$

Den EIM-RBM-Output-Fehlerschätzer definieren wir als

$$\Delta^s_{N,Q} := \|\ell\|_{X'} \cdot \Delta_{N,Q}(\mu). \tag{10.21}$$

Proposition 10.21:
Falls $g(\cdot\,;\mu) \in G_{Q+1}$, dann gilt

(a) $\|u(\mu) - u_{N,Q}(\mu)\|_X \leq \Delta_{N,Q}(\mu)$
(b) $|s(\mu) - s_{N,Q}(\mu)| \leq \Delta^s_{N,Q}(\mu)$

Beweis. Für $e(\mu) := u(\mu) - u_{N,Q}(\mu)$ und $w \in Y$ beliebig gilt

$$\begin{aligned} b(e(\mu), w, g(\cdot\,;\mu)) &= b(u(\mu), w, g(\cdot\,;\mu)) - b(u_{N,Q}(\mu), w, g(\cdot\,;\mu)) \\ &= f(w, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)) + f(w, g_Q(\cdot\,;\mu) \\ &\quad - b(u_{N,Q}(\mu), w, g_Q(\cdot\,;\mu)) - b(u_{N,Q}(\mu), w, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)) \\ &= f(w, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)) + r(w, g_Q(\cdot\,;\mu)) \\ &\quad - b(u_{N,Q}(\mu), w, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)). \end{aligned}$$

Nach Voraussetzung und Lemma 10.19 gilt:

$$|g(x_{Q+1};\mu) - g_Q(x_{Q+1};\mu)| = \hat{\varepsilon}_Q(\mu) \overset{\text{Lem. 10.19}}{\underset{(ii)}{=}} \varepsilon_Q(\mu) = \|g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)\|_\infty,$$

also:

$$\begin{aligned} &|f(w, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)) - b(u_{N,Q}(\mu), w, g(\cdot\,;\mu) - g_Q(\cdot\,;\mu))| \\ &\qquad = |r(w, \underbrace{g(\cdot\,;\mu) - g_Q(\cdot\,;\mu)}_{=\pm\hat{\varepsilon}_Q(\mu)})| \\ &\qquad \leq \|r(\cdot, g_{Q+1})\|_{Y'} \cdot \|w\|_Y \cdot \hat{\varepsilon}_Q(\mu). \end{aligned}$$

Der Rest ist klar mit der inf-sup-Bedingung. □

Bemerkung 10.22:
Wenn die Voraussetzung „$g \in G_{Q+1}$“ nicht *erfüllt ist, dann ist der zweite Term in (10.19) nicht rigoros - man kann eine Art „Sicherheits-Bedingung“ einführen die garantiert, dass Q „hinreichend groß“ gewählt wird.*

Bemerkung 10.23:
Die online-Berechnung der Dualnormen geht wieder über die Riesz-Repräsentatoren der einzelnen Terme in der EIM und dann per linearer Superposition - der Genauigkeitsverlust durch den „Wurzel-Effekt“ bleibt und sollte bei der EIM mit in Betracht gezogen werden.

11 Effiziente Berechnung der Konstanten

Für die konstruierten Fehlerschätzer benötigen wir die Parameter $\alpha(\mu), \beta(\mu), \gamma(\mu)$ bzw. berechenbare Schranken. Wir wollen Alternativen zum min-θ-Verfahren kennen lernen. Sei $w \in X$ gegeben, betrachte den *supremierenden Operator*

$$T_\mu w := \arg\sup_{v \in Y} \frac{b(w,v;\mu)}{\|v\|_Y} \in Y, \qquad \text{vgl. Def. 1.5,} \tag{11.1}$$

der berechnet wird durch (vgl. Bem 1.6)

$$T_\mu w \in Y, \qquad (T_\mu w\,, v)_Y = b(w,v;\mu) \qquad \forall v \in Y. \tag{11.2}$$

Wegen

$$\|T_\mu w\|_Y \overset{\text{Bem. 1.6}}{=} \|b(w,\cdot\,;\mu)\|_{Y'} = \sup_{v \in Y} \frac{b(w,v;\mu)}{\|v\|_Y}$$

gilt

$$\beta(\mu) = \inf_{w \in X} \frac{\|T_\mu w\|_Y}{\|w\|_X}, \qquad \gamma(\mu) = \sup_{w \in X} \frac{\|T_\mu w\|_Y}{\|w\|_X},$$

also lassen sich β, γ über Rayleigh-Quotienten bestimmen. Die Matrix-Darstellung von $b(\cdot,\cdot\,;\mu)$ kennen wir im Wesentlichen aus (2.4)

$$b(w,v;\mu) = \underline{w}^T\, \underline{\mathcal{B}}^{\mathcal{N}}(\mu)\, \underline{v}, \qquad \underline{\mathcal{B}}^{\mathcal{N}}(\mu) := \left[b(\varphi_i^{\mathcal{N}}, \varphi_j^{\mathcal{N}};\mu)\right]_{i,j=1}^{\mathcal{N}}$$

und analog für die Gram-Matrix von Y,

$$\underline{Y} := \left[(\psi_i^{\mathcal{N}}\,, \psi_j^{\mathcal{N}})_Y\right]_{i,j=1}^{\mathcal{N}}.$$

Also berechnet sich

$$T_\mu w = \sum_{j=1}^{\mathcal{N}} t_j(\mu) \psi_j^{\mathcal{N}}, \qquad \underline{T}(\mu) := (t_j(\mu))_{j=1}^{\mathcal{N}}$$

aus

$$\underline{t}(\mu) = \underline{Y}^{-1}\, \underline{\mathcal{B}}^{\mathcal{N}}(\mu)^T\, \underline{w},$$

denn aus (11.2) folgt

$$(\underline{Y}\,\underline{t}(\mu))_j \overset{\underline{Y}=\underline{Y}^T}{=} \sum_{i=1}^{\mathcal{N}} (\psi_i^{\mathcal{N}}, \psi_j^{\mathcal{N}})_Y \, t_i(\mu) = (T_\mu w, \psi_j^{\mathcal{N}})_Y$$
$$\overset{(11.2)}{=} b(w, \psi_j^{\mathcal{N}}; \mu) = \sum_{i=1}^{\mathcal{N}} w_i \, b(\varphi_i^{\mathcal{N}}, \psi_j^{\mathcal{N}}; \mu) = (\underline{\mathcal{B}}^{\mathcal{N}}(\mu)^T \, \underline{w})_j.$$

Weiter gilt:

$$\begin{aligned}
\|T_\mu w\|_Y^2 &= (T_\mu w, T_\mu w)_Y = \underline{t}(\mu)^T \, \underline{Y} \, \underline{t}(\mu) \\
&\overset{\underline{Y}=\underline{Y}^T}{=} \underline{w}^T \, \underline{\mathcal{B}}^{\mathcal{N}}(\mu) \, \underline{Y}^{-1} \, \underline{Y} \, \underline{Y}^{-1} \, \underline{\mathcal{B}}^{\mathcal{N}}(\mu)^T \, \underline{w} \\
&= \underline{w}^T \, \underline{\mathcal{B}}^{\mathcal{N}}(\mu) \, \underline{Y}^{-1} \, \underline{\mathcal{B}}^{\mathcal{N}}(\mu)^T \, \underline{w} =: \underline{w}^T \, \underline{Z}(\mu) \, \underline{w}, \\
\|w\|_X^2 &= (w, w)_X = \underline{w}^T \, \underline{X} \, \underline{w}, \qquad \text{mit } \underline{X} := \left[(\varphi_i^{\mathcal{N}}, \varphi_j^{\mathcal{N}})_X\right]_{i,j=1}^{\mathcal{N}}.
\end{aligned}$$

Wir erhalten also ein verallgemeinertes Eigenwert-Problem

$$\underline{Z}(\mu) \, \underline{v} = \lambda \, \underline{X} \, \underline{v}, \tag{11.3}$$

denn:

$$\frac{\|T_\mu w\|_Y^2}{\|w\|_X^2} = \frac{\underline{w}^T \, \underline{Z}(\mu) \, \underline{w}}{\underline{w}^T \, \underline{X} \, \underline{w}} \Rightarrow \frac{\|T_\mu v\|_Y^2}{\|v\|_X^2} = \lambda$$

und dann folgt $\beta(\mu) = \sqrt{\lambda_{\min}}$, $\gamma(\mu) = \sqrt{\lambda_{\max}}$.

Probleme:

- Komplexität von (11.3) ist mindestens $\mathcal{O}(\mathcal{N})$ *für jedes* $\mu \in \mathcal{D}$ ↯
- selbst wenn b parametrisch-affin ist, d.h.

$$\underline{\mathcal{B}}^{\mathcal{N}}(\mu) = \sum_{q=1}^{Q} \theta_b^q(\mu) \, \underline{B}^q,$$

  kann man $\lambda_{\min}$, $\lambda_{\max}$ nicht einfach aus den Eigenwerten von $\underline{B}^q$, $1 \le q \le Q$, berechnen.

Alternative: *Successive Contraints Method* (SCM) (Huynh, Rozza, Sen, Patera, 2007)

Wir wollen diese in einem etwas allgemeineren Rahmen einführen, ohne dass dies die Sache komplizierter machen würde.

Seien $h_q : X \to \mathbb{R}$, $1 \le q \le Q$ gegeben (nicht notwendigerweise $h_q \in X'$) und wir wollen

$$\sigma(\mu) := \inf_{v \in X} \sum_{q=1}^{Q} \theta_q(\mu) \, h_q(v) \tag{11.4}$$

berechnen.

Beispiel 11.1:
(a) Seien $X = Y$ und b parametrisch-affin und koerziv. Mit

$$h_q(v) := \frac{b^q(v,v)}{\|v\|_X^2}$$

erhalten wir $\sigma(\mu) = \alpha(\mu)$.
(b) Sei b parametrisch-affin und

$$h_q(v) := \sup_{w \in Y} \frac{b^q(v,w)}{\|w\|_Y \, \|v\|_X},$$

dann erhalten wir $\sigma(\mu) = \beta(\mu)$. Mit der bekannten Darstellung über den supremierenden Operator T_μ von b und

$$T_\mu w = \sum_{q=1}^{Q} \theta_q(\mu)\, T_q w$$

mit T_q dem suprimierenden Operator von b_q, $1 \le q \le Q$. Dann erhalten wir eine Darstellung der Form(11.4) mit entsprechendem Q', aber ohne Supremum.
(c) Mit leichten Modifikationen von (11.4) erhalten wir $\gamma(\mu)$.

Wir definieren ein Zielfunktional ($\underline{z} = (z_1, \ldots, z_q) \in \mathbb{R}^Q$)

$$J : \mathcal{D} \times \mathbb{R}^Q \to \mathbb{R} \qquad \text{mit} \quad (\mu, \underline{z}) \mapsto J(\mu, \underline{z}) := \sum_{q=1}^{Q} \theta_q(\mu)\, z_q \tag{11.5}$$

und mit

$$S_0 := \left\{ \underline{z} \in \mathbb{R}^Q : \exists v \in X \text{ mit } z_q = h_q(v) \quad \forall 1 \le q \le Q \right\} \tag{11.6}$$

(gleiches v für alle q) gilt

$$\sigma(\mu) \overset{z_q = h_q(v)}{=} \inf_{\underline{z} \in S_0} \sum_{q=1}^{Q} \theta_q(\mu)\, z_q = \inf_{\underline{z} \in S_0} J(\mu, \underline{z}).$$

Man schränkt S_0 auf ein Polytop ein:

$$B_Q := \prod_{q=1}^{Q} \left[\sigma_q^-, \sigma_q^+\right] \subset \mathbb{R}^Q, \qquad \sigma_q^- := \inf_{v \in X} h_q(v), \quad \sigma_q^+ = \sup_{w \in X} h_q(v). \tag{11.7}$$

Wir definieren zwei Parameter-Mengen (zur Wahl später)

$$C_K := \{w_k \in \mathcal{D} : 1 \le k \le K\}, \qquad \Xi_J := \{v_j \in \mathcal{D} : 1 \le j \le J\}, \tag{11.8}$$

sowie für $M \in \mathbb{N}$ und $E \subset \mathcal{D}$

$$P_M(\mu, E) := \{M \text{ nächste Nachbarn von } \mu \text{ in } E \text{ bzgl. } \|\cdot\|_2\}$$

(mit $P_0(\mu, E) := \emptyset$ und $P_{\widetilde{M}}(\mu, E) := E$ für $\widetilde{M} \geq |E|$).
Weiter wählen wir natürliche Zahlen

$$\begin{aligned} M_\sigma &\in \mathbb{N} \quad (\hat{=} \quad \# \text{ von Satbilitätsbedingungen}) \\ M_+ &\in \mathbb{N} \quad (\hat{=} \quad \# \text{ von Positivitäts-Bedingungen}) \end{aligned}$$

und damit abgeschwächte Nebenbedingungen

$$\begin{aligned} (11.9) \quad S_{LB}(\mu, C_K, \Xi_J) \quad &= \quad \{\underline{z} \in B_Q : \\ &\quad (1) \quad J(\overline{\mu}, \underline{z}) \geq \sigma(\overline{\mu}) \quad \forall\, \overline{\mu} \in P_{M_\sigma}(\mu, C_K) \\ &\quad (2) \quad J(\overline{\mu}, \underline{z}) \geq 0 \qquad \forall\, \overline{\mu} \in P_{M_+}(\mu, \Xi_J)\big\}\,, \end{aligned}$$

$$(11.10) \qquad S_{UB}(C_K) \quad := \quad \left\{ \underline{z}^*(w_k) : 1 \leq k \leq K,\ \underline{z}^*(w) = \arg\min_{\underline{z} \in S_0} J(w, \underline{z}) \right\}$$

und schließlich die Schranken

$$\begin{aligned} (11.11) \qquad \sigma_{LB}(\mu) &= \min_{\underline{z} \in S_{LB}(\mu, C_K, \Xi_J)} J(\mu, \underline{z}) \\ \sigma_{UB}(\mu) &= \min_{\underline{z} \in S_{UB}(C_K)} J(\mu, \underline{z}). \end{aligned}$$

Satz 11.2:
Sei S_0 kompakt, dann gilt für alle $\mu \in \mathcal{D}$

$$\sigma_{LB}(\mu) \leq \sigma(\mu) \leq \sigma_{UB}(\mu).$$

Beweis.

- Sei $\underline{z} \in S_{UB}(C_K)$, dann $\exists w_k \in C_K$ mit $\underline{z} = \underline{z}^*(w_k)$, also

$$J(w_k, \underline{z}) = \min_{\tilde{\underline{z}} \in S_0} J(w_k, \tilde{\underline{z}}).$$

  Da S_0 kompakt folgt $\underline{z} \in S_0 \Rightarrow S_{UB}(C_K) \subset S_0$.

- Sei $\underline{z} \in S_0$, dann ist $\underline{z} \in B_Q$ nach Konstruktion. Weiter gilt

  (1) $\sigma(\tilde{\mu}) = \min_{\tilde{\underline{z}} \in S_0} J(\tilde{\mu}, \tilde{\underline{z}}) \qquad \forall\, \tilde{\mu} \in P_{M_\sigma}(\mu, C_K) \subset \mathcal{D} \quad (\text{sogar } \forall\, \tilde{\mu} \in \mathcal{D}),$

  (2) $J(\tilde{\mu}, \underline{z}) \geq \sigma(\tilde{\mu}) \geq 0 \qquad \forall \tilde{\mu} \in P_{M_+}(\mu, \Xi_J) \subset \mathcal{D}$
  nach Voraussetzung $\quad (\text{sogar } \forall\, \tilde{\mu} \in \mathcal{D}),$

  also $\underline{z} \in S_{LB}(\mu, C_K, \Xi_J)$ und damit

$$S_{UB}(C_K) \subset S_0 \subset S_{LB}(\mu, C_K, \Xi_J).$$

  Damit gilt

$$\underbrace{\min_{\underline{z} \in S_{LB}(\mu, C_K, \Xi_J)} J(\mu, \underline{z})}_{=: \sigma_{LB}(\mu)} \leq \underbrace{\min_{\underline{z} \in S_0} J(\mu, \underline{z})}_{=: \sigma(\mu)} \leq \underbrace{\min_{\underline{z} \in S_{UB}(C_K)} J(\mu, \underline{z})}_{=: \sigma_{UB}(\mu)}$$

□

Bemerkung 11.3:
(a) Für ein festes $\mu \in \mathcal{D}$ ist $J(\mu, \cdot)$ linear, es ist also ein lineares Programm zu lösen.
(b) Die Kompaktheitsvoraussetzung an S_0 garantiert

$$\inf_{\underline{z} \in S_0} J(\mu, z) = \min_{\underline{z} \in S_0} J(\mu, z).$$

Für die Beispiele in 11.1 ist dies erfüllt.

Online-Offline-Porzduren

Sei $\Xi_{\text{train}} \subset \mathcal{D}$ eine endliche „Trainingsmenge“. Für alle $1 \leq k \leq n_{\text{train}}$ und C_k wie oben, definiere die relative Fehlerschranke ($\Xi_j = \Xi_{\text{train}}$)

$$\varepsilon_k^*(\mu) := \frac{\sigma_{UB}(\mu, C_k) - \sigma_{LB}(\mu, C_k, \Xi_{\text{train}})}{\sigma_{UB}(\mu, C_k)}, \tag{11.12}$$

also ein Maß für die Schärfe der Schranken.

Algorithmus 11.4 (Greedy-Algorithmus zur Bestimmung von C_k)**:**

$C_1 := \{\omega_1\}$, $\omega_1 \in \Xi_{\text{train}}$ *beliebig*
$S_{UB}(C_1) := \{\underline{z}^*(\omega_1)\}$
while $\varepsilon_k^*(\mu) > \varepsilon_{tol}$ **do**
  $\omega_{k+1} = \arg\max_{\mu \in \Xi_{\text{train}}} \varepsilon_k^*(\mu)$
  $C_{k+1} = C_k \cup \{\omega_{k+1}\}$
  $S_{UB}(C_{k+1}) = S_{UB}(C_k) \cup \{\underline{z}^*(\omega_{k+1})\}$
  $k = k + 1$
end

Offline Aufwand:

(i) $2Q$ EgenwertProbleme bzgl. $\underline{B}^q \to \sigma_q^-$, σ_q^+ $\forall q$
(ii) K Eigenwertprobleme zur Bestimmung von $\sigma(\omega_k)$ in (11.9) bzw. $\underline{z}^*(\omega_k)$ in (11.10). Die $\sigma(\omega_k)$ kann man sehr effizient mit dem Lanczos-Algorithmus berechnen.
(iii) $\mathcal{O}(\mathcal{N} \cdot Q\, K)$ (Matrix-Vektor) Operationen für $S_{UB}(C_k)$
(iv) $(n_{\text{train}} \cdot K)$ lineare Programme der Größe $\mathcal{O}(2\,Q + M_\sigma + M_+)$

Online-Aufwand:

(i) Berechne $\theta_q(\mu)$: $\mathcal{O}(Q)$
(ii) Bestimme $P_{M_\sigma}(\mu, C_k)$: $\mathcal{O}(K \cdot M_\sigma)$
(iii) Berechne $\sigma_{LB}(\mu)$: $\mathcal{O}(2\,Q + M_\sigma + M_+)$ (Anzahl der Nebenbedingungen)
(iv) (Wurzel ziehen (für β))

12 Zeitabhängige Probleme

Betrachte das Anfangsrandwertproblem der Wärmeleitung für $t \in [0,T],\ T > 0$, $\Omega \subset \mathbb{R}^d$

$$\begin{aligned} &u_t - \nabla \cdot (a(t,x) \cdot \nabla u) = g(t,x) && \text{in } \Omega_T := (0,T) \times \Omega, \\ &u(t,x) = g_\Gamma(t,x) && \forall x \text{ in } \Gamma_T := (0,T) \times \partial\Omega, \\ &u(0,x) = u_0(x) && \forall x \in \Omega. \end{aligned}$$

in allgeminer Form

$$u_t + Au = f \qquad \forall t \in (0,T), \quad u(0) = u_0 \tag{12.1}$$

mit

$$A : H_0^1(\Omega) \to H^{-1}(\Omega), \quad u_0 \in H_0^1(\Omega).$$

Ist $A = A(t)$, so nennt man (12.1) auch *LTV* (linear time variant), ansonsten *LTI* (linear time invariant).

