

Vorlesungsmitschrift

# REDUZIERTE BASIS METHODEN

UNIVERSITÄT STUTTGART, SS15  
Prof. Dr. Bernard Haasdonk

AUTOREN:  
Stefan Simeonov  
Frank Schneider

STAND:  
7. Mai 2015

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>2</b>
1.1	Modellreduktion . . . . .	3
<b>2</b>	<b>Grundlagen</b>	<b>8</b>
<b>3</b>	<b>RB-Methoden für lineare koerzive Probleme</b>	<b>16</b>
3.1	Primales RB-Problem . . . . .	16
3.2	Fehleranalyse . . . . .	19

# 1 Einleitung

## Parameterabhängige Probleme

### Beispiel 1.1 (Parameterabhängige PDE)

Sie  $\Omega \subseteq \mathbb{R}^d$  polygonales Gebiet. Zu Parametervektor  $\mu \in \mathcal{P} \subset \mathbb{R}^p$  aus einer Menge  $\mathcal{P}$  von „erlaubten“ Parametern ist eine Funktion (z. B. „Temperatur“)  $u(\mu) : \Omega \rightarrow \mathbb{R}$ , s. d.:

$$\begin{aligned} \nabla \cdot (\kappa(\mu) \nabla u) &= q(\mu) && \text{in } \Omega \\ u(\mu) &= 0 && \text{auf } \delta\Omega \end{aligned}$$

mit  $\kappa(\mu) : \Omega \rightarrow \mathbb{R}$  (z. B. „Wärmeleitungskoeffizient“)

und  $q(\mu) : \Omega \rightarrow \mathbb{R}$  (z. B. „Wärmequelle/-senke“)

$$\text{z. B. } q(x; \mu) := \begin{cases} 1 & \text{für } x \in \Omega_q \\ 0 & \text{sonst} \end{cases}$$

Weiter kann Ausgabe erwünscht, z. B. mittlere Temperatur

$$s(\mu) = \frac{1}{|\Omega_s|} \int u(x; \mu) dx$$

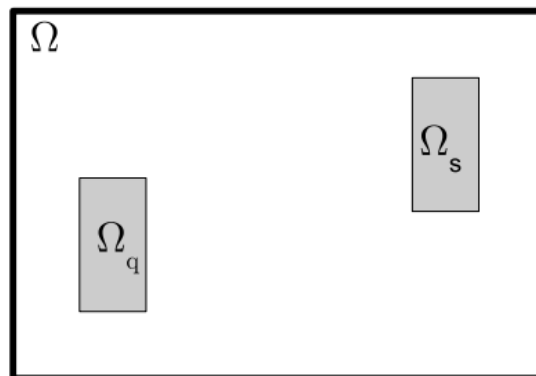


Abbildung 1: Beispiel Wärmeleitung mit Quelle  $\Omega_q$  und Messbereich  $\Omega_s$   
(aus B. Haasdonk, Reduzierte-Basis-Methoden, Skript zur Vorlesung SS 2011, Universität Stuttgart, IANS-Report 4/11, 2011.)

### Beispiel 1.2 (Parametrisches stationäres System)

Zu Parameter  $\mu \in \mathcal{P} \subseteq \mathbb{R}^p$  ist Zustandsvektor  $u(\mu) \in \mathbb{R}^n$  und Ausgabe  $s(\mu) \in \mathbb{R}^k$  gesucht, s. d.:

$$\begin{aligned} 0 &= A(\mu) \cdot u(\mu) + B(\mu)w(\mu) \\ s(\mu) &= l(\mu) \cdot u(\mu) \end{aligned}$$

mit parameterabhängigen Matrizen  $A(\mu) \in \mathbb{R}^{n \times n}$ ,  $B(\mu) \in \mathbb{R}^{n \times m}$ ,  $C(\mu) \in \mathbb{R}^{k \times n}$  mit Eingabevektor  $w \in \mathbb{R}^m$ .