Typische Diskretisierung:

- Zeit: Wähle $K \in \mathbb{N},\ \Delta t := \frac{T}{K},\ t^k := k \cdot \Delta t,\ k = 0, \dots, K.$
- Ort: $X_h \subset X$endlich dimensional$(= H_0^1(\Omega))$, z.B. FE, FV, FD, Wavelets, Fourier
- suche dann eine Folge $\mathcal{U} := (u^k)_{k=0}^K \in X_h^{K+1}$ mit $u^k \approx u(t_k, \cdot),\ k = 0, \dots, K.$

Mögliche Parameterabhängigkeiten:

- $A = A(\mu)$
- $g = g(\mu)$
- $u_0 = u_0(\mu)$
- Diskretisierung

Definition 12.1:
Sei $X^{\mathcal{N}} \subset X,\ X$ sei ein Hilbert-Raum, $\mu \in \mathcal{D} \subset \mathbb{R}^P,\ \Delta t$ wie oben sowie $u_0 \in Y$ (nicht

unbedingt in X, z.B. H_0^1),

$$\mathcal{L}_{\Delta t}^I(t^k;\mu),\ \mathcal{L}_{\Delta t}^E(t^k;\mu) \in \mathcal{L}(X,X)$$
$$P_X : Y \to X \text{ eine beliebige stetige Projektion.}$$

Dann heißt das Problem: Suche $\mathcal{U}(\mu) := (u^k(\mu))_{k=0}^K \in (X^{\mathcal{N}})^{K+1}$ mit

$$\text{(12.2)} \qquad \begin{cases} \mathcal{L}_{\Delta t}(t^k;\mu)u^{k+1}(\mu) & = \mathcal{L}_{\Delta t}^E(t^k;\mu)\, u^k(\mu) + b_{\Delta t}(t^k;\mu),\ k = 0,\dots,K-1 \\ u_0(\mu) & = P_x(u_0) \end{cases}$$

das parabolische Problem (truth) mit

$$\begin{aligned} \mathcal{L}_{\Delta t}^I &: \text{ implizieter Anteil,} \\ \mathcal{L}_{\Delta t}^E &: \text{ explizieter Anteil} \\ b_{\Delta t} &: \text{ Inhomogenität.} \end{aligned}$$

Beispiel 12.2:
(a) Impliziter Euler in der Zeit: $u_t(t^{k+1}) \approx \frac{u(t^{k+1})-u(t^k)}{\Delta t}$, also

$$u(t^{k+1}) = \Delta t \cdot g(t^{k+1}) - \Delta t \cdot Au(t^{k+1}) + u(t^k)$$

in Variationsformulierung für alle $v \in X$

$$\left(u(t^{k+1}),v\right)_0 + \Delta t\, a\left(u(t^{k+1}),v\right) = \Delta t \cdot \left(g(t^{k+1}),v\right)_0 + \left(u(t^k),v\right)_0.$$

In diskreter Form für $X^{\mathcal{N}} = \operatorname{span}\left\{\varphi_1^{\mathcal{N}},\dots,\varphi_{\mathcal{N}}^{\mathcal{N}}\right\}$

$$\underline{M}^{\mathcal{N}}\underline{u}^{k+1} + \Delta t\, A^{\mathcal{N}}\underline{u}^{k+1} = \Delta t \cdot \underline{g}^{k+1} + \underline{M}^{\mathcal{N}}\underline{u}^k$$

mit der Massematrix $\underline{M}^{\mathcal{N}} := \left[(\varphi_i^{\mathcal{N}},\varphi_j^{\mathcal{N}})_0\right]_{i,j=1,\dots,\mathcal{N}}$ und der Steifigkeitsmatrix $\underline{A}^{\mathcal{N}} = \left[a(\varphi_i^{\mathcal{N}},\varphi_j^{\mathcal{N}})\right]_{i,j=1,\dots,\mathcal{N}}$ sowie der rechten Seite $\underline{g}^{k+1} = \left((g(t^{k+1}),\varphi_j^{\mathcal{N}})_0\right)_{j=1,\dots,\mathcal{N}}$, also

$$\mathcal{L}_{\Delta t}^I \cong \underline{M}^{\mathcal{N}} + \Delta t\, \underline{A}^{\mathcal{N}}, \qquad \mathcal{L}_{\Delta t}^E \cong \underline{M}^{\mathcal{N}}, \qquad b_{\Delta t} \cong \Delta t \cdot \underline{g}^{k+1}.$$

(b) Crank-Nicolson:

$$\frac{u(t^{k+1})-u(t^k)}{\Delta t} + \frac{1}{2}(A\,u(t^{k+1}) + A\,u(t^k)) = \frac{1}{2}(g(t^{k+1}) + g(t^k)).$$

Das RB-Evolutionsproblem ergibt sich durch $X_N \subset X^{\mathcal{N}}$, $\dim X_N = N << \mathcal{N}$ und der orthogonalen Projektion $P_N : X \to X_N$ bzgl. $(\cdot,\cdot)_X$ und

$$\mathcal{L}_{\Delta t,N}^{I,E} := P_n \circ \mathcal{L}_{\Delta t}^{I,E}, \qquad b_{\Delta t,N} := P_N \circ b_{\Delta t}.$$

Suche $\mathcal{U}_N(\mu) := \left(u_N^k(\mu)\right)_{k=0}^K \in X_N^{K+1}$ mit

$$\text{(12.3)} \qquad \begin{cases} \mathcal{L}_{\Delta t,N}^I(t^k;\mu)\, u_N^{k+1}(\mu) = \mathcal{L}_{\Delta t,N}^E(t^k;\mu)\, u_N^k(\mu) + b_{\Delta t,N}(t^k;\mu) \\ u_N^0 := P_N(P_X(u_0)). \end{cases}$$

Annahme 12.3:
(a) $\mathcal{L}^I_{\Delta t}$ *sei koerziv, d.h.* $\exists\, \alpha^I_{\Delta t} \in \mathbb{R}^+$ *mit* $(\mathcal{L}^I_{\Delta t}(t^k\,;\mu)v\,, v)_X \geq \alpha^I_{\Delta t}\,\|v\|^2_X$.
(b) $\mathcal{L}^I_{\Delta t},\ \mathcal{L}^E_{\Delta t}$ *seien stetig mit Stetigkeitskonstanten* $\gamma^I_{\Delta t},\ \gamma^E_{\Delta t} \in \mathbb{R}^+$.
(c) $b_{\Delta t}$ *sei beschränkt:* $\|b_{\Delta t}\|_X \leq \gamma^b_{\Delta t}$.

Satz 12.4:
Unter Annahme 12.3 sind die Probleme (12.2) und (12.3) wohlgestellt und es gilt (die Abhängigkeit von μ *wird weggelassen)*

$$\|u^k\|_X,\ \|u^k_N\|_X \leq \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^k \|u_0\|_X + \frac{\gamma^b_{\Delta t}}{\alpha^I_{\Delta t}} \sum_{i=0}^{k-1} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^i . \tag{12.4}$$

Beweis. Jede Iteration k besitzt nach dem Satz von Lax-Milgram wegen Annahme 12.3 eine eindeutige Lösung $u^k,\ u^k_N$, und es gilt

$$\|u^k\|_X \leq \frac{1}{\alpha^I_{\Delta t}} \left(\gamma^E_{\Delta t}\,\|u^{k-1}\|_X + \gamma^b_{\Delta t}\right). \tag{12.5}$$

Beweis von (12.4) per Induktion über k:

$$\text{(IA)} \qquad \|u_o\|_X = \underbrace{\left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^0}_{=1} \|u_0\|_X + \underbrace{\frac{\gamma^b_{\Delta t}}{\alpha^I_{\Delta t}} \sum_{i=0}^{-1} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^i}_{=0}$$

$$\begin{aligned}
\text{(IS)} \qquad \|u^{k+1}\|_X &\overset{(12.5)}{\leq} \frac{1}{\alpha^I_{\Delta t}} \left(\gamma^E_{\Delta t}\,\|u^k\|_X + \gamma^b_{\Delta t}\right) \\
&\overset{\text{IV}}{\leq} \frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \left\{ \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^k \|u_0\|_X + \frac{\gamma^b_{\Delta t}}{\alpha^I_{\Delta t}} \sum_{i=0}^{k-1} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^i \right\} + \frac{\gamma^b_{\Delta t}}{\alpha^I_{\Delta t}} \\
&= \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^{k+1} \|u_0\|_X + \frac{\gamma^b_{\Delta t}}{\alpha^I_{\Delta t}} \sum_{i=0}^{k-1} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^i ,
\end{aligned}$$

also (12.4) für u^k. Wegen $\|P_N v\|_X \leq \underbrace{\|P_N\|}_{=1} \cdot \|v\|_X$ folgt auch (12.4) für u^k_N. □

Bemerkung 12.5:
Falls $\gamma^E_{\Delta t} \geq \alpha^I_{\Delta t}$ *wächst die Norm der Lösung und ist für* $k \to \infty$ *unbeschränkt. Für* $\gamma^E_{\Delta t} < \alpha^I_{\Delta t}$ *ist die Lösung für alle Zeiten beschränkt.*

Beispiel 12.6:
Betrachte die Euler-Diskretisierung und bezeichne mit $\lambda^M_{\min},\ \lambda^M_{\max}$ *den kleinsten bzw. größten Eigenwert von* $\underline{M}^{\mathcal{N}}$. *Dann gilt* $\gamma^E_{\Delta t} = \lambda^M_{\max}$ *sowie* $\alpha^I_{\Delta t} = \lambda^M_{\min} + \Delta t\,\alpha_A$, *also erhalten wir*

$$\gamma^E_{\Delta t} < \alpha^I_{\Delta t} \qquad \textit{für } \Delta t > \frac{\lambda^M_{\max} - \lambda^M_{\min}}{\alpha_A}.$$

Ist $\Phi^{\mathcal{N}}$ *eine Orthonormalbasis von* X, *so ist* $\lambda^M_{\min} = \lambda^M_{\max} = 1$ *und die obige Bedingung ist immer erfüllt.*

Korollar 12.7:
Zusätzlich zu den Voraussetzungen von Satz 12.4 sei $\gamma^E_{\Delta t} \le 1$ *und* $\alpha^I_{\Delta t} = 1+\alpha\,\Delta t$, $\gamma^b_{\Delta t} = c\cdot\Delta t$. *Dann gilt*

$$\lim_{k\to\infty} \|u^k\|_X, \qquad \lim_{k\to\infty} \|u_N^k\|_X \le e^{-\alpha T}\,\|u_0\|_X + c\,T. \tag{12.6}$$

Beweis. Es gilt

- $\left(\frac{\gamma^b_{\Delta t}}{\alpha^I_{\Delta t}}\right)^K \le \left(\frac{1}{1+\Delta t\,\alpha}\right)^K = \left(\frac{1}{1+\alpha\,\frac{T}{K}}\right)^K = \Bigg(\underbrace{\left(\frac{1}{1+\alpha\,\frac{T}{K}}\right)^{\frac{K}{\alpha T}}}_{\overset{K\to\infty}{\longrightarrow} e^{-1}}\Bigg)^{\alpha T} \longrightarrow e^{-\alpha T}$
- $\sum_{i=0}^{K-1} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^i \le K$

und damit gilt für die rechte Seite in (12.4) also

$$\underbrace{\left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^k}_{\to e^{-\alpha\,T}} \|u_0\|_X + \frac{\gamma^b_{\Delta t}}{\alpha^I_{\Delta t}} \underbrace{\sum_{i=0}^{k-1} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}}\right)^i}_{\le K\cdot\frac{c\,\Delta t}{1} = c\,T}.$$

□

Bemerkung 12.8:
(a) Ist $\Phi^{\mathcal{N}}$ *eine Orthonormalbasis von* X *oder* $\|\cdot\|_X = \|\cdot\|_{L_2(\Omega)}$ *(schwächere Norm), dann können die Voraussetzungen erfüllt werden.*
(b) Ist $\|\cdot\|_X \equiv \|\cdot\|_{H^1(\Omega)}$ *und* $\Phi^{\mathcal{N}}$ *die normale FE-Basis, so ist* $\alpha << 1$ *sehr klein und die Schranke wächst über alle Grenzen!*

Beispiel 12.9 (FEM mit Crank-Nicolson)**:**
Wir betrachten den einfachen Fall $g|_\Gamma \equiv 0$ *und zeitunabhängige* a, g *mit* $a \in L_\infty(\Omega)$ *und*

$$a(t,x) = a(x) \ge a_0 > 0.$$

Es sei $X = X^{\mathcal{N}} = P_{1,0}(\mathcal{T}_h)$ *(lineare FE auf einem Gitter* $\mathcal{T}_h$ *mit homogenen Dirichlet-Randbedingungen). Sei*

$$(\cdot\,,\cdot)_X \equiv (\cdot\,,\cdot)_{H^1(\Omega)}, \qquad \|\cdot\|_X = \|\cdot\|_1 = \|\cdot\|_{H^1(\Omega)}.$$

Das Problem lautet dann: Suche $(u^k)_{k=0}^K \subset X^{K+1}$ *mit*

$$\begin{cases} u^0 = P_X(u_0) \\ \frac{1}{\Delta t}(u^{k+1}-u^k\,,v)_0 + \frac{1}{2}(a\nabla u^{k+1}\,,\nabla v)_0 + \frac{1}{2}(a\nabla u^k\,,\nabla v)_0 = (g\,,v)_0 \quad \forall v\in X, \end{cases} \tag{12.7}$$

also:

- $(\mathcal{L}^I_{\Delta t} u\,, v)_X := (u\,, v)_0 + \frac{\Delta t}{2}\,(a\nabla u\,, \nabla v)_0, \quad v \in X,$
- $(\mathcal{L}^E_{\Delta t} u\,, v)_X := (u\,, v)_0 - \frac{\Delta t}{2}\,(a\nabla u\,, \nabla v)_0, \quad v \in X,$
- $(b_{\Delta t}\,, v)_X := \Delta t (g\,, v)_0, \quad v \in X.$

Wir wollen die Annahmen 12.3 überprüfen:

Stetigkeit von $\mathcal{L}^E_{\Delta t}$:

Untersuche beide Teile getrennt:

$$\begin{aligned}(\mathcal{L}^E_1 u\,, v) &:= (u\,, v)_0,\\ (\mathcal{L}^E_2 u\,, v) &:= -\frac{\Delta t}{2}\,(a\nabla u\,, \nabla v)_0.\end{aligned}$$

Es gilt:

$$\begin{aligned}\|\mathcal{L}^E_1 u\|^2_X &= (\mathcal{L}^E_1 u\,, \mathcal{L}^E_1 u)_X = (u\,, \mathcal{L}^E_1 u)_0\\ &\overset{\substack{Cauchy-\\ Schwartz}}{\leq} \|u\|_0 \cdot \|\mathcal{L}^E_1 u\|_0 \leq \|u\|_0 \cdot \|\mathcal{L}^E_1 u\|_X\\ \Rightarrow \quad &\|\mathcal{L}^E_1 u\|_X \leq \|u\|_0.\end{aligned}$$

Als nächstes verwende Poincaré-Friedrichs:

$$\|v\|_0 \leq C\,\|\nabla v\|_0, \qquad \forall v \in H^1_0(\Omega),\ C = C(\Omega)$$

$$\begin{aligned}\Rightarrow \|\mathcal{L}^E_1 u\|^2_X &\leq \|u\|^2_0 = \alpha\,\|u\|^2_0 + (1-\alpha)\,\|u\|^2_0 \quad \textit{mit bel. } \alpha \in (0,1)\\ &\leq \alpha\,\|u\|^2_0 + C^2(1-\alpha)\|\nabla u\|^2_0.\end{aligned}$$

Wähle nun $\alpha \in (0,1)$ speziell so, dass $\alpha = C^2(1-\alpha)$, also $\alpha = \frac{C^2}{1+C^2} \in (0,1)$.

$$\Rightarrow \|\mathcal{L}^E_1 u\|^2_X \leq \frac{C^2}{1+C^2}\,(\underbrace{\|u\|^2_0 + \|\nabla u\|^2_0}_{= \|u\|^2_1 = \|u\|^2_X}),$$

d.h.

$$\|\mathcal{L}^E_1\|_{\mathcal{L}(X,X)} \leq \frac{C}{\sqrt{1+C^2}} =: \gamma^E_1 < 1.$$

Für den zweiten Teil gilt:

$$\begin{aligned}\|\mathcal{L}^E_2 u\|^2_X &= (\mathcal{L}^E_2 u\,, \mathcal{L}^E_2 u)_X = \frac{\Delta t}{2}\,(a\nabla u\,, \nabla(\mathcal{L}^E_2 u))_0\\ &\overset{\substack{C.-S.\\ a\in L_\infty(\Omega)}}{\leq} \frac{\Delta t}{2}\,\|a\|_\infty \cdot \|\nabla u\|_0 \cdot \|\nabla(\mathcal{L}^E_2 u)\|_0\\ &\leq \frac{\Delta t}{2}\,\|a\|_\infty \cdot \|u\|_X \cdot \|\mathcal{L}^E_2 u\|_X,\\ \Rightarrow \|\mathcal{L}^E_2\|_{\mathcal{L}(X,X)} &\leq \frac{\Delta t}{2}\,\|a\|_\infty =: \gamma^E_2.\end{aligned}$$

Also:

$$\|\mathcal{L}^E_{\Delta t}\|_{\mathcal{L}(X,X)} \le \|\mathcal{L}^E_1\|_{\mathcal{L}(X,X)} + \|\mathcal{L}^E_2\|_{\mathcal{L}(X,X)} \le \gamma^E_1 + \gamma^E_2 =: \gamma^E_{\Delta t}.$$

Wegen $\gamma^E_1 < 1$ gilt $\gamma^E_{\Delta t} < 1$ für Δt hinreichend klein. Offenbar können wir $\gamma^I_{\Delta t} = \gamma^E_{\Delta t}$ wählen.

Beschränktheit von $b_{\Delta t}$:

$$\begin{aligned} \|b_{\Delta t}\|^2_X &= (b_{\Delta t}\,, b_{\Delta t})_X = \Delta t\,(g\,, b_{\Delta t})_0 \overset{C.-S.}{\le} \Delta t \cdot \|g\|_0 \cdot \underbrace{\|b_{\Delta t}\|_0}_{\le \|b_{\Delta t}\|_X} \\ \Rightarrow \|b_{\Delta t}\|_X &\le \Delta t\,\|g\|_0 =: \gamma^b_{\Delta t}. \end{aligned}$$

Koerzivität von $\mathcal{L}^I_{\Delta t}$:

$$(\mathcal{L}^I_{\Delta t} u\,, u)_X \ge \|u\|^2_0 + \frac{\Delta t}{2}\, a_0 \|\nabla u\|^2_0 \ge \underbrace{\min\left\{1, \frac{\Delta t}{2}\, a_0\right\}}_{=\alpha_{\Delta t}} \underbrace{\|u\|^2_1}_{\|u\|^2_X}$$

Beachte:

$$\alpha_{\Delta t} \overset{\Delta t \to 0}{\longrightarrow} 0 \quad \lightning$$

Anderer Ansatz: $X^{\mathcal{N}}$ ist endlich-dimensional und $|\cdot|_1,\ \|\cdot\|_0,\ \|\cdot\|_1$ sind Normen auf $X^{\mathcal{N}}$

$$\Rightarrow \exists c_1,\ c_2 \text{ mit } \ \|u\|_0 \ge c_1 \|u\|_X, \qquad \|\nabla u\|_0 \ge c_2 \|u\|_X.$$

Aber:

$$c_1 = c_1(\mathcal{N}),\ c_2 = c_2(\mathcal{N}), \qquad c_1,\ c_2 \overset{\mathcal{N}\to\infty}{\longrightarrow} 0.$$

$$\Rightarrow (\mathcal{L}^I_{\Delta t} u\,, u)_X \ge \underbrace{(c_1^2 + \frac{\Delta t}{2}\, a_0\, c_2^2)}_{=: \alpha^I_{\Delta t}} \|u\|^2_X$$

und $\alpha^I_{\Delta t} \overset{\Delta t \to 0}{\longrightarrow} c_1^2 > 0$, aber bei realistischem $\mathcal{N}$ gilt $c_1^2 << 1$. Diese Abschätzung ist auch nicht zu verbessern.

Fazit:
Annamhe 12.3 ist erfüllt, aber die Vorraussetzung von Korollar 12.7 bzgl. $\mathcal{N}^I_{\Delta t}$ NICHT*! Wählt man $\|\cdot\|_X = \|\cdot\|_0$, dann gilt mit obiger Argumentation: $\alpha^I_{\Delta t} \ge 1 + \tilde{c}\,\Delta t$.*

Beispiel 12.10 (Konvektion-Diffusion mit finiten Volumen)**:**
Sei $\Omega = (0,1),\ T > 0$ und löse

$$\begin{cases} \partial_t u - \Delta u + \nabla \cdot (\beta \cdot u) = 0 \quad \text{auf } [0,T] \times \Omega, \\ u(0,\cdot) = 0 \qquad\qquad\qquad\quad \text{auf } \Omega,\ u_0 \in L_1(\Omega), \\ u(t,0) = u(t,1) = 0. \end{cases} \tag{12.8}$$

Sei $H \in \mathbb{N}$, $\Delta x := \frac{1}{H}$, $e_i := ((i-1)\Delta x, i \cdot \Delta x)$ *und definiere:*

$$\begin{aligned} x_i &:= i \cdot \Delta x + \frac{\Delta x}{2} = (i + \frac{1}{2})\,\Delta x = \frac{1}{2}(i \cdot \Delta x + (i+1)\,\Delta x), \\ x_{i-\frac{1}{2}} &:= x_i - \frac{\Delta x}{2} = i \cdot \Delta x, \\ x_{i+\frac{1}{2}} &:= x_i + \frac{\Delta x}{2} = i \cdot \Delta x, \\ X^{\mathcal{N}} &:= \operatorname{span}\{\mathbb{1}_{e_i}\}_{i=1,\dots,H} \subset \underbrace{L_1(\Omega) \cup L_2(\Omega)}_{= L_2(\Omega)},\ \mathcal{N} \equiv H. \end{aligned}$$

$\|\cdot\|_X = \|\cdot\|_0$, *suche also* $u^k \in X^{\mathcal{N}}$ *mit* $u^k = \sum_{i=1}^{\mathcal{N}} u_i^k \underbrace{\mathbb{1}_{e_i}}_{=\varphi_i^{\mathcal{N}}}$ *bzw.* $\underline{u}^k := (u_i^k)_{i=1}^H \in \mathbb{R}^H$ *(Vektor der Freiheitsgrade). Ziel:*

$$\begin{aligned} u_i^k &\approx \frac{1}{|e_i|}\int_{e_i} u(t^k, x)\,dx \quad \textit{(Zellmittel)}, \\ u_i^0 &:= P_{L_2}(u_0|e_i) = \frac{1}{\Delta x}\int_{e_i} u_0(x)\,dx \quad (L_2 - \textit{Projektion}). \end{aligned}$$

<u>Beachte</u>*:*

Für ein Kontrollvolumen $R := (t^k, t^{k+1}) \times e_i$ *gilt:*

$$\begin{aligned} 0 &\overset{(12.8)}{=} \int_R \left[\partial_t\, u(t,x) - \Delta u(t,x) + \nabla \cdot (\beta(t,x)\, u(t,x))\right] dx\, dt \\ &= \int_{e_{i+1}} \int_{t^k}^{t^{k+1}} \begin{pmatrix} \partial_x \\ \partial_t \end{pmatrix} \cdot \begin{pmatrix} -\partial_x u(t,x) + \beta(t,x)\, u(t,x) \\ u(t,x) \end{pmatrix} dx\, dt \\ &= \begin{cases} \int_{e_{i+1}} u(t^{k+1}, x)\,dx - \int_{e_{i+1}} u(t^k, x)\,dx \\ + \int_{t^k}^{t^{k+1}} \left[-\partial_x u(t, x_{i+\frac{1}{2}}) + \beta(t, x_{i+\frac{1}{2}})\, u(t, x_{i+\frac{1}{2}})\right] dt \\ - \int_{t^k}^{t^{k+1}} \left[-\partial_x u(t, x_{i-\frac{1}{2}}) + \beta(t, x_{i-\frac{1}{2}})\, u(t, x_{i-\frac{1}{2}})\right] dt \end{cases} \qquad (12.9) \\ &= \textit{Massen-Differenz + Fluss über } x_{i+\frac{1}{2}} \textit{ - Fluss über } x_{i-\frac{1}{2}}. \end{aligned}$$

→ *finite Volumen sind „konservativ“ (Erhaltungseigenschaft). Ersetze nun die Flüsse durch „numerische Flüsse“, die nur von den Zellmitteln abhängen:*

$$g_{i+\frac{1}{2}}^k \approx \frac{1}{\Delta t}\int_{t^k}^{t^{k+1}} \left[-\partial_x u(t, x_{i+\frac{1}{2}}) + \beta(t, x_{i+\frac{1}{2}})\, u(t, x_{i+\frac{1}{2}})\right] dt$$

und analog $g_{i-\frac{1}{2}}^k$*. Aus* (12.9) *ergibt sich dann ein Verfahren:*

$$u_i^{k+1} = u_i^k - \frac{\Delta t}{\Delta x}\,(g_{i+\frac{1}{2}}^k - g_{i-\frac{1}{2}}^k). \qquad (12.10)$$

Die Definition der Flüsse führt zu einem speziellen Verfahren, z.B.:

$$g_{i+\frac{1}{2}}^k := \left(-\frac{1}{\Delta x}\right)\left(u_{i+1}^{k+1} - u_i^{k+1}\right) + \underbrace{\beta_{i+\frac{1}{2}}^k\, u_i^k}_{\text{„Upwind“}}. \qquad (12.11)$$

Mit ghost nodes („Geister-Knoten“) $u_0^k = u_{H+1}^k := 0$ kann man (12.11) auch für die Randzellen verwenden. Einsetzen und sortieren liefert für $\beta(t,x) \equiv \beta$

$$(12.12)\quad -\frac{\Delta t}{\Delta x^2}\,u_{i+1}^{k+1} + \left(1 + 2\,\frac{\Delta t}{\Delta x^2}\right) u_i^{k+1} - \frac{\Delta t}{\Delta x^2}\,u_{i-1}^{k+1} = \left(1 - \frac{\Delta t}{\Delta x}\,\beta\right) u_i^k + \frac{\Delta t}{\Delta x}\,\beta u_{i-1}^k.$$

Für $\Delta t \le \beta \Delta x$ (CFL) ist das Verfahren stabil.

$$\begin{aligned}
\Rightarrow (\mathcal{L}_{\Delta t}^I \cong)\ \underline{L}_{\Delta t}^I &= \operatorname{tridiag}\left(-\frac{\Delta t}{\Delta x^2}, 1 + 2\frac{\Delta t}{\Delta x^2}, -\frac{\Delta t}{\Delta x^2}\right) \in \mathbb{R}^{H\times H},\\
\underline{L}_{\Delta t}^E &= \operatorname{tridiag}\left(\frac{\Delta t}{\Delta x}\,\beta, 1 - \frac{\Delta t}{\Delta x}\,\beta, 0\right) \in \mathbb{R}^{H\times H},\\
\underline{b}_{\Delta t} &= 0 \quad \Rightarrow \gamma_{\Delta t}^b = 0.
\end{aligned}$$

Stetigkeit von $\mathcal{L}_{\Delta t}^E$:

Es gilt

$$\|u^k\|_X^2 = \|u^k\|_0^2 = (u^k, u^k)_0 = (\underline{u}^k)^T\, \underline{M}\, \underline{u}^k$$

mit der Massenmatrix

$$\begin{aligned}
M_{ij} &= (\varphi_i^{\mathcal{N}}, \varphi_j^{\mathcal{N}})_0 = \int_\Omega e_i(x)\, e_j(x)\, dx,\\
&\overset{e_i = \mathbb{1}_{e_i}}{=} \delta_{ij} \int_{e_i} dx = \delta_{ij} \cdot \Delta x\\
\Rightarrow \underline{M} &= \Delta x \cdot \underline{I}.
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \|\mathcal{L}_{\Delta t}^E\|^2 &= \sup_{\|u\|_X = 1} \|\mathcal{L}_{\Delta t}^E u\|_X^2 = \sup_{\|u\|_X = 1} \left(\mathcal{L}_{\Delta t}^E, \mathcal{L}_{\Delta t}^E\right)_X\\
&= \sup_{\substack{\underline{u}\in\mathbb{R}^H\\ \underline{u}^T \underline{M}\, \underline{u} = 1}} \underline{u}^T \left(\underline{L}_{\Delta t}^E\right)^T \underline{M}\, \underline{L}_{\Delta t}^E\, \underline{u},
\end{aligned}$$

und mit $\underline{v} := \underline{M}^{\frac{1}{2}}\, \underline{u}$, $\underline{M}^T = \underline{M}$ folgt

$$\begin{aligned}
\|\mathcal{L}_{\Delta t}^E\|^2 &= \sup_{\substack{\underline{v}\in\mathbb{R}^H\\ \underline{v}^T \underline{v} = 1}} \underline{v}^T \underbrace{\underline{M}^{-\frac{1}{2}} \left(\underline{L}_{\Delta t}^E\right)^T \underline{M}\, \underline{L}_{\Delta t}^E\, \underline{M}^{-\frac{1}{2}}}_{=:\,\underline{A}}\, \underline{v}\\
&= \lambda_{\max}(A).
\end{aligned}$$

Da $\underline{M} = \Delta x \cdot \underline{I}$ und $\underline{M}^{-\frac{1}{2}} = \frac{1}{\sqrt{\Delta x}}\, \underline{I}$ gilt mit $\nu := \left(1 - \frac{\Delta t}{\Delta x}\beta\right)$

$$\underline{A} = \left(\underline{L}_{\Delta t}^E\right)^T \underline{L}_{\Delta t}^E = \operatorname{tridiag}\left(\frac{\Delta t}{\Delta x}\,\beta\nu, \nu^2 + \frac{\Delta t^2}{\Delta x^2}\,\beta^2, \frac{\Delta t}{\Delta x}\,\beta\nu\right).$$

Mit der Notation $\mu := \frac{\Delta t}{\Delta x}$ liefert der Satz von Gerschgorin

- *Mittelpunkt:* $(1-\mu)^2 + \mu^2 = 1 - 2\mu + 2\mu^2 > 0$

- *Radius:* $\sum_{k\neq j} |\underline{A}_{kj}| = 2\mu\,(1-\mu) = 2\mu - 2\mu^2$

$\Rightarrow$ *Alle Eigenwerte sind positiv, falls*

$$2\mu - 2\mu^2 < 1 - 2\mu + 2\mu^2 \Leftrightarrow 4\mu^2 - 4\mu + 1 > 0 \Leftrightarrow \mu > \frac{1}{2}.$$

und für alle Eigenwerte gilt

$$|\lambda| \leq \underbrace{(1-\mu)^2 + \mu^2}_{\textit{Mittelpunkt}} + \underbrace{2\mu\,(1-\mu)}_{\textit{Radius}} = 1 \Rightarrow \lambda_{\max}(A) \leq 1$$

$\Rightarrow \gamma^E_{\Delta t} := 1.$

Stetigkeit von $\mathcal{L}^I_{\Delta t}$: analog

Koerzivität von $\mathcal{L}^I_{\Delta t}$:

$$\begin{aligned}\frac{\left(\mathcal{L}^I_{\Delta t} u\,, u\right)_X}{\|u\|^2_X} &= \frac{\underline{u}^T \left(\underline{L}^I_{\Delta t}\right)^T \underline{M}\,\underline{u}}{\underline{u}^T \underline{M}\,\underline{u}} \overset{\underline{M}=\Delta t\cdot I}{=} \frac{\underline{u}^T \left(\underline{L}^I_{\Delta t}\right)^T \underline{u}}{\underline{u}^T\,\underline{u}} \\ &\overset{\underline{L}^I_{\Delta t}\ \textit{symm.}}{\geq} \inf_{\substack{\underline{u}\in\mathbb{R}^+ \\ \underline{u}^T\underline{u}=1}} \frac{\underline{u}^T\,\underline{L}^I_{\Delta t}\,\underline{u}}{\underline{u}^T\,\underline{u}} = \lambda_{\min}\left(\underline{L}^I_{\Delta t}\right).\end{aligned}$$

*Für EW symmetrischer Tridiagonalmatrizen (*tridiag(b,a,b)*) gilt mit $\overline{\mu} := \frac{\Delta t}{\Delta x^2} > 0$:*

$$|\lambda| \leq |a| + 2|b| = 1 + 2\overline{\mu} + 2\overline{\mu} = 1 + 4\overline{\mu} =: \alpha^I_{\Delta t}.$$

Fazit: Annahme 12.3 und *die Voraussetzungen von Korollar 12.7 sind mit $\alpha = \frac{4}{\Delta x^2}$ und $c = 0$ erfüllt.*

$$\begin{aligned}\underline{L}^I_{\Delta t} &= \underline{I} + \overline{\mu}\underbrace{\text{tridiag}(-1,2,-1)}_{=:\underline{B}_H},\ \lambda_{\min}(\underline{B}_H) = 2 - 2\cos\frac{\pi}{H+1} =: \overline{\alpha} > 0 \\ &\rightarrow \lambda_{\min}\left(\underline{L}^I_{\Delta t}\right) = 1 + \Delta t\,\underbrace{\frac{\overline{\alpha}}{\Delta x^2}}_{=:\alpha}.\end{aligned}$$

Lemma 12.11:
Falls $\left\{u^k(\mu)\right\}_{k=0}^K \subset X_N$, so gilt $u^k_N(\mu) = u^k(\mu) \quad \forall k.$

Beweis. Per Induktion:
(IA) $u^0(\mu) \in X_N \Rightarrow u^0_N = P_N\,u^0 = u^0$
(IS) Sei $u^k_N = u^k$ (IV), dann folgt aus $\mathcal{L}^I_{\Delta t,N} u^{k+1}_N - \mathcal{L}^E_{\Delta t,N} u^k_N - b_{\Delta t,N} = 0$

$$
\begin{aligned}
0 &= P_N \left(\mathcal{L}^I_{\Delta t,N} u_N^{k+1} - \mathcal{L}^E_{\Delta t,\,N} u_N^k - b_{\Delta t,\,N} \right) \\
&\overset{\text{(IV)}}{=} \mathcal{L}^I_{\Delta t,\,N} P_N\, u_N^{k+1} - \underbrace{\left(\mathcal{L}^E_{\Delta t,\,N} u^k + b_{\Delta t,\,N} \right)}_{= \mathcal{L}^I_{\Delta t,\,N} u^{k+1}} \\
&= P_N\, \mathcal{L}^I_{\Delta t,\,N} (u_N^{k+1} - u^{k+1}), \\
\Rightarrow 0 &= \left(\mathcal{L}^I_{\Delta t,\,N} \underbrace{(u_N^{k+1} - u^{k+1})}_{\in X_N}, \underbrace{(u_N^{k+1} - u^{k+1})}_{\in X_N} \right)_X \geq \alpha^I_{\Delta t} \, \| u_N^{k+1} - u^{k+1} \|_X^2 \geq 0 \\
\Rightarrow u_N^{k+1} &= u^{k+1}.
\end{aligned}
$$

□

Satz 12.12:
Unter Annahme 12.3 gilt

$$
\| u^k(\mu) - u_N^k(\mu) \|_X \leq \Delta_N(\mu, t^k) := \sum_{i=1}^{k} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \right)^{k-i} \frac{\Delta t}{\alpha^I_{\Delta} t} \| r^i(\mu) \|_X + \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \right)^k \| e^0(\mu) \|_X
$$

mit dem Residuum

$$
r^i(\mu) := \frac{1}{\Delta t} \left\{ \mathcal{L}^E_{\Delta t}(t^{i-1}; \mu) u_N^{i-1}(\mu) + b_{\Delta t}(t^i; \mu) - \mathcal{L}^I_{\Delta t}(t^i; \mu) u_N^i(\mu) \right\} \in X \ (\text{nicht} X')
$$

und dem Anfangsfehler $e^0(\mu) := u^0(\mu) - u_N^0(\mu)$.

Beweis. Es gilt

$$
\begin{aligned}
\mathcal{L}^I_{\Delta t} u^{k+1} &= \mathcal{L}^E_{\Delta t} u^k + b_{\Delta t}, \\
\mathcal{L}^I_{\Delta t,\,N} u_N^{k+1} &= \mathcal{L}^E_{\Delta t,\,N} u_N^k + b_{\Delta t,\,N}.
\end{aligned}
$$

Subtraktion ergibt für $e^k := u^k - u_N^k$

$$
\begin{aligned}
\mathcal{L}^I_{\Delta t} e^{k+1} &= \mathcal{L}^I_{\Delta t} u^{k+1} - \mathcal{L}^I_{\Delta t} u_N^{k+1} \\
&= \mathcal{L}^E_{\Delta t} u^k + b_{\Delta t} + \Delta t\, r^{k+1} - \mathcal{L}^E_{\Delta t} u_N^k - b_{\Delta t} \\
&= \mathcal{L}^E_{\Delta t} e^k + \Delta t\, r^{k+1},
\end{aligned}
$$

also ein Evolutionsproblem mit dem Residuum als Inhomogenität. Aus Satz 12.4, (12.5) folgt dann

$$
\begin{aligned}
\| e^{k+1} \|_X &\leq \frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \| e^k \|_X + \frac{1}{\alpha^I_{\Delta t}} \Delta t \, \| r^{k+1} \|_X \\
&\leq \frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \| e^{k-1} \|_X + \frac{\Delta t}{\alpha^I_{\Delta t}} \| r^k \|_X \right) + \frac{\Delta t}{\alpha^I_{\Delta t}} \| r^{k+1} \|_X \\
&\leq \ldots \leq \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \right)^{k+1} \| e^0 \|_X + \frac{\Delta t}{\alpha^I_{\Delta t}} \sum_{i=1}^{k+1} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \right)^{k+1-i} \| r^i \|_X \\
&= \Delta_N(\mu, t^{k+1}).
\end{aligned}
$$

□

Bemerkung 12.13:
Einer der entscheidenden Vorteile des allgemeinen Rahmens (Haasdonk, Ohlberger) ist, dass man ganz ähnliche Argumente wie bei stationären Problemen verwenden kann.

Bemerkung 12.14:
Oft gilt $\gamma^E_{\Delta t} = 1$ *und* $\alpha^I_{\Delta t} \leq 1$. *Dann wächst der Fehler monoton in* k. *Der echte Fehler kann aber durchaus fallen, so dass der Fehlerschätzer für lange Zeiten oft unbrauchbar wird. Dies kann man in diesem Rahmen vermeiden, wenn man Verfahren so konstruiert, dass* $\alpha^I_{\Delta t} > 1$ *gilt.*

Bemerkung 12.15:
Falls $\mathcal{L}^I_{\Delta t}$, $\mathcal{L}^E_{\Delta t}$, $b_{\Delta t}$ *affin zerlegbar bzgl.* Parameter und Zeit *sind, kann man analog effiziente offline/online-Zerlegungen herleiten.*

13 Basis-Generierung für zeitabhängige Probleme

Situation:
Suche RB-Raum $X_N \subset X$, der für alle $\mu \in \mathcal{D}$ *und* alle Zeiten $t^k,\ k = 0, \dots, K$, „gut“ ist in dem Sinne, dass $\|u^k(\mu) - u_N^k(\mu)\|_X$ möglichst klein ist für alle $\mu \in \mathcal{D}$ und alle $k = 0, \dots, K$. Es muss also die *gesamte Trajektorie* mit einem X_N gut approximiert werden, vgl. Lemma 12.11, dort war vorausgesetzt, dass $\left\{u^k(\mu)\right\}_{k=0}^K \subset X_N$
$(\Rightarrow u_N^k = u^k)$.

1. Ansatz: Lagrange-RB-Räume

- wähle $\mathcal{D}_{\text{train}} \subset \mathcal{D}$ endlich
- berechne $\mathcal{U}^{\mathcal{N}}(\mu) := \left(u_k^{\mathcal{N}}(\mu)\right)_{k=0}^K \in \mathbb{R}^{\mathcal{N}\times(K+1)} \quad \forall \mu \in \mathcal{D}_{\text{train}}$
- bestimme linear unabhängige Mengen $\Phi_N \subseteq \left\{\mathcal{U}^{\mathcal{N}}(\mu) : \mu \in \mathcal{D}_{\text{train}}\right\}$ und setze $X_N := \operatorname{span}\{\Phi_N\}$

Nachteile	Vorteile
(1) Berechnung & Speicherung von $\lvert\mathcal{D}_{\text{train}}\rvert \cdot (K+1)$ snapshots der Dim. $\mathcal{N}$	(i) Einfache Umsetzbarkeit
(2) N ggf. sehr groß (max. $\lvert\mathcal{D}_{\text{train}}\rvert \cdot (K+1)$)	(ii) Einfaches Debugging (Lem. 12.11) bei der online-Wahl von $\mu \in \mathcal{D}_{\text{train}}$
(3) Evtl. starke lineare Abhängigkeiten in $\left\{\mathcal{U}^{\mathcal{N}}(\mu) : \mu \in \mathcal{D}_{\text{train}}\right\}$	
(4) Aufgrund Dimensionen muss $\mathcal{D}_{\text{train}}$ klein sein $\Rightarrow$ evtl. schlechte Approx.	