## Schwache Formulierung in Hilberträumen

Sie  $X$  reeller Hilbertraum (reel, seperabel). Zu  $\mu \in \mathcal{P}$  ist gesucht ein  $u(\mu) \in X$  und  $s(\mu) \in \mathbb{R}$

$$\begin{aligned} a(u(\mu), v; \mu) &= f(v; \mu) \\ s(\mu) &= l(u(\mu); \mu) \end{aligned} \quad \forall v \in X$$

Mit Bilinearform  $a(\cdot, \cdot; \mu)$  und Linearform  $f(\cdot; \mu)$ ,  $l(\cdot; \mu)$ . Beide Beispiele lassen sich so formulieren.

z. B. 1.1:

$$\begin{aligned} X &= H_0^1(\Omega) := \{f \in L^2(\Omega) \mid \frac{\partial}{\partial x_i} f \in L^2(\Omega), f|_{\partial\Omega} = 0\} \\ \underbrace{\int_{\Omega} \kappa(x; \mu) \nabla u(x; \mu) \cdot \nabla v(x) dx}_{a(u(\mu), v; \mu)} &= \underbrace{\int_{\Omega} q(x; \mu) \cdot v(x) dx}_{f(v; \mu)} \quad \forall v \in X \\ s(\mu) &= \frac{1}{|\Omega_s|} \int_{\Omega_s} u(x; \mu) =: l(u(\mu); \mu) \end{aligned}$$

Zu Bsp. 1.2 ( $k = 1$ , „single output“)  $X = \mathbb{R}^n$

$$\begin{aligned} \underbrace{v^T A(\mu) u(\mu)}_{a(u(\mu), v; \mu)} &= \underbrace{-v^T B w}_{f(v; \mu)} \\ s(\mu) &:= \underbrace{C(\mu) u(\mu)}_{l(u(\mu); \mu)} \end{aligned}$$

In der Vorlesung werden weitere Verallgemeinerungen zu  $a : X_1 \times X_2 \rightarrow \mathbb{R}$  mit  $X_1 \neq X_2$ , nichtlinear und instationäre Probleme behandelt.

## 1.1 Modellreduktion

### Grundidee/Motivation

- $\mathcal{M} := \{u(\mu) \mid \mu \in \mathcal{P}\} \subset X$  für  $\mathcal{P} \subseteq \mathbb{R}^p$  ist die durch  $\mu$  parametrisierte Lösungsmanigfaltigkeit.
- $X$  ist im allgemeinen  $\infty$ -dimensional Sobolev-Raum) oder endlich- aber sehr hoch-dimensional (FEM, FV, FD-Raum).  $\mathcal{M}$  ist aber höchstens  $p$ -dimensional.  
 $\Rightarrow$  Motivation für Suche nach einem niedrigdimensionalen Teilraum  $X_n \subseteq X$  zur Approximation von  $\mathcal{M}$  und einer Approximation  $u_N(\mu) \approx u(\mu)$ ,  $u_N(\mu) \in X_N$
- Insbesondere bei Reduzierten-Basis-Methoden (RB-Methoden):  
 $X_N$  durch Beispiellösungen erzeugt, sog. „Snapshots“  
 $X_N \subseteq \text{span}\{u(\mu_1), \dots, u(\mu_n)\}$  für geeignete Parameterwerte  $\mu_i \in \mathcal{P}$ .  
Ziel ist außerdem Fehlerkontrolle durch Schranken  $\Delta_N(\mu)$ :