2. Ansatz: PoD

- berechne $\mathcal{U}^{\mathcal{N}}(\mu), \quad \mu \in \mathcal{D}_{\text{train}}$ wie oben
- wähle $N \leq \lvert\mathcal{D}_{\text{train}}\rvert \cdot (K+1)$ und bestimme Φ_N per PoD

$$X_N = \arg \min_{\substack{Y \subset X \\ \dim(Y)=N}} \sum_{\mu \in \mathcal{D}_{\text{train}}} \sum_{k=0}^{K} \frac{1}{\lvert\mathcal{D}_{\text{train}}\rvert \cdot (K+1)} \|u_k^{\mathcal{N}}(\mu) - P_Y u_k^{\mathcal{N}}(\mu)\|_X^2$$

mit der orthogonalen Projektion $P_Y : X \to Y$ und einer ONB Φ_N

Nachteile	Vorteile
(1) wie oben (4) wie oben (großer offline-Aufwand) (5) Optimalität bzgl. μ nur im Mittel, nicht im worst case, kann für einzelne $\mu \in \mathcal{D}$ schlecht sein	(i) Problem (2) gelöst (ii) Problem (3) gelöst (iii) einfache Umsetzbarkeit

3. Ansatz: PoD-Greedy (Haasdonk, Ohlberger, 2008)
Ziel: betrachte

$$\arg \min_{\substack{Y \subset X \\ \dim(Y)=N}} \sup_{\mu \in \mathcal{D}_{\text{train}}} \frac{1}{K+1} \sum_{k=0}^{K} \|u_k^{\mathcal{N}}(\mu) - P_Y u_k^{\mathcal{N}}(\mu)\|_X^2$$

als guten Kandidaten für X_N, also L_∞-Parameter/L_2-Zeit $\to$ Nachteil (5) wäre gelöst. Wir brauchen dazu einen geeigneten Fehlerschätzer $\Delta(\mu, Y) \geq 0 \; \forall Y \subseteq X$.

Algorithmus 13.1 (PoD-Greedy)**:**

Wähle $\mathcal{D}_{\text{test}} \subset \mathcal{D}$ *endlich,* $\varepsilon_{\text{tol}} > 0$, $N_0 \in \mathbb{N}$, $\Phi_{N_0} \subset X$ ONB, $N_0 = |\Phi_{N_0}|$
Setze $N := N_0$, $X_N := \operatorname{span}(\Phi_N)$
while $\varepsilon_N := \max_{\mu \in \mathcal{D}_{\text{train}}} \Delta(\mu, X_N) > \varepsilon_{\text{tol}}$ **do**
  $\mu_{N+1} := \arg\max_{\mu \in \mathcal{D}_{\text{train}}} \Delta(\mu, X_N)$
  berechne $\mathcal{U}^{\mathcal{N}}(\mu_{N+1})$, $E_N(\mu_{N+1}) = \left(e_N^k(\mu_{N+1})\right)_{k=0}^K$
  mit $e_N^k(\mu) := u_k^{\mathcal{N}}(\mu) - P_{X_N} u_k^{\mathcal{N}}(\mu)$
  $\varphi_{N+1} := \arg\max_{\substack{\varphi \in X^{\mathcal{N}} \\ \|\varphi\|_X = 1}} \sum_{k=0}^{K} \|e_N^k(\mu_{N+1}) - (\varphi, e_N^k(\mu_{N+1}))_X \varphi\|_X^2$
  $\Phi_{N+1} := \Phi_N \cup \{\varphi_{N+1}\}$, $X_{N+1} := \operatorname{span}(\Phi_{N+1})$
end

Bemerkung 13.2:

(a) φ_{N+1} *ist die PoD-Mode der Trajektorie des Projektions*fehlers*.*

(b) $\mu \in \mathcal{D}_{\text{train}}$ *kann mehrfach ausgewählt werden. Wegen unterschiedlicher* N *führt dies aber im Allgemeinen zu unterschiedlichen* φ_N*.*

(c) Man kann mit $\Phi_N = \emptyset$ *starten, sinnvoller ist aber folgendes Vorgehen: Seien* $u^{0,q} \in X$, $q = 1, \ldots, Q_0$ *mit*

$$u^0(\mu) = \sum_{q=1}^{Q_0} \theta_0^q(\mu)\, u^{0,q}.$$

Wähle $X_{N_0} := \operatorname{span}\{u^{0,q}, q = 1, \ldots, Q_0\}$, $N_0 = \dim X_{N_0}$*. Dann gilt* $u^0(\mu) \in X_{N_0}$ *und damit* $\|u^0(\mu) - u_N^0(\mu)\|_X = 0 \; \forall \mu \in \mathcal{D}$*, die Anfangsbedingung wird also immer exakt repräsentiert.*

(d) Die PoD in (6) hat die Dimension $K+1$*, nicht* $|\mathcal{D}_{\text{train}}| \cdot (K+1) \Rightarrow$ *Problem (2) gelöst, Problem* (4) *entschärft.*

Lemma 13.3:
Φ_N ist eine ONB für X_N.

Beweis. Induktiv bezüglich N:

$(IA):$ $N = N_0$ klar

$(IS):$ $\|\varphi_{N+1}\|_X = 1$ nach Konstruktion und für $v_N \in X_N$ gilt:
$\left(e_N^k(\mu_{N+1}), v_N\right)_X = \left(u_k^{\mathcal{N}}(\mu_{N+1}, v_N\right)_X - \left(\cancel{\mathrm{P}_{X_N}} u_k^{\mathcal{N}}(\mu_{N+1}), v_N\right)_X = 0$
$\Rightarrow e_N^k(\mu_{N+1}) \in X_N^\perp$
Weiter:
$\varphi_{N+1} \underset{(6)}{\overset{(IV)}{\in}} \operatorname{span}\left\{e_N^k(\mu_{N+1})\right\}_{k=0}^K \subset X_N^\perp = (\operatorname{span}\{\Phi_N\})^\perp$
$\Rightarrow \varphi_{N+1} \perp \varphi_n,\ n = 1, \ldots, N \ \Rightarrow \Phi_{N+1}$ ist ONB.

□

Bemerkung 13.4:
Durch Verwendung der PoD ist Problem (3) gelöst.

Definition 13.5:
a) Wir bezeichnen mit

$$\langle U, V\rangle_{X^{K+1}} := \sum_{k=0}^{K} \left\langle u^k, v^k\right\rangle_X, \qquad \|U\|_{X^{K+1}} := \sqrt{\langle U, V\rangle_{X^{K+1}}}$$

ein Skalarprodukt und Norm auf $X^{K+1} = (X)^{K+1}$ mit $U = (u^k)_{k=0,\ldots,K}$, $V = (v^k)_{k=0,\ldots,K} \in X^{K+1}$.

b) Der (der POD zugrunde liegende) Projektionsfehler lautet:

$$E(\mu; Y) := \|U(\mu) - P_{Y^{K+1}} U(\mu)\|_{X^{K+1}}.$$

Definition 13.6:
a) Gilt für $\Delta(\mu; Y)$ in Algorithmus 13.1, dass

$$E(\mu_{n+1}; X_n) = \max_{\mu \in \mathcal{D}_{\text{train}}} E(\mu; X_n) \tag{13.1}$$

so heißt der Algorithmus 13.1 strong POD-Greedy-Verfahren.

b) Existiert ein $\gamma \in (0, 1)$ (unabhängig von $\mathcal{D}_{\text{train}}$) mit

$$E(\mu_{n+1}; X_n) \geq \gamma \max_{\mu \in \mathcal{D}_{\text{train}}} E(\mu; X_n) \tag{13.2}$$

so heißt der Algorithmus 13.1 weak POD-Greedy-Verfahren.

Proposition 13.7:
$\Delta(\mu; Y) := E(\mu; Y)$ erfüllt (13.1).

Beweis.

$$E(\mu_{N+1}; X_N) = \Delta(\mu_{N+1}; X_N) = \max_{\mu \in \mathcal{D}_{\text{train}}} \Delta(\mu; X_N) = \max_{\mu \in \mathcal{D}_{\text{train}}} E(\mu; X_N).$$

□

Ein Beispiel für ein weak POD-Greedy lernen wir im kommenden Paragraphen kennen. Die Güte des Fehlerschätzers ist ausschlaggebend für das gesamte Verfahren. Gibt es „instabile Parameter“ μ, so werden die Moden instabil.

Bemerkung 13.8:
In der Praxis verwendet man typischerweise den Fehlerschätzer $\Delta_N(\mu; t^k)$ aus Satz 12.12 für $\|u^k - u_N^k\|_X$ und setzt

$$\Delta_N(\mu; X_N) := \sum_{k=0}^{K} \Delta_N(\mu; t^k).$$

Bei Wahl von X_{N_0} gemäß Bemerkung 13.2(c) erhalten wir also

$$\Delta_N(\mu; X_N) = \sum_{k=0}^{K} \sum_{i=1}^{k} \left(\frac{\gamma^E_{\Delta t}}{\alpha^I_{\Delta t}} \right)^{k-i} \frac{\Delta t}{\alpha^I_{\Delta t}} \|r^i(\mu)\|_X, \tag{13.3}$$

wobei die X_N-Abhängigkeit im Residuum $r^i(\mu)$ gemäß Satz 12.12 enthalten ist. Offenbar wächst (13.3) *in der Zeit. Allgemein ist (wohl) offen, ob dies zu einem weak-Greedy-Verfahren führt (für Beispiel 12.10 kann man es zeigen), jedoch mit einem zeitlichen Verfahren, das klar suboptimal ist.*

Wir stellen nun eine Alternative vor: Seien für $X = Y$, $X_1, X_2 \subset X^{K+1}$ Hilbert-Räume mit Skalarprodukten $\langle \cdot, \cdot \rangle_{X_i}$, $i = 1, 2$ gemäß Definition 13.5. Definiere für $U \in X_1$ und $V \in X_2$

$$a(U, V; \mu) := \sum_{k=0}^{K-1} \left(\mathcal{L}^I_{\Delta t}(\mu; t^k) u^{k+1} - \mathcal{L}^E_{\Delta t}(\mu; t^k) u^k, v^{k+1} \right)_X + \left(u^0(\mu), v^0 \right)_X \tag{13.4}$$

$$f(V; \mu) := \sum_{k=0}^{K-1} \left(b_{\Delta t}(\mu; t^k), v^{k+1} \right)_X + \left(P_X u_0(\mu), v_0 \right). \tag{13.5}$$

Betrachte dann:

$$\text{Suche } U(\mu) \in X_1 : \quad a\left(U(\mu), V; \mu\right) = f(V; \mu) \qquad \forall V \in X_2. \tag{13.6}$$

Weiter seien $X_{1,N} \subset X_1, X_{2,N} \subset X_2$ endlich-dimensionale Räume sowie

$$\text{Suche } U_N(\mu) \in X_{1,N} : \quad a\left(U_N(\mu), V; \mu\right) = f(V; \mu) \qquad \forall V \in X_{2,N}. \tag{13.7}$$

Falls a, f stetig sind, ist (13.6) äquivalent zu (12.2) („wahres parabolisches Problem“) und (13.7) zu (12.3) (RB-Evolutions-Problem), jeweils für $X_1 = X_2$, $X_{1,N} = X_{2,N}$.

Satz 13.9 (Raum-Zeit-Evolutionsschema)**:**
Sei a inf-sup-stabil auf $X_1 \times X_2$ mit inf-sup-Konstante $\beta(\mu) > 0$. Dann gilt

$$\|U(\mu) - U_N(\mu)\|_{X_1} \leq \frac{\|\hat{R}_N\|_{X_2}}{\beta(\mu)} =: \Delta_N(\mu), \tag{13.8}$$

wobei $\hat{R}_N$ der Riesz-Repräsentant des Residuums ist, d.h.

$$\left\langle \hat{R}_N, V \right\rangle_{X_2} = f(V;\mu) - a\left(U_N, V;\mu\right), \qquad V \in X_2.$$

Für die Effektivität gilt

$$\frac{\Delta_N(\mu)}{\|U(\mu) - U_N(\mu)\|_{X_1}} \leq \frac{\gamma_a(\mu)}{\beta(\mu)} =: \eta(\mu), \tag{13.9}$$

mit der Stetigkeits-Konstanten $\gamma_a(\mu)$ von $a(\cdot,\cdot;\mu)$

Beweis. Ungleichung (13.8) folgt direkt aus Lemma 6.3 und den Eigenschaften von Riesz-Repräsentanten, da (13.6),(13.7) Petrov-Galerkin-Verfahren sind. Weiter gilt für $E_N := U - U_N$, $e_N := u - u_N$, $r_N(v) := f(v) - a(u_N, v),\ v \in Y$ und $a(u,v) = f(v)$

$$\begin{aligned}
\gamma_a \|e_N\| \|v\| &\geq a(e_N, v) \\
&= a(u,v) - a(u_N, v) \\
&= f(v) - a(u_N, v) \\
&= r_N(v), \qquad\qquad r_N \in Y'.
\end{aligned}$$

Weiter ist $\hat{r}_N \in Y$ mit $(\hat{r}_N, v)_Y = r_N(v)$ für alle $v \in Y$ und $\|\hat{r}_N\|_Y = \|r_N\|_{Y'}$. Daraus folgt $\gamma_a \|e_N\| \geq \sup_{v \in Y} \frac{r_N(v)}{\|v\|_Y} = \|r_N\|_{Y'} = \|\hat{r}_N\|_Y$ und folglich $\gamma_a(\mu)\|E_N\|_{X_1} \geq \|\hat{R}_N\|_{X_2}$. Weiter ist $\gamma_a(\mu)\|E_N\|_{X_1}\|V\|_{X_2} \geq a(E_N, V;\mu) = R_N(V)$, daraus folgt

$$\frac{\|\hat{R}_N\|_{X_2}}{\|E_N\|_{X_1}} \leq \gamma_a(\mu) \qquad \text{und} \qquad \frac{\Delta_N(\mu)}{\|E_N\|_{X_1}} = \frac{1}{\beta(\mu)} \frac{\|\hat{R}_N\|_{X_2}}{\|E_N\|_{X_1}} \leq \frac{\gamma_a(\mu)}{\beta(\mu)}$$

□

Bemerkung 13.10:
Der Ansatz in (13.6),(13.7) *koppelt Raum und Zeit, X_N enthält Raum-Zeit-Funktionen. Dies geht auf KU, Patera 2011-2012 zurück und wird im kommenden Abschnitt näher beschrieben.*

Proposition 13.11:
Der Fehlerschätzer $\Delta_N(\mu)$ aus (13.8) *führt zu einem weak POD-Greedy-Verfahren.*

Beweis. Nach Satz 13.9 gilt (sei $X_1 = X_2 = X$)

$$\|U(\mu) - U_N(\mu)\|_{X^{K+1}} \leq \Delta_N(\mu) \leq \eta(\mu)\|U(\mu) - U_N(\mu)\|_{X^{K+1}},$$

Weiterhin gilt:

$$\begin{aligned}\|U(\mu) - P_{X_N^{K+1}} U(\mu)\|_{X^{K+1}} &\leq \|U(\mu) - U_N(\mu)\|_{X^{K+1}} \\ &\leq \left(1 + \frac{\gamma_a(\mu)}{\beta(\mu)}\right) \|U(\mu) - P_{X_N^{K+1}} U(\mu)\|_{X^{K+1}}\end{aligned}$$

denn allgemein für $v \in X_N$

$$\begin{aligned}\beta(\mu)\|v_N - u_N\|^2 &\leq a(v_N - u_N, v_N - u_N; \mu) \\ &\leq a(v_N - u_N + (u_N - u), v_N - u_N; \mu) \\ &= a(v_N - u, v_N - u_N; \mu) \\ &\leq \gamma_a(\mu)\|v_N - u\|\|v_N - u_N\|\end{aligned}$$

folglich $\beta(\mu)\|v_N - u_N\| \leq \gamma_a(\mu)\|v_N - u\|$ und damit

$$\begin{aligned}\|U(\mu) - U_N(\mu)\|_{X^{K+1}} &\leq \|U(\mu) - P_{X_N^{K+1}} U(\mu)\|_{X^{K+1}} + \|P_{X_N^{K+1}} U(\mu) - U_N(\mu)\|_{X^{K+1}} \\ &\leq \left(1 + \frac{\gamma_a(\mu)}{\beta(\mu)}\right) \|U(\mu) - P_{X_N^{K+1}} U(\mu)\|_{X^{K+1}}.\end{aligned}$$

Daraus folgt:

$$\begin{aligned}E(\mu_{n+1}, X_n) &= \|U(\mu_{n+1}) - P_{X_N^{K+1}} U(\mu_{n+1})\|_{X^{K+1}} \\ &\geq \left(1 + \frac{\gamma_a(\mu_{n+1})}{\beta(\mu_{n+1})}\right)^{-1} \|U(\mu_{n+1}) - U_N(\mu_{n+1})\|_{X^{K+1}} \\ &\geq \frac{1}{\eta(\mu_{n+1})} \left(1 + \frac{\gamma_a(\mu_{n+1})}{\beta(\mu_{n+1})}\right)^{-1} \Delta_N(\mu) \\ &= \frac{1}{\eta(\mu_{n+1})} \left(1 + \frac{\gamma_a(\mu_{n+1})}{\beta(\mu_{n+1})}\right)^{-1} \max_{\mu \in \mathcal{D}_{\text{train}}} \Delta_N(\mu) \\ &\geq \gamma_{n+1} \max_{\mu \in \mathcal{D}_{\text{train}}} \|U(\mu_{n+1}) - P_{X_N^{K+1}} U(\mu_{n+1})\|_{X^{K+1}} \\ &= \gamma_{n+1} \max_{\mu \in \mathcal{D}_{\text{train}}} E(\mu, X_N)\end{aligned}$$

wobei die vorletzte Ungleichung aus Algorithmus 13.1 folgt. Weiter gilt die Abschätzung $\Delta_N(\mu) \geq \|U(\mu) - U_N(\mu)\|_{X^{K+1}}$ und

$$\gamma_{n+1} := \frac{1}{\eta(\mu_{n+1})} \left(1 + \frac{\gamma_a(\mu_{n+1})}{\beta(\mu_{n+1})}\right)^{-1} \in (0, 1),$$

da

$$\frac{1}{\eta(\mu)} < 1 \qquad \forall \mu \in \mathcal{D}_{\text{train}}.$$

Weiter ist $\gamma_a(\mu_{n+1}) \leq \gamma_a^{UB} < \infty$ und $\beta(\mu_{n+1}) \geq \beta^{LB} > 0$ und somit ist

$$1 + \frac{\gamma_a(\mu_{n+1})}{\beta(\mu_{n+1})} > 1$$

und folglich $\left(1 + \frac{\gamma_a(\mu_{n+1})}{\beta(\mu_{n+1})}\right)^{-1} < 1$. □

Bemerkung 13.12:

- *Strong POD-Greedy:*
  - immernoch offline teuer, snapshots werden gebraucht
  + allgemeines Verfahren zur Approximation von Lösungs-Sequenzen
  + Fehlersequenz fällt monoton
  → Verwendung falls keine Fehlerschätzer vorhanden!

- *Weak POD-Greedy:*
  + offline sehr effizient $\Rightarrow |\mathcal{D}_{\text{train}}| \nearrow$
  + alle obigen Probleme gelöst
  - Fehlersequenz muss nicht fallen
  ∘ abhängig von der Güte des Fehlerschätzers

14 Space-time-Diskretisierungen

Seien $V \hookrightarrow H \hookrightarrow V'$ Hilberträume, $I := (0,T]$, $T > 0$ und $A \in \mathcal{L}(V, V')$ definiert durch

$$\langle A\Phi, \Psi \rangle_{V' \times V} = a(\Phi, \Psi), \qquad a : V \times V \to \mathbb{R} \quad \text{BLF.}$$

Gegeben $g : I \times \Omega \to \mathbb{R}$, $u_0 \in H$, suche $u : I \times \Omega \to \mathbb{R}$ ($u(t) : \Omega \to \mathbb{R}\ \forall t \in I$) mit

$$u_t(t) + Au(t) = g(t) \quad \forall t \in I, \qquad u(0) = u_0 \quad \text{in } H, \tag{14.1}$$

vergleiche (12.1). Die erste Gleichung ist offenbar in V'. Um eine space-time Variationsformulierung von (14.1) zu formulieren, brauchen wir eine Verallgemeinerung des Lebesque-Integrals für Banach-Raum-wertige Funktionen, das Bochner-Integral.

Sei $\left(\tilde{\Omega}, \mathcal{A}, \mu\right)$ ein σ-endlicher Maßraum und $(B, \|\cdot\|_B)$ ein Banach-Raum sowie $s : \tilde{\Omega} \to B$ eine Treppenfunktion der Form

$$s(x) := \sum_{i=1}^{m} \alpha_i \chi_{\Omega_i}(x), \quad x \in \tilde{\Omega}, \qquad \Omega_i \in \mathcal{A} \text{ messbar}, \quad \alpha_i \in B,$$

dann setzt man

$$\int_\Omega s \, d\mu := \sum_{i=1}^{m} \alpha_i \mu(\Omega_i)$$

und dann weiter wie beim Lebesque-Integral. Wir verwenden diese Defintion für $\tilde{\Omega} = I$ und $B \in \{V', H, V\}$. Dann setze

$$L_2(I;B) := \left\{ f : I \to B \;\middle|\; \int_I \|f(t)\|_B^2 \, dt < \infty \right\}$$

mit der Norm $\|f\|_{L_2(I;B)}^2 = \int_I \|f(t)\|_B^2 \, dt$. Sei dann

$$\begin{aligned} X :=& \{v \in L_2(I;V) \mid \dot{v} \in L_2(I;V'),\ v(0) = 0\} \\ =& L_2(I;V) \cap H^1_{(0)}(I,V') \ \ (= W_0(I,V,V')) \\ Y :=& L_2(I;V) \end{aligned}$$

mit

$$H^1_{(0)}(I,V') := \left\{ v \in H^1(I;V') \mid v(0) = v_0 \right\}$$

und

$$H^1(I;V') := \{v \in L_2(I,V') \mid \dot{v} \in L_2(I;V')\}$$

und Normen

$$\|w\|_X^2 := \|w\|_{L_2(I;V)}^2 + \|\dot{w}\|_{L_2(I;V')}^2 + \|w(T)\|_H^2$$
$$\|v\|_Y := \|v\|_{L_2(I;V)}$$

Bemerkung 14.1:
a) Es gilt $X \hookrightarrow \mathcal{C}(I;H)$, Punktauswertung bezüglich t macht also Sinn.
b) Die stärkere Norm für X (mit $\|w(T)\|_H$) dient der Kontrolle am Endzeitpunkt.

Mit $[w,v]_{\mathcal{H}} := \int_I \langle w(t), v(t)\rangle_{V'\times V}\ dt$, $\quad \mathcal{A}[w,v] := \int_I a\,(w(t),v(t))\ dt$ setze

$$b(w,v) := [\dot{w},v]_{\mathcal{H}} + \mathcal{A}[w,v]\,, \qquad f(v) = [g,v]_{\mathcal{H}} \tag{14.2}$$

und die space-time-Variationsformulierung lautet dann

$$\text{Suche } u \in X \text{ mit} \qquad b(u,v) = f(v) \qquad \forall v \in Y. \tag{14.3}$$

Satz 14.2:
Angenommen, es gibt Konstanten $M_a < \infty, \alpha > 0$ und $\lambda \geq 0$ so dass

$$|a(\Phi,\Psi)| \leq M_a\|\Phi\|_V\|\Psi\|_V, \qquad \textit{(Beschränktheit)} \tag{14.4}$$
$$a(\Psi,\Psi) + \lambda\|\Psi\|_H^2 \geq \alpha\|\Psi\|_V^2, \qquad \textit{(Gårding-Ungleichung)} \tag{14.5}$$

dann ist (14.3) *korrekt gestellt.*

Beweis. Nach dem Satz von Babuška und Aziz haben wir folgende drei Bedingungen zu zeigen

$$M_b := \sup_{w\in X}\sup_{v\in Y}\frac{|b(w,v)|}{\|w\|_X\|v\|_Y} < \infty, \qquad \text{(Stetigkeit)} \tag{14.6}$$
$$\beta := \inf_{w\in X}\inf_{v\in Y}\frac{|b(w,v)|}{\|w\|_X\|v\|_Y} > 0, \qquad \text{(inf-sup)} \tag{14.7}$$
$$\forall 0 \neq v \in Y \qquad \sup_{w\in X}|b(w,v)| > 0, \qquad \text{(Surjektivität)} \tag{14.8}$$

Es gilt mit Hölder

$$\begin{aligned}
|b(w,v)| &\leq \|\dot{w}\|_{L_2(I;V')}\|v\|_{L_2(I;V)} + M_a\|w\|_{L_2(I;V)}\|v\|_{L_2(I;V)} \\
&\leq \underbrace{\max\{1,M_a\}\sqrt{2}}_{=:M_b}\ \sqrt{\|\dot{w}\|_{L_2(I;V')}^2 + \|w\|_{L_2(I;V)}^2}\ \|v\|_{L_2(I;V)} \\
&\leq M_b\|w\|_X\|v\|_Y,
\end{aligned}$$

also (14.6). Nun zu (14.7): Sei $0 \neq w \in X$ gegeben. Definiere $z_w(t) := (A^*)^{-1}\,\dot{w}(t)$, wobei $A^* : V \to V'$ die adjungierte Abbildung ist, d.h. $\langle A^*\Phi,\Psi\rangle_{V'\times V} = a(\Psi,\Phi)$. Daraus folgt

$$\begin{aligned}
\|A^*\Phi\|_{V'} &= \sup_{\Psi\in V}\frac{\langle A^*\Phi,\Psi\rangle}{\|\Psi\|_V} = \sup_{\Psi\in V}\frac{a(\Psi,\Phi)}{\|\Phi\|_V} \\
&\geq \frac{a(\Phi,\Phi)}{\|\Phi\|_V} \overset{(14.5)}{\geq} \frac{1}{\|\Phi\|_V}\left(\alpha\|\Phi\|_V^2 - \lambda\|\Phi\|_H^2\right) \\
&\geq \left(\alpha - \lambda\rho^2\right)\|\Phi\|_V,
\end{aligned}$$

wobei $\rho := \sup_{0\neq\Phi\in V} \frac{\|\Phi\|_H}{\|\Phi\|_V}$. Also gilt $\| (A^*)^{-1} \| \leq (\alpha - \lambda\rho^2)^{-1}$. Damit gilt für die Variable $v_w(t) := z_w(t) + w(t)$:

$$\begin{aligned} \|v_w\|_Y^2 &\leq 2\left(\|z_w\|_Y^2 + \|w\|_Y^2\right) \leq 2\left(\left(\alpha-\lambda\rho^2\right)^{-2}\|\dot{w}\|_Y^2 + \|w\|_Y^2\right) \\ &\leq 2\max\left\{1, \left(\alpha-\lambda\rho^2\right)^{-2}\right\}\|w\|_X^2. \end{aligned} \tag{14.9}$$

Weiter gilt:

$$b(w, v_w) = [\dot{w}, z_w]_{\mathcal{H}} + \mathcal{A}[w, z_w] + [\dot{w}, w]_{\mathcal{H}} + \mathcal{A}[w, w]$$

und es gelten die folgenden Gleichungen und Ungleichungen:

$$\begin{aligned} [\dot{w}, z_w]_{\mathcal{H}} &= \int_0^T \left\langle \dot{w}(t), (A^*)^{-1}\dot{w}(t)\right\rangle_{V'\times V}\, dt \\ &= \int_0^T a\left(z_w(t), z_w(t)\right)\, dt \geq \left(\alpha-\lambda\rho^2\right)\|z_w\|_Y^2 \\ &\geq \left(\alpha-\lambda\rho^2\right) M_a^{-2}\|\dot{w}\|_{L_2(I;V')}^2, \end{aligned}$$

denn $\|z_w\|_V = \| (A^*)^{-1}\dot{w}(t)\|_V \geq \frac{1}{M_a}\|\dot{w}(t)\|_{V'}$.

$$\begin{aligned} \mathcal{A}[w, z_w] &= \int_0^T a\left(w(t), (A^*)^{-1}\dot{w}(t)\right)\, dt \\ &= \int_0^T \langle w(t), \dot{w}(t)\rangle_{V\times V'}\, dt = \frac{1}{2}\int_0^T \frac{d}{dt}\|w(t)\|_H^2\, dt \\ &= \frac{1}{2}(\|w(T)\|_H^2 - \underbrace{\|w(0)\|_H^2}_{=0}) = \frac{1}{2}\|w(T)\|_H^2. \end{aligned}$$

$$[\dot{w}, w]_{\mathcal{H}} = \int_0^T \langle \dot{w}(t), w(t)\rangle_{V'\times V}\, dt \overset{s.o.}{=} \frac{1}{2}\|w(T)\|_H^2.$$

$$\mathcal{A}[w, w] = \int_0^T a\left(w(t), w(t)\right)\, dt \geq \left(\alpha-\lambda\rho^2\right)\|w\|_Y^2.$$

Insgesamt:

$$\begin{aligned} b(w, v_w) &\geq \left(\alpha-\lambda\rho^2\right) M_a^{-2}\|\dot{w}\|_{L_2(I;V')}^2 + \|w(T)\|_H^2 + \left(\alpha-\lambda\rho^2\right)\|w\|_Y^2 \\ &\geq \min\left\{\min\left\{M_a^{-2}, 1\right\}\left(\alpha-\lambda\rho^2\right), 1\right\}\|w\|_X^2 \\ &\geq \underbrace{\frac{\min\left\{\min\left\{M_a^{-2}, 1\right\}\left(\alpha-\lambda\rho^2\right), 1\right\}}{\sqrt{2}\max\left\{1, \left(\alpha-\lambda\rho^2\right)^{-2}\right\}}}_{=:\beta^{\mathrm{LB}}>0}\|v_w\|_Y\|w\|_X \end{aligned} \tag{14.10}$$

also $\beta \geq \beta^{\mathrm{LB}} > 0$ und damit (14.7). Schließlich wird (14.8) durch eine endlichdimensionale Galerkin-Approximation gezeigt, vgl. [C. Schwab, R. Stirenson, Adaptive Wavelet Methods for Parabolic Problems, Math. Comp. 78, No. 267, 1293-1318 (2009)], S. 1315-1316. □

Bemerkung 14.3:
a) Offenbar liefert Ungleichung (14.10) *sogar eine berechenbare untere Schranke für* β. *Zwar ist diese Schranke in der Praxis oft zu ungenau, es zeigt aber das Potenzial von Space-Time-Diskretisierungen.*
b) Obiger Beweis funktioniert auch für LTV-Operatoren.
c) Für die Graphennorm auf X *erhält man eine ähnliche Aussage.*

Satz 14.4:
Für $A = -\Delta, V = H_0^1(\Omega)$, $H = L_2(\Omega)$, $\|\Phi\|_V^2 = a(\Phi,\Phi) = \|\nabla v\|_{L_2(\Omega)}^2$ *gilt:*

$$\beta = \gamma := \sup_{w\in X}\sup_{v\in Y}\frac{b(w,v)}{\|w\|_X\|v\|_Y} = 1.$$

Beweis. (1) Wir zeigen zunächst $\beta \geq 1$. Gegeben sei dazu $0 \neq w \in X$, setze wie oben $z_w := A^{-1}\dot{w}$, $v_w := z_w + w$. Wir erhalten $\|v_w\|_{L_2(I;V)}^2 = \|z_w\|_{L_2(I;V)}^2 + \|w\|_{L_2(I;V)}^2 + 2\int_I (z_w(t), w(t))_V \; dt$ und

$$\begin{aligned}(z_w(t), w(t))_V &= a\left(z_w(t), w(t)\right) = a\left(A^{-1}\dot{w}(t), w(t)\right)\\ &= \left\langle A^{-1}\dot{w}(t), w(t)\right\rangle_{V'\times V} \overset{s.o.}{=} \frac{1}{2}\frac{d}{dt}\|w(t)\|_H^2\\ \overset{(*)}{\Rightarrow} \|v_w\|_{L_2(I;V)}^2 &= \|\dot{w}\|_{L_2(I;V')}^2 + \|w(T)\|_H^2 = \|w\|_X^2\end{aligned}$$

und damit wie im Beweis zu Satz 14.2

$$b(w, v_w) \geq \|w\|_X^2 = \|w\|_X\|v_w\|_Y.$$

(2) Für $w \in X$ und $v \in Y$ gilt $b(w,v) = \int_I a\left(A^{-1}\dot{w}(t) + w(t), v(t)\right) \; dt$. Gegeben sei $v \in Y \Rightarrow Av \in L_2(I;V') = Y'$, also gibt es nach (1) genau ein $z \in X$ mit $\dot{z} + Az = Av$, $z(0) = 0$, d.h. $v = A^{-1}\dot{z} + z$

$$\begin{aligned}\Rightarrow \sup_{v\in Y}\frac{b(w,v)}{\|v\|_Y} &= \sup_{z\in X}\frac{b(w, A^{-1}\dot{z} + z)}{\|A^{-1}\dot{z} + z\|}\\ &= \sup_{z\in X}\frac{\int_I a\left(A^{-1}\dot{w}(t) + w(t), A^{-1}\dot{z}(t) + z(t)\right) \; dt}{\|A^{-1}\dot{z} + z\|_Y}\\ &= \|A^{-1}\dot{w} + w\|_Y \overset{(*)}{=} \|w\|_X,\end{aligned}$$

da $v_w = A^{-1}\dot{w} + w$. [K. Urban, A.T. Patera, A new error bound for Reduced Basis Approximation of Parabolic PDEs, C.R. Acad Sci I, 350 (3-4), 203-207 (2012)]. □

Bemerkung 14.5:
Oft transformiert man (14.1) *für analytische Zwecke:*

$$\begin{aligned}&\hat{u}(t) := e^{-\lambda t}u(t) \qquad \hat{v}(t) := e^{\lambda t}v(t) \qquad \hat{g}(t) := e^{-\lambda t}g(t)\\ &\Rightarrow \hat{b}(\hat{w}, \hat{v}) = \hat{f}(\hat{v}), \qquad\qquad \forall \hat{v} \in Y,\end{aligned}$$

mit

$$\hat{b}(\hat{w},\hat{v}) := \int_0^T \left\langle \frac{d}{dt}\hat{w}(t), \hat{w}(t) \right\rangle_{V'\times V} dt + \int_0^T \hat{a}\left(\hat{w}(t),\hat{v}(t)\right)\ dt$$

$$\hat{a}\left(\hat{w}(t),\hat{v}(t)\right) := a\left(\hat{w}(t),\hat{v}(t)\right) + \lambda\left(\hat{w}(t),\hat{v}(t)\right)_H$$

$$\hat{f}(\hat{v}) := \int_0^T \langle \hat{g}(t),\hat{v}(t)\rangle_{V'\times V}\ dt$$

$\Rightarrow \hat{a}$ *erfüllt Gårding* (14.5) *mit* $\lambda = 0$.

$$\Rightarrow \beta \geq \hat{\beta}^{LB} := \frac{e^{-2\lambda t}}{\max\left\{\sqrt{1+2\lambda^2\rho^4}, \sqrt{2}\right\}} \cdot \frac{\min\left\{1, \alpha \min\left\{1, M_a^{-2}\right\}\right\}}{\sqrt{2}\max\left\{1, \left(\beta_a^*\right)^{-1}\right\}}$$

[UP], Cor 2.7.

Nun zur (truth) Diskretisierung für $H = L_2(\Omega), V = H_0^1(\Omega)$.

$$\delta := (\Delta t, h)$$

$$x_\delta := S_{\Delta t} \otimes V_h, \qquad S_{\Delta t} := span\left\{\sigma^1,\ldots,\sigma^K\right\},\ \Delta t = \frac{T}{K}$$

$$y_\delta := Q_{\Delta t} \otimes V_h, \qquad Q_{\Delta t} := span\left\{\tau^1,\ldots,\tau^K\right\},\ t^k = k\cdot\Delta t$$

$$V_h := span\left\{\Phi_1,\ldots,\Phi_{n_h}\right\}, \qquad I^k := \left(t^{k-1}, t^k\right).$$

Dann gilt für $w_\delta = \sum_{k=1}^K \sum_{i=1}^{n_h} w_i^k \sigma^k \otimes \Phi_i \in X_\delta$, $v_\delta = \sum_{l=1}^K \sum_{j=1}^{n_h} v_j^l \tau^k \otimes \Phi_j \in Y_\delta$:

$$b(w_\delta, v_\delta) = \sum_{k,l=1}^{K}\sum_{i,j=1}^{n_h} w_i^k v_j^l \Big[\underbrace{\left(\dot{\sigma}^k, \tau^l\right)_{L_2(I)}}_{\delta_{k,l}-\delta_{k+1,l}} \langle\Phi_i,\Phi_j\rangle_H + \underbrace{\left(\sigma^k,\tau^l\right)_{L_2(I)}}_{\frac{\Delta t}{2}\left(\delta_{k,l}-\delta_{k+1,l}\right)} a\left(\Phi_i,\Phi_j\right)\Big]$$

$$\Rightarrow b(w_\delta, \tau^l\otimes\Phi_j) = \sum_{i=1}^{n_h}\left[\left(w_i^l - w_i^{l-1}\right)\langle\Phi_i,\Phi_j\rangle_H + \frac{\Delta t}{2}\left(w_i^l + w_i^{l-1}\right) a\left(\Phi_i,\Phi_j\right)\right]$$

$$= \Delta t\left[\underline{M}_h^{\text{space}} \frac{1}{\Delta t}\left(\underline{w}^l - \underline{w}^{l-1}\right) + \underline{A}_h^{\text{space}}\underline{w}^{l-\frac{1}{2}}\right]$$

mit

$$\underline{w}^l = \left(w_i^l\right)_{1\leq i\leq n_h} \qquad \underline{M}_h^{\text{space}} = \left[\langle\Phi_i,\Phi_j\rangle_H\right]_{i,j=1,\ldots,n_h}$$

$$\underline{A}_h^{\text{space}} = \left[a(\Phi_i,\Phi_j)\right]_{i,j=1,\ldots,n_h} \qquad \underline{w}^{l-\frac{1}{2}} = \frac{1}{2}\left(\underline{w}^l - \underline{w}^{l-1}\right).$$

Falls

$$f(\tau^l\otimes\Phi_j) = \int_0^T \left\langle g(t), \tau^l\otimes\Phi_j\right\rangle_H dt$$

$$\approx \frac{\Delta t}{2}\left\langle g(t^{l-1}) + g(t^l), \Phi_j\right\rangle_H = \Delta t \underline{g}_j^{l-\frac{1}{2}},$$

dann erhalten wir

$$\frac{1}{\Delta t}\underline{M}_h^{\text{space}}\left(\underline{w}^l - \underline{w}^{l-1}\right) + \underline{A}_h^{\text{space}}\underline{w}^{l-\frac{1}{2}} = \underline{g}^{l-\frac{1}{2}}, \qquad \underline{w}^0 := 0 \tag{14.11}$$

Crank-Nicolson-Verfahren.

Definition 14.6:
Für $w \in X_\delta$ setze $\overline{w}^k := \frac{1}{\Delta t}\int_{I^k} w(t)\ dt \in V$, wobei

$$\overline{w} := \sum_{k=1}^{K} \chi_{I^k} \otimes \overline{w}^k \in L_2(I;V) \qquad \textit{und}$$
$$|||w|||_{X,\delta}^2 := \|\dot{w}\|_{L_2(I;V')}^2 + \|\overline{w}\|_{L_2(I;V)}^2 + \|w(T)\|_H^2$$

Bemerkung 14.7:
Für $w \in X_\delta$ gilt $\overline{w} \in Y_\delta$.