$$\|u(\mu) - u_N(\mu)\| \leq \Delta_N(\mu)$$

## Illustration

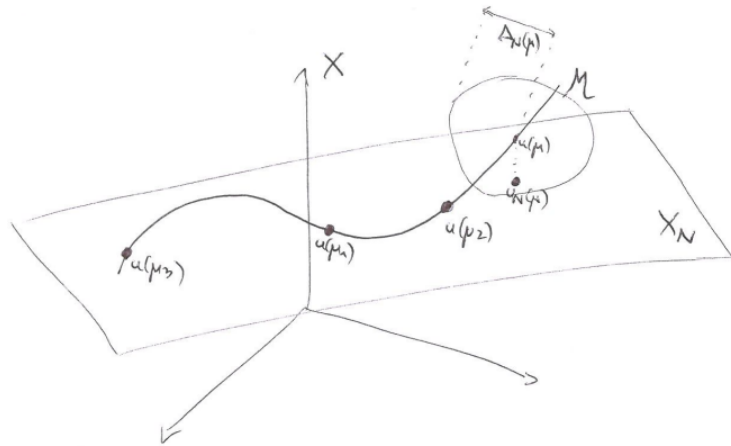


Abbildung 2: Parametrisierte niedrigdimensionale Lösungsmenge  
(aus dem Online-Skript von Prof. Dr. Haasdonk zu Reduzierte Basen 2015)

### Beispiel 1.3

Gesucht ist  $u(\mu) \in C^2([0,1])$  mit

$$\begin{aligned} (1 + \mu)u'' &= 1 & \text{auf } (0,1) \\ u(0) &= u(1) = 1 \end{aligned}$$

Für  $\mu \in [0,1] =: \mathcal{P} \subseteq \mathbb{R}$ . Spezielle Lösungen („Snapshots“)

$$\begin{aligned} \mu = 0 &\Rightarrow u_0(x) = u(x; \mu = 0) = \frac{1}{2}x - \frac{1}{2}x + 1 \\ \mu = 1 &\Rightarrow u_1(x) = u(x; \mu = 1) = \frac{1}{4}x - \frac{1}{4}x + 1 \end{aligned}$$

RB-Raum:  $X_N := \text{span}(u_0, u_1)$  Reduzierte Lösung gegeben durch

$$\begin{aligned} u_N(\mu) &:= \alpha_0(\mu)u_0 + \alpha_1(\mu)u_1 \\ \alpha_0(\mu) &= \frac{2}{\mu + 1} - 1; \quad \alpha_1(\mu) = 2 - \frac{2}{\mu + 1} \end{aligned}$$

Diese erfüllt

$$\|u_N(\mu) - u(\mu)\|_\infty = \sup_\mu \|u(\mu) - u_N(\mu)\| = 0$$

ist somit exakt.  $\mathcal{M}$  ist enthalten in 2-dimensionalem Unterraum  $X_N$ : Genauer  $\alpha_0 + \alpha_1 = 1, 0 \leq \alpha_0, \alpha_1 \leq 1$ , also ist  $\mathcal{M}$  Menge der Konvexkombinationen von  $u_0, u_1$ .

## Begriffe

- Eine PDE ist ein *analytisches* Modell, welches die *exakte Lösung*  $u(\mu) \in X$  in einem typischerweise  $\infty$ -dimensionalen Funktionenraum  $X$  charakterisiert.
- Ein *detailliertes Modell* (auch *hochdimensionales Modell*) ist ein Berechnungsverfahren oder charakterisiert eine Approximation  $u(\mu) \in X$  in hochdimensionalen Raum mit sehr allgemeinen Approximationseigenschaften. (z.B. FEM/FV/FD,  $\dim X = 10^3 - 10^8$ ). In dieser Vorlesung kann  $u(\mu)$  sowohl eine analytische als auch eine detaillierte Lösung darstellen.
- Ein *reduziertes* Modell ist ein Berechnungsverfahren bzw. eine Charakterisierung einer reduzierten Lösung  $u_N(u)$  in einem sehr problemangepassten Raum  $X_N$  ( $\dim X_N = 1 - 10^3$ ).
- *Modellreduktion* beschäftigt sich mit Methoden der Erzeugung reduzierter Modelle und Untersuchung ihrer Eigenschaften
- Modellreduktion ist ein modernes Gebiet der angewandten Mathematik und Ingenieurwissenschaften (Schwerpunkt in SimTech PN3, MOR-Seminar)