Satz 14.8:
Sei $a(\cdot,\cdot)$ s.p.d., beschränkt und $\|\Phi\|_V^2 := a(\Phi,\Phi)$. Dann gilt:

$$\beta_\delta \equiv \inf_{w_\delta \in X_\delta} \sup_{v_\delta \in Y_\delta} \frac{b(w_\delta, v_\delta)}{|||w_\delta|||_{X,\delta}\|v_\delta\|_Y} = \gamma_\delta \equiv \sup_{w_\delta \in X_\delta} \sup_{v_\delta \in Y_\delta} \frac{b(w_\delta, v_\delta)}{|||w_\delta|||_{X,\delta}\|v_\delta\|_Y} = 1.$$

Beweis. Da $v_\delta \in Y_\delta$ bezüglich t stückweise konstant ist, gilt:

$$\begin{aligned}
\int_I a\left(w(t), v_\delta(t)\right)\ dt &= \sum_{k=1}^{K} \int_{I^k} a\left(w(t), v_\delta(t)|_{I^k}\right)\ dt \\
&= \sum_{k=1}^{K} a\left(\int_{I^k} w(t)\ dt, v_\delta|_{I^k}\right) \\
&= \sum_{k=1}^{K} a\left(\Delta t\overline{w}^k, v_\delta|_{I^k}\right) = \Delta t \sum_{k=1}^{K} a\left(\overline{w}^k, v_\delta|_{I^k}\right) \\
&= \int_0^T a\left(\overline{w}^k(t), v_\delta(t)\right)\ dt,
\end{aligned}$$

da der vorletzte Term die Rechteckregel für eine Trapezregel für eine Treppenfunktion ist. Damit gilt für $v_\delta \in Y_\delta$, $w_\delta \in X_\delta$:

$$b(w_\delta, v_\delta) = \int_I a\left(A_h^{-1}\dot{w}_\delta(t) + \overline{w}_\delta(t), v_\delta(t)\right)\ dt,$$

wobei $z_\delta := A_h^{-1}\dot{w}_\delta$ definiert ist durch

$$a\left(z_\delta, \Phi_h\right) = \langle \dot{w}_\delta, \Phi_h\rangle_{V'\times V} \qquad \forall \Phi_h \in V_h.$$

Für $\tilde{v} \in V'$ gilt

$$\|A_h^{-1}\tilde{v}\|_V^2 = a\left(A_h^{-1}\tilde{v}, A_h^{-1}\tilde{v}\right) = \left\langle \tilde{v}, A_h^{-1}\tilde{v}\right\rangle_{V'\times V} = (\tilde{v},\tilde{v})_{V'} = \|\tilde{v}\|_{V'}.$$

Wir zeigen später, dass $\forall v_\delta \in Y_\delta\ \exists! z_\delta \in X_\delta$ mit

$$\int_I a\left(A_h^{-1}\dot{z}_\delta(t) + \overline{z}_\delta(t), q_\delta(t)\right)\ dt = \int_I a\left(v_\delta(t), q_\delta(t)\right)\ dt \qquad \forall q_\delta \in Y_\delta. \tag{14.12}$$

Nun setze $v_\delta := A_h^{-1}\dot{z}_\delta + \overline{z}_\delta \in Y_\delta$ für $z_\delta \in X_\delta$

$$\begin{aligned}
\Rightarrow \sup_{v_\delta\in Y_\delta} \frac{b(w_\delta, v_\delta)}{\|v_\delta\|_Y} &= \sup_{z_\delta\in X_\delta} \frac{b(w_\delta, A_h^{-1}\dot{z}_\delta + \overline{z}_\delta)}{\|A_h^{-1}\dot{z}_\delta + \overline{z}_\delta\|_Y} \\
&= \sup_{z_\delta\in X_\delta} \frac{\int_0^T a\left(A_h^{-1}\dot{w}_\delta(t) + \overline{w}_\delta(t), A_h^{-1}\dot{z}_\delta(t) + \overline{z}_\delta(t)\right)\ dt}{\|A_h^{-1}\dot{z}_\delta + \overline{z}_\delta\|_Y} \\
&= \|A_h^{-1}\dot{z}_\delta + \overline{z}_\delta\|_Y,
\end{aligned}$$

denn "$\leq$"mit Cauchy-Schwarz und "$\geq$"durch Wahl von $z_\delta = w_\delta$. Damit gilt:

$$\begin{aligned}
\|A_h^{-1}\dot{w}_\delta + \overline{w}_\delta\|_Y^2 &= \|A_h^{-1}\dot{w}_\delta\|_Y^2 + \|\overline{w}_\delta\|_Y^2 + 2\int_0^T \langle \dot{w}_\delta(t), \overline{w}_\delta(t)\rangle_{V'\times V}\ dt \\
&= \|\dot{w}_\delta\|_{Y'}^2 + \|\overline{w}_\delta\|_Y^2 + \|w_\delta(T)\|_H^2 = |||w|||_{X,\delta}^2
\end{aligned}$$

$$\Rightarrow \sup_{v_\delta\in Y_\delta} \frac{b(w_\delta, v_\delta)}{\|v_\delta\|_Y} = \|A_h^{-1}\dot{w}_\delta + \overline{w}_\delta\|_Y = |||w|||_{X,\delta}^2,$$

also $\beta_\delta = \gamma_\delta = 1$. Es bleibt (14.12) zu zeigen: Seien $\lambda_j > 0$, $e_j \in \mathbb{R}^{n_h}$, $j = 1,\dots,n_h$ Eigenwerte und Eigenvektoren von A_h, d.h.

$$a\left(e_j, \Phi_h\right) = \lambda_j \left(e_j, \Phi_h\right)_H \qquad \forall \Phi_h \in V_h, \quad \|e_j\|_H = 1.$$

Gegeben sei

$$v_\delta = \sum_{k=1}^{K} v^k \tau^k \in Y_\delta, \qquad v^k = \sum_{j=1}^{n_h} v_j^k e_j \in V_h$$

und bestimme ζ_j^k, $k = 1,\dots,K$, eindeutig durch die Iteration

$$\zeta_j^0 = 0, \qquad \frac{1}{\Delta t}\left(\zeta_j^k - \zeta_j^{k-1}\right) + \frac{\lambda_j}{2}\left(\zeta_j^k + \zeta_j^{k-1}\right) = \lambda_j v_j^k$$

und definiere:

$$\begin{aligned}
z_\delta &:= \sum_{k=1}^{K}\sum_{j=1}^{n_h} \zeta_j^k e_j \sigma^k \in X_\delta \\
\Rightarrow \overline{z}_\delta &= \sum_{k=1}^{K} \overline{z}_\delta^k \chi_{I^k} \overset{\substack{Trapezr.\\ =\\ stk.lin.}}{} \frac{\Delta t}{2}\sum_{k=1}^{K}\left(z^k + z^{k-1}\right)\tau^k, \qquad \text{mit } z^k := z_\delta(t^k) \\
\dot{z}_\delta &= \frac{1}{\Delta t}\sum_{k=1}^{K}\sum_{j=1}^{n_h} \zeta_j^k e_j\left(\tau^k - \tau^{k+1}\right) = \frac{1}{\Delta t}\sum_{k=1}^{K}\sum_{j=1}^{n_h}\left(\zeta_j^k - \zeta_j^{k-1}\right)e_j
\end{aligned}$$

$\Rightarrow$ für $q_\delta = \sum_{k=1}^K q^k \tau^k \in Y_\delta$, $q^k = q_\delta(t^k)$ gilt:

$$\begin{aligned}
\int_I a\left(v_\delta(t), q_\delta(t)\right)\ dt &= \sum_{k,l=1}^K a\left(v^k, q^l\right) \underbrace{\int_I \tau^k(t)\tau^l(t)\ dt}_{=\Delta t \delta_{k,l}} \\
&= \Delta t \sum_{k=1}^K a\left(v^k, q^k\right) = \Delta t \sum_{k=1}^K \sum_{j=1}^{n_h} v_j^k \underbrace{a\left(e_j, q^k\right)}_{=\lambda_j \left(e_j, q^k\right)_H} \\
&= \sum_{k=1}^K \sum_{j=1}^{n_h} \Delta t \left(e_j, q^k\right)_H \left[\frac{1}{\Delta t}\left(\zeta_j^k - \zeta_j^{k-1}\right) + \frac{\lambda_j}{2}\left(\zeta_j^k + \zeta_j^{k-1}\right)\right] \\
&= \sum_{k=1}^K \left(\sum_{j=1}^{n_h} \left(\zeta_j^k - \zeta_j^{k-1}\right) e_j, q^k\right)_H \\
&\quad + \Delta t \sum_{k=1}^K \sum_{j=1}^{n_h} \underbrace{\lambda_j \left(\frac{\zeta_j^k + \zeta_j^{k-1}}{2} e_j, q^k\right)_H}_{=a\left(\frac{\zeta_j^k + \zeta_j^{k-1}}{2} e_j, q^k\right)} \\
&= \int_I \left\langle \dot{z}_\delta(t), q_\delta(t)\right\rangle_{V' \times V}\ dt + \int_I a\left(z_\delta(t), q_\delta(t)\right)\ dt,
\end{aligned}$$

denn $z_\delta = \sum_{k=1}^K \sum_{j=1}^{n_h} \zeta_j^k e_j \sigma^k$, also

$$\begin{aligned}
\int_I a\left(z_\delta(t), q_\delta(t)\right)\ dt &= \sum_{k=1}^K \int_{I^k} a\left(z_\delta(t), q_\delta(t)\right)\ dt \\
&= \sum_{k=1}^K \int_{I^k} a\left(z_\delta(t), q^k\right)\ dt \\
&\overset{Trapezr.}{=} \sum_{k=1}^K \frac{\Delta t}{2} \cdot a(\underbrace{z_\delta(t^k) + z_\delta(t^{k-1})}_{=\sum_{j=1}^{n_h}\left(\zeta_j^k - \zeta_j^{k-1}\right)e_j}, q^k)
\end{aligned}$$

Damit ist die Existenz in (14.12) gezeigt. Zur Eindeutigkeit: Seien $z_\delta, w_\delta \in X_\delta$ zwei Lösungen von (14.12), dann gilt:

$$\int_I a\left(A_h^{-1}\left(\dot{z}_\delta(t) - \dot{w}_\delta(t)\right) + \overline{z}_\delta(t) - \overline{w}_\delta(t), q_\delta(t)\right)\ dt = 0, \qquad \forall q_\delta \in Y_\delta.$$

Es ist $A_h^{-1}\left(\dot{z}_\delta(t) - \dot{w}_\delta(t)\right) + \overline{z}_\delta(t) - \overline{w}_\delta(t) \in Y_\delta$, damit verwenden wir das erste Argument als Testfunktion und erhalten:

$$\|\dot{z}_\delta - \dot{w}_\delta\|^2_{L_2(I;V')} + \|\overline{z}_\delta - \overline{w}_\delta\|^2_{L_2(I;V)} = 0$$

woraus $z_\delta = w_\delta$ in X_δ folgt. $\square$

Proposition 14.9:
Unter den Voraussetzungen von Satz 14.8 gilt für

$$\beta_\delta^* \equiv \inf_{v_\delta \in Y_\delta} \sup_{w_\delta \in X_\delta} \frac{b(w_\delta, v_\delta)}{|||w_\delta|||_{X,\delta} \|v_\delta\|_Y}$$

Beweis. [UP, 2012], Satz von Nečes. □

Nun zu parameter-abhängigen Problemen:

$$\mathcal{A}[w, v; \mu] := \int_I a(w(t), v(t); \mu)\ dt, \qquad b(w, v; \mu) := [\dot{w}, v]_{\mathcal{H}} + \mathcal{A}[w, v; \mu]. \tag{14.13}$$

- $\beta_\delta(\mu), \gamma_\delta(\mu)$ seien inf-sup- bzw. Stetigkeits-Konstanten.
- Die übliche Annahme der affinen Zerlegung von a wird getroffen.

Sei $V_N := span\{\xi_1, \dots, \xi_N\} \subset V_h$ ein RB-Raum bezüglich X, setze

$$X_{\Delta t,N} := S_{\Delta t} \otimes V_N$$
$$Y_{\Delta t,N} := Q_{\Delta t} \otimes V_N.$$

Definiere dann:

(14.14)	Residuum	$r_N(v;\mu) := f(v;\mu) - b(u_N(\mu), v; \mu) = b(e_N(\mu), v; \mu).$
(14.15)	Duales Residuum	$\tilde{r}_{\tilde{N}}(w;\mu) := -\ell(w) - b(w, z_{\tilde{N}}(\mu); \mu) = b(w, \tilde{e}_{\tilde{N}}(\mu); \mu).$
(14.16)	Primal-dualer Output	$s_N(\mu) := \ell(u_N(\mu)) - r_N(z_{\tilde{N}}(\mu)).$

Proposition 14.10:
(a) $|||u_\delta(\mu) - u_N(\mu)|||_{X,\delta} \le \frac{1}{\beta_\delta^{LB}} \|r_N(\mu)\|_{Y'}$
(b) $|s_\delta(\mu) - \ell(u_N(\mu))| \le \frac{\sqrt{T}}{\beta_\delta^{LB}} \|\ell\|_{V'} \|r_N(\mu)\|_{Y'}$
(c) $|s_\delta(\mu) - s_N(\mu)| \le \frac{\gamma_\delta}{(\beta_\delta^{LB})^2} \|r_N(\mu)\|_{Y'} \cdot |||\tilde{r}_{\tilde{N}}(\mu)|||_{X',\delta}$

Beweis.

$$\beta_\delta^{\mathrm{LB}} |||u_\delta(\mu) - u_N(\mu)|||_{X,\delta} \le \sup_{v_\delta \in Y_\delta} \frac{b(e_N(\mu), v_\delta)}{\|v_\delta\|_Y} = \sup_{v_\delta \in Y_\delta} \frac{r(v_\delta; \mu)}{\|v_\delta\|_Y} = \|r_N(\mu)\|_{Y'}.$$

$$\begin{aligned} |s_\delta(\mu) - \ell(u_N(\mu))| &\le \int_I |\ell(u_\delta(t;\mu)) - \ell(u_N(t;\mu))|\ dt \\ &\le \int_I \|\ell\|_{V'} \|u_\delta(t,\mu) - u_N(t;\mu)\|_V\ dt \\ &\le \|\ell\|_{V'} \sqrt{T} \underbrace{\|u_\delta(\mu) - u_N(\mu)\|_Y}_{\le |||\ell_N(\mu)|||_{X,\delta}} \end{aligned}$$

und dann mit (a). Bezüglich (c) gilt:

$$\begin{aligned}
|s_\delta(\mu) - s_N(\mu)| &= |\ell\left(u_\delta(\mu)\right) - \ell\left(u_N(\mu)\right) - r_N\left(z_{\tilde{N}}(\mu)\right)| \\
&= |\ell\left(e_N(\mu)\right) - b\left(e_N(\mu), z_{\tilde{N}}(\mu); \mu\right)| \\
&= |b\left(e_N(\mu), z_\delta(\mu); \mu\right) - b\left(e_N(\mu), z_{\tilde{N}}(\mu); \mu\right)| \\
&= |b\left(e_N(\mu), \tilde{e}_{\tilde{N}}(\mu); \mu\right)| \\
&\leq \gamma_\delta \|e_N(\mu)\|_X \|\tilde{e}_{\tilde{N}}\|_Y
\end{aligned}$$

und weiter mit Standard-Argumenten. □

Bemerkung 14.11:

(a) Die Normen, Bilinearformen, Supremierer können durch den Tensorprodukt-Ansatz weitgehend entkoppelt werden.

(b) Die meisten Techniken aus dem elliptischen Fall lassen sich direkt übertragen (Greedy, SSCM, ...).

(c) Erweiterungen für Systeme (→ Wellengleichung) sind möglich.

(d) Insbesondere interessant auch für zeit-periodische Probleme (Kristina Steih).

15 Quadratisch Nichtlineare Probleme

Nun sei $b(\cdot, \cdot\,; \mu)$ nur noch semilinear, also linear in der zweiten Komponente.

Beispiel 15.1:
(a) Quadratische Nichtlinearitäten:

$$b(v, w; \mu) = a_0(v, w; \mu) + a_1(v, v; w; \mu)$$

mit einer Bilinearform $a_0 : X \times Y \times \mathcal{D} \to \mathbb{R}$ und einer Trilinearform $a_0 : X \times X \times Y \times \mathcal{D} \to \mathbb{R}$.
(b) Polynomiale Nichtlinearitäten durch Multilinearformen.
(c) Allgemeinere Nichtlinearitäten, z.B.: $b(v, w; \mu) = (e^{\mu v}, w)$.

Spezieller:

Beispiel 15.2:
(a) Nichtlineare Reaktion:

$$b(w, v; \mu) = \mu_1(\nabla w, \nabla v)_0 + \mu_2(w^2, v)_0$$

hier also $a_0(w, v; \mu) = \mu_1(\nabla w, \nabla v)_0$, $a_1(v, w, z; \mu) = \mu_2(vw, z)_0$. Mann kann zeigen, dass a_1 stetig ist (mittels $H_0^1 \hookrightarrow L_4(\Omega)$).
(b) Viskose Burgerss-Gleichung:

$$b(w, v; \mu) = \mu_1(\nabla w, \nabla v)_0 \; \mu_2(w^2, \nabla v)_0.$$

(c) Navier-Stokes:

$$\begin{aligned} -\mu_1 \Delta u + \mu_2 u \cdot \nabla u + \nabla \varphi &= f \\ \nabla \cdot u &= 0 \end{aligned}$$

Zunächst stellt sich die Frage der Wohlgestelltheit - lineare Funktionalanalysis wie etwa Lax-Milgram hilft hier nicht viel.

Satz 15.3 (Brouwer'scher Fixpunktsatz)**:**
Sei X endlich-dimensional, $\emptyset \neq C \subset X$ konvex und kompakt sowie $F : C \to C$ stetig. Dann besitzt F mindestens einen Fixpunkt.

Korollar 15.4:
Sei X ein endlich-dimensionaler Hilbertraum, $P : X \to X$ stetig mit

$$\exists \xi > 0 : \ (P(f), f)_X > 0 \qquad \forall f \in X, \ \|f\|_X = \xi. \tag{15.1}$$

Dann existiert ein $f \in H$ mit $\|f\|_X \leq \xi$ und $P(f) = 0$.

Beweis. Per Widerspruch: Angenommen $P(f) \neq 0$ auf $D := \{f \in X \mid \|f\|_X \leq \xi\}$. Dann ist die Abbildung

$$f \mapsto -\xi \frac{P(f)}{\|P(f)\|_X}, \qquad f : D \to D$$

stetig. Da D konvex und kompakt ist, existiert nach Satz 15.3 ein Fixpunkt $f \in X$, also

$$f = -\xi \frac{P(f)}{\|P(f)\|_X}.$$

Damit gilt $\|f\|_X = \xi$ und

$$(P(f), f)_X = -\xi \left(P(f), \frac{P(f)}{\|P(f)\|_X} \right)_X = -\underbrace{\|f\|_X}_{>0} \underbrace{\|P(f)\|_X}_{>0} < 0 \qquad \lightning$$

□

Damit können wir folgendes Existenz-Resultat zeigen:

Satz 15.5:
Sei X separabel, $a : X \times X \times X \to \mathbb{R}$ trilinear mit

(i) $\exists \alpha > 0 : \ a(v, v, v) \geq \alpha \|v\|_X^2 \quad \forall v \in X.$
(ii) Die Abbildung $u \mapsto a(u, u, v), \quad v \in X$ ist schwach stetig in X, d.h.

$$u_m \rightharpoonup u \text{ in } X \Rightarrow \lim_{m \to \infty} a(u_m, u_m, v) = a(u, u, v) \ \forall v \in X. \tag{15.2}$$

Dann besitzt das Problem

$$\text{Suche } u \in X : \qquad a(u, u, v) = (f, v)_X \quad \forall v \in X \tag{15.3}$$

mindestens eine Lösung in X.

Beweis. Konstruieren eine Folge von Näherungslösungen mit dem Galerkin-Verfahren. Da X separabel, existiert eine Folge $(w_i)_{i \geq 1}$ in X mit

(a) $X_m := \{w_1, \ldots, w_m\}$ ist linear unabhängig $\forall m \geq 1$.
(b) $span\,\{w_i \mid i \in \mathbb{N}\}$ ist dicht in X.

Betrachte nun das m-dimensionale Problem:

$$\text{(15.4)} \qquad \text{Suche } u_m \in X_m : \qquad a(u_m, u_m, v) = (f, v)_X \quad \forall v \in X_m.$$

Wir wollen die Existenz einer Lösung von (15.4) zeigen. Betrachte dazu den Operator $P_m : X_m \to X_m$ definiert durch

$$(P_m(v), w_i)_X := a(v, v, w_i) - (f, w_i)_X \qquad \forall 1 \le i \le m,\ \forall v \in X_m.$$

$$\begin{aligned} \Rightarrow (P_m(v), v)_X :=& a(v, v, v) - (f, v)_X \overset{(i),C.S.U.}{\geq} \alpha \|v\|_X^2 - \|f\|_{X'} \cdot \|v\|_X \\ =& \|v\|_X \left(\alpha \|v\|_X - \|f\|_{X'}\right). \end{aligned}$$

Wenn wir also $v \in X$ so wählen, dass $\|v\|_X = \xi > \frac{\|f\|_{X'}}{\alpha}$, dann gilt $(P_m(v), v)_X > 0$. Wegen (ii) ist P_m auf V_m ($dim V_m = m < \infty$) stetig. Mit Korollar 15.4 existiert u_m mit

$$\text{(15.5)} \qquad 0 = (P_m(u_m), u_m)_X \geq \|u_m\|_X \left(\alpha \|u_m\|_X - \|f\|_{X'}\right)$$

$$\text{(15.6)} \qquad \|u_m\|_X \leq \xi = \frac{1}{\alpha} \|f\|_{X'},$$

also ist $(u_m)_m$ für $m \to \infty$ beschränkt $\Rightarrow \exists$ Teilfolge $\left(u_{m_p}\right)_{p \in \mathbb{N}}$, die in X schwach konvergiert, also existiert ein $u \in X$ mit $u_{m_p} \rightharpoonup u$ in X mit $p \to \infty$. Nach (ii) folgt

$$\lim_{p \to \infty} a\left(u_{m_p}, u_{m_p}, v\right) = a(u, u, v) \qquad \forall v \in X.$$

Wegen (15.4) und (a) folgt im Grenzwert $a(u, u, w_i) = (f, w_i)_X \ \forall i$ und mit (b) dann $a(u, u, w) = (f, w)_X \ \forall w \in X$. Also löst $u \in X$ (15.3). □

Für die Eindeutigkeit brauchen wir stärkere Voraussetzungen:

Satz 15.6:
Angenommen, es gelte

(i) Die Form b ist gleichmäßig elliptisch bezüglich des zweiten und dritten Arguments, d.h.

$$b(w, v, v) \geq \alpha \|v\|_X^2 \qquad \forall v, w \in X.$$

(ii) Die Abbildung $w \mapsto B(w) \in X'$ mit $(B(w)u, v)_X := b(w, u, v)$ ist lokal Lipschitzstetig in X, d.h.

$$\begin{aligned} &\exists L : \mathbb{R}^+ \to \mathbb{R}^+ \text{ monoton wachsend und stetig mit} \\ &|b(w_1, u, v) - b(w_2, u, v)| \leq L(\xi) \|u\|_X \|v\|_X \|w_1 - w_2\|_X \qquad \forall u, v \in X \end{aligned}$$

für alle $\xi > 0$ und alle $w_1, w_2 \in D_\xi := \{v \in X \mid \|v\|_X \leq \xi\}$.

Falls $\frac{\|f\|_{X'}}{\alpha^2} L\left(\frac{\|f\|_{X'}}{\alpha}\right) < 1$, dann besitzt (15.3) genau eine Lösung $u \in X$.

Beweis. Aus (i) folgt mit Lax-Milgram, dass $B(w) \in \mathcal{L}(X, X')$ für alle $w \in X$ invertierbar ist. Für $T(w) := B(w)^{-1}$ gilt $T(w) \in \mathcal{L}(X', X)$ und mit (i): $\|T(w)\|_{\mathcal{L}(X',X)} \leq \frac{1}{\alpha}$. Damit lautet (15.3) $B(u)u = f$ in X' beziehungsweise $u = T(u)f$ in X. Ziel: Zeige, dass $v \mapsto T(v)f$ eine Kontraktion von D_ξ nach D_ξ mit $\xi = \frac{1}{\alpha}\|f\|_{X'}$ ist:

(1) Selbstabbildung: Sei $v \in D_\xi$, dann gilt

$$\|T(v)f\|_X \leq \|T(v)\|_{\mathcal{L}(X',X)}\|f\|_{X'} \leq \frac{1}{\alpha}\|f\|_{X'} = \xi,$$

also $T(v)f \in D_\xi$.

(2) Kontraktion: Seien $u, v \in D_\xi$, dann gilt $T(u) - T(v) = T(u)\,[B(v) - B(u)]\,T(v)$, (denn $T(u)B(u) = I$) und damit $\|T(u) - T(v)\|_{\mathcal{L}(X',X)} \leq \frac{1}{\alpha^2}\|B(v) - B(u)\|_{\mathcal{L}(X,X')}$, also

$$\|T(u)f - T(v)f\|_X \leq \frac{1}{\alpha^2} \underbrace{\|B(v) - B(u)\|_{\mathcal{L}(X,X')}}_{\overset{(ii)}{<} L(\xi)\|v-u\|_X} \|f\|_{X'} \overset{VS}{<} \|v - u\|_X, \quad \forall f \in X',$$

Rest mit Brouwer'schen Fixpunktsatz. □

Korollar 15.7:
Sei $b(v, w; \mu) := a_0(v, w; \mu) + a_1(v, v, w; \mu)$ wie in Beispiel 15.1 (a) mit

(a) Beschränktheit: $\exists 0 < \rho_i < \infty$, $i = 0, 1$ mit

$$|a_0(u, v; \mu)| \leq \rho_0(\mu)\|u\|_X\|v\|_X, \qquad \rho_0(\mu) \leq \rho_0$$
$$|a_1(u, w, v; \mu)| \leq \rho_1(\mu)\|u\|_X\|v\|_X\|w\|_X, \qquad \rho_1(\mu) \leq \rho_1$$

(b) Gleichmäßige inf-sup: $\exists \beta_0 > 0$ mit $\beta\,(u(\mu); \mu) \geq \beta_0$ für alle $\mu \in \mathcal{D}$ mit

$$\beta(z; \mu) := \inf_{v \in X} \sup_{w \in X} \frac{db(v, w; \mu)\,[z]}{\|v\|_X\|w\|_X}$$

mit der Fréchet-Ableitung $db(v, w; \mu)\,[z]$ von $b(v, w; \mu)$ an der Stelle $z \in X$ (in diesem Fall hier: $= a_0(v, w; \mu) + a_1(v, z, w; \mu) + a_1(z, v, w; \mu)$).

Dann existiert lokal nahe $u(\mu)$ eine eindeutige Lösung.

Beweis. Es müssen die Voraussetzungen von Satz 15.6 verifiziert werden (Übung). □

Bemerkung 15.8:
Die Bedingung (b) ist problematisch, da sie die unbekannte Lösung $u(\mu)$ enthält. Die Annahme

$$\beta(z; \mu) \geq \beta_0 \qquad \forall z \in X$$

wäre aber viel zu stark und oft unrealistisch. Man versucht die Korrektgestelltheit a-posteriori für einen gegebenen (neuen) Parameter $\mu \in \mathcal{D}$ zu sichern (Brezzi-Rappaz-Raviart, BRR-Theorie).

Im folgenden sei stets

$$b(v, w; \mu) = a_0(v, w; \mu) + a_1(v, v, w; \mu) \tag{15.7}$$

mit

$$a_1(u, v, w; \mu) = a_1(v, u, w; \mu) \qquad \forall u, v, w \in X,\ \forall \mu \in \mathcal{D}. \tag{15.8}$$

Damit gilt $db(v, w; \mu)[z] = 2a_1(v, z, w; \mu) + a_0(v, w; \mu)$.

Algorithmus 15.9 (Newton-Verfahren)**:**

Wähle $u^{(0)} \in X$
repeat
  Bestimme $z^{(k)}$ *als Lösung von* $db(u^{(k)}, w; \mu)\left[z^{(k)}\right] = -b(u^{(k)}, w; \mu) + f(w)\ \forall w \in X$
  also

$$2a_1\left(u^{(k)}, z^{(k)}, w; \mu\right) = a_0\left(u^{(k)}, w; \mu\right) + f(w) - b\left(u^{(k)}, w; \mu\right)$$

  $u^{(k+1)} := u^{(k)} + z^{(k)}$
until $\|u^{(k+1)} - u^{(k)}\| < \varepsilon_{tol}$;

Falls das Verfahren konvergiert, ist die Existenz einer Lösung ebenfalls gesichert.

Das RB-Problem lautet: Sei $X_N \subset X$ ein RB-Raum, suche $u_N(\mu) \in X_N$, $S_N(\mu) \in \mathbb{R}$:

$$a_0\left(u_N(\mu), v_N; \mu\right) + a_1\left(u_N(\mu), u_N(\mu), v_N; \mu\right) = f(v_N; \mu) \qquad \forall v_N \in X_N \tag{15.9}$$

$$S_N(\mu) = \ell\left(u_N(\mu)\right) \tag{15.10}$$

Das Newton-Verfahren für (15.9) ist dann ganz analog.

Bezüglich der offline-online-Zerlegung nehmen wir an, dass a_0 parametrisch affin ist und

$$a_1(u, v, w; \mu) = \sum_{q=1}^{Q_1} \Theta_1^q(\mu) a_1^q(u, v, w) \tag{15.11}$$

Man berechnet dann offline

$$\mathbb{A}_N^{1,q} := \left(a_1\left(\xi_i, \xi_j, \xi_k\right)\right)_{i,j,k=1,\ldots,N} \in \mathbb{R}^{N \times N \times N},$$

dann ist offline-online wie bisher. Nun zu Fehlerschätzern. Setze dazu $F(v; \mu) \in X'$ definiert durch

$$\langle F(v; \mu), w\rangle_{X' \times X} := a_1(v, v, w; \mu) + a_0(v, w; \mu) - f(w; \mu).$$

Dann ist die Lösung $u(\mu)$ Nullstelle von $F(\cdot\,; \mu)$, $\quad F_N := F|_{X_N}$.

Weiterhin führen wir die Bezeichnung $dF(u) = dF(u; \mu)$ für die Fréchet-Ableitung ein, definiert durch

$$\begin{aligned}\langle dF(v; \mu)[z], w\rangle_{X' \times X} :=& db(v, w; \mu)[z] \\ =& a_0(v, w; \mu) + 2a_1(v, z, w; \mu), \qquad w \in X,\end{aligned}$$

d.h. $dF(v;\mu)[z] \in X'$, $dF(v;\mu) \in \mathcal{L}(X, X')$. Der Newton-Schritt lautet dann

$$z^{(k)} \in X : \qquad dF\left(u^{(k)}\right)\left[z^{(k)}\right] = -F\left(u^{(k)}\right).$$

Der folgende Satz gilt für allgemeine Funktionen $F(v;\mu)$ - der obige Fall (quadratische Nichtlinearität) ist ein Spezialfall.

Satz 15.10 (Fehlerschätzer für Newton-Iteration)**:**
Seien $u^{(0)} := u_N^{(0)} = 0$ *und* $dF\left(u^{(k)}\right), dF_N\left(u_N^{(k)}\right)$ *invertierbar (also* $\left(u^{(k)}\right)_k, \left(u_N^{(k)}\right)_k$ *gemäß Newton-Verfahren wohl-definiert). Weiter seien*

(i) $\|u^{(k)}\|_X, \|u_N^{(k)}\|_X \le \gamma_k(\mu) \quad \forall k.$
(ii) $\left\|\left(dF\left(u^{(k)}\right)\right)^{-1}\right\|_{\mathcal{L}(X',X)} \le \frac{1}{\beta(\mu)} \quad \forall k.$
(iii) $dF : X \to X'$ *sei Lipschitz-stetig auf* $B := B_0(\gamma_k(\mu))$, *d.h.*

$$\|dF(u) - dF(v)\|_{\mathcal{L}(X,X')} \le L_{dF}\|u - v\|_X \qquad \forall u, v \in B.$$

(iv) F *sei auf* B *Lipschitz-stetig, d.h.*

$$\|F(u) - F(v)\|_{X'} \le L_F\|u - v\|_X \qquad \forall u, v \in B.$$

Dann gilt

(15.12)

$$\|u^{(k)}(\mu) - u_N^{(k)}(\mu)\|_X \le \Delta_N^{(k)}(\mu) := \sum_{i=1}^{k} \frac{\|r^{(i-1)}\|_{X'}}{\beta(\mu)} \prod_{j=i}^{k-1} \left(1 + \frac{L_F}{\beta(\mu)} + \frac{L_{dF}}{\beta(\mu)}\|z_N^{(j)}\|\right)$$

mit dem Residuum $r^{(k)} \in X'$ *definiert durch*

$$r^{(k)} \in X' := dF\left(u_N^{(k)}\right)\left[z_N^{(k)}\right] + F\left(u_N^{(k)}\right).$$

Beweis. Für $k \ge 0$ gilt

$$\begin{aligned} dF\left(u^{(k)}\right)\left[z^{(k)} - z_N^{(k)}\right] &= \overbrace{dF\left(u^{(k)}\right)\left[z^{(k)}\right]}^{=-F\left(u^{(k)}\right)} - dF\left(u^{(k)}\right)\left[z_N^{(k)}\right] \\ &= \overbrace{-F\left(u_N^{(k)}\right) - dF\left(u_N^{(k)}\right)\left[z_N^{(k)}\right]}^{=-r^{(k)}} \\ &\quad + dF\left(u_N^{(k)}\right)\left[z_N^{(k)}\right] - dF\left(u^{(k)}\right)\left[z_N^{(k)}\right] \\ &\quad + F\left(u_N^{(k)}\right) - F\left(u^{(k)}\right) \end{aligned}$$

$$\overset{(ii)}{\Rightarrow} \|z^{(k)} - z_N^{(k)}\|_X \leq \frac{1}{\beta(\mu)} \Big\{ \|r^{(k)}\|_{X'} + \underbrace{\|F\left(u_N^{(k)}\right) - F\left(u^{(k)}\right)\|_{X'}}_{\overset{(iv)}{\leq} L_F \|u_N^{(k)} - u^{(k)}\|_X} + \underbrace{\left\| dF\left(u_N^{(k)}\right)\left[z_N^{(k)}\right] - dF\left(u^{(k)}\right)\left[z_N^{(k)}\right] \right\|_{\mathcal{L}(X,X')}}_{\overset{(iii)}{\leq} L_{dF} \|u_N^{(k)} - u^{(k)}\|_X \|z_N^{(k)}\|_X} \Big\}$$

und damit

$$\begin{aligned} \|u^{(k+1)}(\mu) - u_N^{(k+1)}(\mu)\|_X &= \|u^{(k)}(\mu) + z^{(k)}(\mu) - u_N^{(k)}(\mu) - z_N^{(k)}(\mu)\|_X \\ &\leq \|u^{(k)}(\mu) - u_N^{(k)}(\mu)\|_X + \|z^{(k)}(\mu) - z_N^{(k)}(\mu)\|_X \\ &\leq \left\{ 1 + \frac{L_F}{\beta(\mu)} + \frac{L_{dF}}{\beta(\mu)} \|z_N^{(k)}\|_X \right\} \|u^{(k)}(\mu) - u_N^{(k)}(\mu)\|_X + \frac{\|r^{(k)}\|_{X'}}{\beta(\mu)} \end{aligned} \tag{15.13}$$

Der Rest folgt mit vollständiger Induktion über k. Für $\underline{k = 0}$:

$$\|u^{(k)}(\mu) - u_N^{(k)}(\mu)\|_X \overset{n.V.}{=} 0 = \Delta_N^{(0)}(\mu)$$

Für $\underline{k \to k+1}$: Es gelte (15.12) für ein k

$$\begin{aligned} \|u^{(k+1)}(\mu) - u_N^{(k+1)}(\mu)\|_X &\overset{(15.13),(iv)}{\leq} \left(1 + \frac{L_F}{\beta(\mu)} + \frac{L_{dF}}{\beta(\mu)} \|z_N^{(k)}\|_X \right) \Delta_N^{(k)}(\mu) + \frac{\|r^{(k)}\|_{X'}}{\beta(\mu)} \\ &= \left(1 + \frac{L_F}{\beta(\mu)} + \frac{L_{dF}}{\beta(\mu)} \|z_N^{(k)}\|_X \right) \sum_{i=1}^{k} \frac{\|r^{(i-1)}\|_{X'}}{\beta(\mu)} \prod_{j=i}^{k-1} \left(1 + \frac{L_F}{\beta(\mu)} + \frac{L_{dF}}{\beta(\mu)} \|z_N^{(j)}\|_X \right) \\ &\quad + \frac{\|r^{(k)}\|_{X'}}{\beta(\mu)} \\ &= \sum_{i=1}^{k} \frac{\|r^{(i-1)}\|_{X'}}{\beta(\mu)} \prod_{j=i}^{k} \left(1 + \frac{L_F}{\beta(\mu)} + \frac{L_{dF}}{\beta(\mu)} \|z_N^{(j)}\|_X \right) + \frac{\|r^{(k)}\|_{X'}}{\beta(\mu)} \\ &= \sum_{i=1}^{k+1} \frac{\|r^{(i-1)}\|_{X'}}{\beta(\mu)} \prod_{j=i}^{k} \left(1 + \frac{L_F}{\beta(\mu)} + \frac{L_{dF}}{\beta(\mu)} \|z_N^{(j)}\|_X \right) \\ &= \Delta_N^{k+1}(\mu). \end{aligned}$$

□

Bemerkung 15.11:

(a) Die Konvergenz des Newton-Verfahrens wird hier <u>nicht</u> vorausgesetzt, die Iterierten müssen nur in B bleiben.

(b) Es gilt

$$\Delta_N^{(k)}(\mu) \geq \sum_{i=1}^{k} \frac{\|r^{(i-1)}\|_{X'}}{\beta(\mu)} (k - i) \nrightarrow 0 \qquad (k \to \infty)$$

Dies ist sinnvoll, angenommen es gelte $\|u(\mu) - u^{(k)}(\mu)\|_X, \|u_N(\mu) - u_N^{(k)}(\mu)\|_X < \varepsilon$ *(Newton-Verfahren konvergiert), dann*

$$\begin{aligned}\|u(\mu) - u_N(\mu)\|_X &\leq \|u(\mu) - u^{(k)}(\mu)\|_X + \|u_N(\mu) - u_N^{(k)}(\mu)\|_X + \|u^{(k)}(\mu) - u_N^{(k)}(\mu)\|_X \\ &\leq 2\varepsilon + \Delta_N^{(k)}(\mu)\end{aligned}$$

$\Rightarrow$ *für kleines* ε *wird* $\Delta_N^{(k)}(\mu)$ *der dominierende Term.*

(c) *Nach (a) wächst* $\Delta_N^{(k)}(\mu)$ *monoton mit* $k \Rightarrow$ *Effektivität ist nicht gut. Falls aber* $u^{(k)}(\mu) \in X_N$*, dann wird die Lösung reproduziert, d.h.* $u_N^{(k)}(\mu) = u^{(k)}(\mu)$.

Offline-online Prozedur für die Dualnorm des Residuums

Es gilt:

$$\begin{aligned} r^{(k)} &= dF\left(u_N^{(k)};\mu\right)\left[z_N^{(k)}\right] + F\left(u_N^{(k)};\mu\right) \\ &\overset{\text{im quadr. Fall}}{=} 2a_1\left(u_N^{(k)}, z_N^{(k)}, \cdot\,;\mu\right) + a_0\left(u_N^{(k)}, \cdot\,;\mu\right) + a_1\left(u_N^{(k)}, u_N^{(k)}, \cdot\,;\mu\right) - f(\cdot) \\ &= a_1\left(u_N^{(k)}, 2z_N^{(k)} + u_N^{(k)}, \cdot\,;\mu\right) + a_0\left(u_N^{(k)}, \cdot\,;\mu\right) - f(\cdot)\end{aligned}$$

- Annahme der affinen Zerlegbarkeit von $a_0, a_1 \Rightarrow$ Residuum ist affin zerlegbar wie in Paragraph 6.
- offline-Berechnung der entsprechenden Riesz-Repräsentanten.
- online: affine Kombination der vorab berechneten Terme.

Bemerkung 15.12:
Zum dualen Problem: Die Lagrange-Funktion lautet

$$\mathcal{L}(u, z) := \ell(u) + \langle F(u), z\rangle_{X' \times X}\,,$$

dann gilt

$$\langle \mathcal{L}_u(u, z), v\rangle := \ell(v) + \langle dF(v)z, u\rangle\,.$$

Dies führt auf ein lineares Problem, also ist die duale nur so teuer wie eine Newton-Iteration!

Damit lautet also das duale Problem

$$\text{Suche } z_{\tilde{N}} \in \tilde{X}_{\tilde{N}}: \quad \langle dF(v)z_{\tilde{N}}, u_N(\mu)\rangle = -\ell(v), \qquad v \in \tilde{X}_{\tilde{N}}. \tag{15.14}$$

Für die a-posteriori-Fehlerschätzung und die a-posteriori-Untersuchung der Korrektgestelltheit benötigen wir einige Bezeichnungen und Abkürzungen:

$$R_N(v;\mu) = \langle F(u_N(\mu)), v\rangle\,, \qquad R_N(\mu) := \|R_N(\cdot\,;\mu)\|_{X'} \quad \text{(Residuum)} \tag{15.15}$$

$$\beta_N(\mu) := \beta\,(u_N(\mu);\mu) \quad \text{(vgl. Korollar 15.7)} \tag{15.16}$$

$$\gamma_N(\mu) := \gamma\,(u_N(\mu);\mu)\,, \qquad \gamma(z;\mu) := \sup_{v\in X} \frac{\langle F(z;\mu), v\rangle}{\|v\|_X} \tag{15.17}$$

$$\text{„Proximity indicator“ :} \qquad \tau_N(\mu) := 4\rho_1(\mu)\frac{R_N(\mu)}{\beta_N(\mu)^2} \tag{15.18}$$

mit $\rho_1(\mu)$ aus Korollar 15.7

$$\Delta_N(\mu) := \frac{\beta_N(\mu)}{2\rho_1(\mu)}\left(1-\sqrt{1-\tau_N(\mu)}\right) \tag{15.19}$$

Der folgende Satz ist auch als BRR-Satz bekannt.

Satz 15.13:
Falls $\tau_N(\mu) < 1$, so existiert eine eindeutige Lösung $u(\mu) \in B_{u_N(\mu)}\left(\beta_N(\mu)\,(2\rho_1(\mu))^{-1}\right)$ von

$$\langle F(u(\mu);\mu), v\rangle_{X'\times X} = 0, \qquad \forall v \in X \tag{15.20}$$

mit F wie in Satz 15.10 und es gilt $\|u(\mu) - u_N(\mu)\|_X \le \Delta_N(\mu)$.

Beweis. Seien $w_1, w_2 \in X \Rightarrow$ Hauptsatz der Differential- und Integralrechnung

$$F(w_2;\mu) - F(w_1;\mu) = \int_0^1 dF(w_2 - w_1;\mu)\,[w_1 + t(w_2 - w_1)]\;\, dt.$$

Weiter gilt für $z_1, z_2 \in X$

$$\begin{aligned} |\langle dF(v;\mu)\,[z_2]\,, w\rangle \,-\, \langle dF(v;\mu)\,[z_1]\,, w\rangle| &= 2\,|a_1(v, z_2 - z_1, w;\mu)| \\ &\overset{Kor15.7}{\le} 2\rho_1(\mu)\|v\|_X\|w\|_X\|z_1 - z_2\|_X. \end{aligned} \tag{15.21}$$

Betrachte nun den Operator $H^\mu : X \to X$ definiert durch

$$\begin{aligned} \langle dF(H^\mu w;\mu)\,[u_N(\mu)]\,, v\rangle_{X'\times X} = &\langle dF(w;\mu)\,[u_N(\mu)]\,, v\rangle_{X'\times X} \\ &- \langle F(w;\mu), v\rangle_{X'\times X}\,, \qquad v \in X \end{aligned}$$

für $w \in X$ gegeben. Nun folgt aus $\tau_N(\mu) < 1$ sofort $\beta_N(\mu) > 0$, also ist H^μ wohldefiniert ($X = X^{\mathcal{N}}$ ist endlich-dimensional).

Ziel: Zeige, dass ein Fixpunkt w^* von H^μ existiert, denn dann gilt $F(w^*;\mu) = 0$. Schränke dazu H^μ auf $\overline{B}_{u_N(\mu)}(\alpha)$ mit geeignetem α ein.

(i) Kontraktion: Seien $w_1, w_2 \in \overline{B}_{u_N(\mu)}(\alpha)$, dann gilt:

$$\begin{aligned} &dF(H^\mu w_2 - H^\mu w_1;\mu)\,[u_N(\mu)] \\ &\quad \overset{\text{Def } H^\mu}{=} dF(w_2 - w_1;\mu)\,[u_N(\mu)] - F(w_2;\mu) + F(w_1;\mu) \\ &\quad \overset{\text{HDI}}{=} dF(w_2 - w_1;\mu)\,[u_N(\mu)] - \int_0^1 dF(w_2 - w_1;\mu)\,[w_1 + t(w_2 - w_1)]\ dt \\ &\quad = \int_0^1 \{dF(w_2 - w_1;\mu)\,[u_N(\mu)] - dF(w_2 - w_1;\mu)\,[w_1 + t(w_2 - w_1)]\}\ dt \end{aligned}$$

und damit

$$\begin{aligned} |dF(H^\mu w_2 - H^\mu w_1;\mu)\,[u_N(\mu)]| \leq& \\ \leq \int_0^1 2\rho_1(\mu)\|w_2 - w_1\|_X \underbrace{\|u_N(\mu) - \underbrace{(w_1 + t(w_2 - w_1))}_{\in \overline{B}_{u_N(\mu)}(\alpha)}\|_X}_{\leq \alpha}\ dt& \\ \leq 2\alpha\rho_1(\mu)\|w_2 - w_1\|_X& \end{aligned}$$

Weiter gilt

$$\begin{aligned} \|H^\mu w_2 - H^\mu w_1\|_X &\leq (\beta_N(\mu))^{-1} \sup_{v\in X} \frac{\langle dF(H^\mu w_2 - H^\mu w_1;\mu)\,[u_N(\mu)]\,, v\rangle}{\|v\|_X} \\ &\leq \underbrace{(\beta_N(\mu))^{-1}\, 2\alpha\rho_1(\mu)}_{<1} \|w_2 - w_1\|_X \end{aligned}$$

$\Rightarrow H^\mu$ ist eine Kontraktion für $\alpha \in I_1 := \left[0, \beta_N(\mu)\cdot(2\rho_1(\mu))^{-1}\right)$

(ii) Selbstabbildung: Es gilt

$$\begin{aligned} &dF(H^\mu w - u_N(\mu);\mu)\,[u_N(\mu)] = dF(w - u_N(\mu);\mu)\,[u_N(\mu)] - F(w;\mu) \\ &= dF(w - u_N(\mu);\mu)\,[u_N(\mu)] - (F(w;\mu) - F(u_N(\mu);\mu)) - F(u_N(\mu);\mu) \\ &= \underbrace{\int_0^1 dF(w - u_N(\mu);\mu)\,[u_N(\mu) - (u_N(\mu) + t(w - u_N(\mu)))]\ dt}_{\leq 2\rho_1(\mu)\|w-u_N(\mu)\|_X^2 \int_0^1 t\ dt \leq 2\rho_1(\mu)\|w-u_N(\mu)\|_X^2 \frac{1}{2}} - F(u_N(\mu);\mu) \end{aligned}$$

Damit gilt:

$$\begin{aligned} &\|dF(H^\mu w - u_N(\mu);\mu)\,[u_N(\mu)]\,\|_{X'} \\ &\qquad \leq \rho_1(\mu) \underbrace{\|w - u_N(\mu)\|_X^2}_{\leq \alpha^2,\ \text{falls } w\in\overline{B}_{u_N(\mu)}(\alpha)} + \underbrace{\|F(u_N(\mu);\mu)\|_{X'}}_{\leq R_N(\mu),\ \text{nach (15.5)}} \leq \alpha^2\rho_1(\mu) + R_N(\mu) \end{aligned}$$

Wie bei (i) folgt nun

$$\|H^\mu w - u_N(\mu)\|_X \leq (\beta_N(\mu))^{-1}\left(\alpha^2\rho_1(\mu) + R_N(\mu)\right),$$

also ist H^μ Selbstabbildung, falls

$$(\beta_N(\mu))^{-1}\left(\alpha^2\rho_1(\mu) + R_N(\mu)\right) \leq \alpha,$$

also für $\alpha \geq \Delta_N(\mu)$.

Damit sind die Voraussetzungen des Brouwer'schen Fixpunktsatzes für

$$\alpha \in \left[\Delta_N(\mu), \beta_N(\mu)\left(\alpha^2\rho_1(\mu)\right)^{-1}\right)$$

erfüllt und die Wahl $\alpha = \Delta_N(\mu)$ ergibt die Fehlerschätzung. □

Bemerkung 15.14:

(a) $\tau_N(\mu)$ *ist also ein a-posteriori Indikator für Korrekt-Gestelltheit. Man kann so auch eine inf-sup untere Schranke bekommen (C. Camto, T. Tonn, K. Urban, SINum 2009).*

(b) Analoge Aussagen bekommt man auch für Output-Funktionale, siehe auch K. Veroy, C. Prud'homme, A.T. Patera, CRAS 2003.

(c) Manchmal hängt zwar b affin vom Parameter ab, dieser geht aber nicht direkt ein, Bsp VSP. Dann hat man die Form $a_0(v, w; h_0(\mu)) + a_1(v, v, w; h_1(\mu))$ *mit* $h_i : \mathcal{D} \times \Omega \to \mathbb{R}$ *(z.B. variable Koeffizienten). Man wendet dann die EIM auf* h_i *an (CTU 2009).*

(d) Man bekommt offline-online-Prozeduren analog.

(e) Die Sampling Strategie muss auch τ_N *berücksichtigen.*

16 Allgemeine Nichtlinearitäten

Sei nun $\mathcal{L} : X \to Y$ ein allgemeiner nichtlinearer Operator und betrachte für $\mu \in \mathcal{D}$

$$u(\mu) \in X : \qquad \mathcal{L}\left(u(\mu); \mu\right) = f(\mu) \tag{16.1}$$

für $f(\mu) \in Y$. Existenz und Eindeutigkeit seien vorausgesetzt.

X, Y seien separable Hilbertäume mit Basen $\left\{\varphi_i^X\right\}_{i \in J^X}$ und $\left\{\psi_j^Y\right\}_{j \in J^Y}$, die normiert seien $\|\varphi_i^X\|_X = \|\psi_j^Y\|_Y = 1, \quad \forall i \in J^X, j \in J^Y, \quad J^X, J^Y \subset \mathbb{N}$.

Definition 16.1:
Für $u \in X$ habe $\mathcal{L}(u) \in Y$ die Darstellung

$$\mathcal{L}(u) = \sum_{j \in J^Y} \ell_j(u) \psi_j^Y.$$

Die Funktionale $\ell_j : X \to \mathbb{R}$ heißen Koeffizientenfunktionale.

Wir betrachten (natürlich) $\mathcal{L}(u; \mu)$ beziehungsweise $l_j(u; \mu), \quad \mu \in \mathcal{D}$.

Definition 16.2:
Für $M \in \mathbb{N}$ sei $T_M := \{j_1, \ldots, j_M\} \subset J^Y$ ("magic indices") und $Y_M \subset Y$ ein Raum der Dimension M mit Basis $\Phi^{(M)} := \left\{\phi_1^{Y,M}, \ldots, \phi_1^{Y,M}\right\}$ und Entwicklung

$$\phi_m^{Y,M} = \sum_{j \in J^Y} \left(\phi_m^{Y,M}\right)_j \psi_j^Y, \qquad \left(\phi_m^{Y,M}\right)_{j_n} = \delta_{m,n}, \qquad 1 \le m, n \le M,$$

dann definiere den interpolierenden Operator $\mathcal{I}_M\left(\mathcal{L}\right) : X \to Y_M$ durch

$$\mathcal{I}_M\left(\mathcal{L}\right)(v) := \sum_{m=1}^{M} \ell_{j_m}(v) \phi_m^{Y,M}. \tag{16.2}$$

Bemerkung 16.3:
Es gilt für

$$\begin{aligned}
\mathcal{I}_M\left(\mathcal{L}\right)(v) &= \sum_{m=1}^{M} \ell_{j_m}(v) \phi_m^{Y,M} = \sum_{m=1}^{M} \ell_{j_m}(v) \sum_{j \in J^Y} \left(\phi_m^{Y,M}\right)_j \psi_j^Y \\
&= \sum_{j \in J^Y} \left(\sum_{m=1}^{M} \ell_{j_m}(v) \left(\phi_m^{Y,M}\right)_j \right) \psi_j^Y = \sum_{j \in J^Y} \left(\mathcal{I}_M\left(\mathcal{L}\right)(v)\right)_j \psi_j^Y
\end{aligned}$$

$$\Rightarrow \left(\mathcal{I}_M\left(\mathcal{L}\right)(v)\right)_{j_n} = \sum_{m=1}^{M} \ell_{j_m}(v) = \underbrace{\left(\phi_m^{Y,M}\right)_{j_n}}_{=\delta_{m,n}} = \ell_{j_n}(v) = \left(\mathcal{L}(v)\right)_{j_n}$$

und daher der Name „interpolierender Operator“, an den „magic indices“ stimmen Koeffizientenfunktional und Interpolation überein.

Bemerkung 16.4:

(a) $\mathcal{I}_M$ *kann als EIM interpretiert werden, jedoch auf der Indexmenge* J^Y *(diskret), nicht auf* Ω *(kontinuierlich).*

(b) $\mathcal{I}_M : \{\mathcal{L} : X \to Y\} \to \{\mathcal{L}_M : X \to Y_M\}$ *ist linear, auch wenn die Argumente* $\mathcal{L}$ *nichtlinear sind.*

(c) Es gilt

$$\partial_v \mathcal{L}(v) = \partial_v \left(\sum_{j \in J^Y} \ell_j(v) \psi_j^Y \right) = \sum_{j \in J^Y} \underbrace{\left(\partial_v \ell_j(v)\right)}_{X \to X'} \psi_j^Y : X \to \mathcal{L}(X,Y)$$

$$\partial_v \left(\mathcal{I}_M\left(\mathcal{L}\right)(v)\right) = \sum_{m=1}^{M} \left(\partial_v \ell_{j_m}(v)\right) \psi_m^{Y,M} = \mathcal{I}_M\left(\partial_v \mathcal{L}\right)(v)$$

(d) J^X, J^Y *können als DOF („Degrees Of Freedom - Freiheitsgrade“) der truth angesehen werden, also* $|J^X| = \mathcal{N}^X, |J^Y| = \mathcal{N}^Y < \infty$*, aber* $\gg 1$*.*

Die Basis-Konstruktion erfolgt dann wie bei der EIM:

Algorithmus 16.5 (DOF-EIM)**:**

Gegeben seien Trainings-Snapshots $G^{train} := \{\mathcal{L}\left(u(\mu);\mu\right) \mid \mu \in \mathcal{D}_{train}\} \subset Y$ *und* $\varepsilon_{tol} > 0$ $Y_0 := \{0\}, \quad M := 0, \quad T_0 := \emptyset, \quad Q_0 := \emptyset, \quad \Phi^{(0)} := \emptyset$

while $\varepsilon_M := \max_{w \in G^{train}} \|w - \sum_{m=1}^{M} w_{j_m} \phi_m^{Y,M}\|_\infty > \varepsilon_{tol}$ **do**

  $w^* := arg\max_{w \in G^{train}} \|w - \sum_{m=1}^{M} w_{j_m} \phi_m^{Y,M}\|_\infty$

  $r_{M+1} := w^* - \sum_{m=1}^{M} w_{j_m}^* \phi_m^{Y,M}$

  $j_{M+1} := arg\max_{j \in J^Y} \left|\left(r_{M+1}\right)_j\right|, \qquad T_{M+1} := T_M \cup \{j_{M+1}\}$

  $q_{M+1} := \frac{r_{M+1}}{\left(r_{M+1}\right)_{j_{M+1}}}, \qquad Q_{M+1} := Q_M \cup \{q_{M+1}\}$

  $Y_{M+1} := spanQ_{M+1}, \qquad \Phi^{M+1} := nodale\ Basis\ zu\ T_{M+1}$

  $M = M + 1$

end

Bemerkung 16.6:

(a) Es gilt $Q_M \subset Q_{M+1}$*, aber* $\Phi_M \not\subset \Phi_{M+1}$*.*

(b) $\|q_{M+1}\|_\infty = 1$ *und* $B_M := \left[\left(q_j\right)_{j_m}\right]_{j,m=1,\ldots,m} \in \mathbb{R}^{M \times M}$ *ist eine untere Dreiecks-Matrix mit Diagonale* 1*. Es gilt* $\Phi^{(M)} = Q_M \cdot \left(B_M\right)^{-1}$*.*

Ziel: Effiziente Auswertung der Funktionale $\ell_j(u)$ und die Ableitungen hiervon. Wegen $u = \sum_{j \in J^X} u_i \varphi_i^X, \quad |J^X| = \mathcal{N}^X \gg 1$ ist dies absolut nicht trivial.

Idee dazu: Differenzialoperatoren sind lokal, wenn die Basis Φ^X auch lokal ist, braucht man eventuell nur „wenig benachbarte“ DOFs zu berücksichtigen.

Definition 16.7:
Für $i \in J^Y$ sei $\overline{N}(i) \subset J^X$ maximal, d.h.

$$\ell_i(u+v) = \ell_i(u), \qquad \forall u \in X, v \in span\left(\varphi_j^X\right)_{j\in\overline{N}(i)}. \tag{16.3}$$

Sei $N(i) := J^X \setminus \overline{N}(i)$ und sei $C \in \mathbb{N}$ minimal mit

$$|N(i)| \leq C \ll dim X$$

Die Menge

$$N_M := \bigcup_{i\in T_M} N(i)$$

heißt Menge der lokalen DOF-Indizes und

$$R_M : X \to span\left(\varphi_j^X\right)_{j\in N_M}, \qquad R_M(u) := \sum_{i\in N_M} u_i\varphi_i^X, \qquad u \in X,$$

Restriktionsoperator.

Bemerkung 16.8:
(a) Die DOFS aus $\overline{N}(i)$ beeinflussen $\ell_i(u)$ also nicht, $\overline{N}(i)$ ist die Menge der „aktiven Indizes“ von denen $\ell_i(u)$ abhängt.
(b) Sinn der Definition: Es reicht die Kenntnis von u_j für $j \in N(i)$ um $\ell_i(u)$ auszuwerten ($i \in J_j^Y$).

$$\ell_i(u) = \ell_i\left(\sum_{j\in N(i)} + \sum_{j\in\overline{N}(i)} u_j\varphi_j^X\right) \overset{(16.3)}{=} \ell_i\left(\sum_{j\in N(i)} u_j\varphi_j^X\right)$$

(c) Für $m = 1,\dots,M$ und $u \in X$ gilt $j_m \in T_M$, also $N(j_m) \subset N_M$. Damit gilt:

$$\ell_{j_m}(u) = \ell_{j_m}\left(\sum_{j\in N(j_m)} u_j\varphi_j^X\right) = \ell_{j_m}\left(\sum_{j\in N_M} u_j\varphi_j^X\right) = \ell_{j_m}\left(R_M(u)\right),$$

also eine lokale DOF-Abhängigkeit für ℓ_{j_m}. Die Kenntnis von $(u_j)_{j\in N_M}$ reicht also aus um alle $\ell_{j_m}(u)$, $m = 1,\dots,M$ auswerten zu können.
(d) Wegen $|N_M| \leq |T_M| \max_{m=1,\dots,M} |N(i_m)| \leq CM$ und C, M sind als klein angenommen, braucht man nur wenige Koeffizienten $\Rightarrow$ schnelle Auswertung
(e) Die „lokale DOF-Abhängigkeit“ ist eine Eigenschaft
- *des Operators und*
- *der Basis.*

Operatoren:	Basen:
$\circ \left\vert \left\{ a(\varphi_i^X, \varphi_j^X) \neq 0 \mid j \in J^X \right\} \right\vert = \mathcal{O}(1)$ für festes i	$\circ supp\, \varphi_i^X$ kompakt und
+ (lineare) Differenzialoperatoren	+ FEM + FV + FD + Wavelets
- Integraloperatoren (i.A.)	-Spektralmethoden

Lemma 16.9:
Für alle $v, w \in X$ und $1 \leq m \leq M$ gilt

(i) $\ell'_{j_m}(v) = \ell'_{j_m}(R_m v) \in X'$.
(ii) $\left\langle \ell'_{j_m}(v), w \right\rangle \in \left\langle \ell'_{j_m}(v), R_m w \right\rangle \in \mathbb{R}$

Beweis. Wegen $\ell_j : X \to \mathbb{R}$ ist $\ell'_j(v) \in X'$ für $v \in X$, also

$$\left\langle \ell'_{j_m}(v), w \right\rangle = \sum_{i \in J^X} r_i w_i \in \mathbb{R}$$

mit $(r_i)_{i \in J^X}, r_i \in \mathbb{R}$ mit $w = \sum_{i \in J^X} w_i \varphi_i^X \in X$. Setze $w = \varphi_i^X$ ein:

$$r_i = \left\langle \ell'_{j_m}(v), \varphi_i^X \right\rangle = \lim_{h \to 0} \frac{1}{h} \left\{ \ell_{j_m} \left(v + h \frac{\varphi_i^X}{\|\varphi_i^X\|} \right) - \ell_{j_m}(v) \right\}$$

Für $i \in \overline{N}(j_m)$ gilt nach (16.3): $\ell_{j_m}(v + \alpha \varphi_i^X) = \ell_{j_m}(v)$ für alle α, also $r_i = 0$ für alle $i \in \overline{N}(j_m)$. Daraus folgt

$$\left\langle \ell'_{j_m}(v), w \right\rangle = \sum_{i \in N(j_m)} r_i w_i = \left\langle \ell'_{j_m}(v), R_M w \right\rangle,$$

da $R_M w = \sum_{j \in N_M} w_j \varphi_j^X$ und $N(j_m) \subset N_M$. □

Satz 16.10 (Fehlerschätzer)**:**
Falls $\mathcal{L}(u) \in Y_{M'}$ für $M' > M$. Dann gilt mit $\underline{\alpha} := \underline{Q}^{-1} \underline{\ell}$, $\underline{Q} := [(q_j)_i]_{i,j=M+1,\ldots,M'}$, $\underline{\ell} := (\mathcal{L}(u) - \mathcal{I}_M(\mathcal{L})(u))_{i=M+1,\ldots,M'}$, $\underline{K}_Q := [(q_i, q_j)_Y]_{i,j=M+1,\ldots,M'}$

(a) $\|\mathcal{L}(u) - \mathcal{I}_M(\mathcal{L})(u)\|_\infty \leq \|\underline{\alpha}\|_1$
(b) $\|\mathcal{L}(u) - \mathcal{I}_M(\mathcal{L})(u)\|_Y \leq \left(\underline{\alpha}^T \underline{K}_Q \underline{\alpha}\right)^{1/2}$

Satz 16.11 (Erhaltungseigenschaft)**:**
Sei $g \in Y'$ mit $g\left(\mathcal{L}(u(\mu); \mu)\right) = 0 \quad \forall \mu \in \mathcal{D}$. Dann ist $g\left(\mathcal{I}_M\left(\mathcal{L}(u(\mu); \mu)\right)\right) = 0$.

Ausblicke

Einige aktuelle Fragestellungen:

- RB und Adaptivität
  + truth
  + Erweiterung der Trainingsmenge (Optimierung)
  + Online-Auswahl von „guten“ snapshots
  + hp-Methoden in $\mathcal{D}$ (adaptive Parameter-Partition)
- Hyperbolische Probleme
- Hochdimensionalität
  + X
  + $\mathcal{D}$ - Parameterfunktionen
  + Stochastische Einflüsse
- Stark gekoppelte Probleme
  + LEGO bei einfachen Kopplungen (+ Maxwell, Schrödinger,...)
  + Multiphysics?
- Turbulenz, hohe Reynolds-Zahlen
- „Echte“ Anwendungen
- ...