## Anwendungen für parametrische reduzierte Modelle

„Kleinere“ Modelle stellen geringere Anforderungen an Rechenzeit und Speicher, daher Einsatz in:

- „multi-query“-Kontext, d. h. Vielfachanfragen unter Parametervariation: Parameterstudien, Design, Parameteridentifikation, Inverse Probleme, Optimierung, statistische Analyse
- Multi-skalen-Modelle (reduzierte Mikrolöser)
- „real-time“-Kontext, d. h. Anwendungen mit schneller Simulationsantwort: Interaktive Benutzeroberfläche, Web-Formulare, Echtzeitsteuerung von Prozessen
- „cool-computing“-Kontext, d. h. Simulation auf „einfacher“ Hardware: elektronische Regler, Smartphones, Ubiquitous Computing

## Demonstration

demo\_thermalblock.m aus RBmatlab, Smartphone App JaRMoS

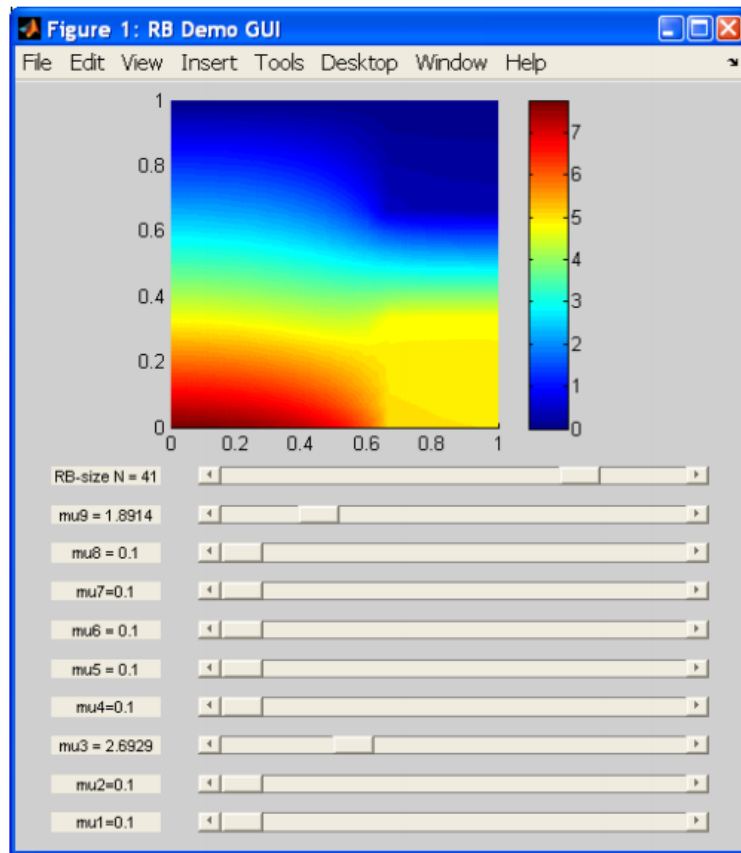


Abbildung 3: Beispiel des Thermischen Blocks aus `demo_thermalblock.m`  
 (aus B. Haasdonk, Reduzierte-Basis-Methoden, Skript zur Vorlesung SS 2011, Universität  
 Stuttgart, IANS-Report 4/11, 2011.)

### Offline/Online Zerlegung

Typischerweise wird eine Verechnungsintensive Generierung des reduzierten Modells akzeptiert, sog. *Offline-Phase*. Dies ermöglicht schnelle Anwendbarkeit des reduzierten Modells in der *Online-Phase*. Offline-Kosten werden gerechtfertigt durch Amortisierung im multi-query-Kontext, d. h. Laufzeitgewinn bei genügend großer Anzahl an Online-Simulationen

multi-query mit detailliertem Modell:



multi-query mit reduziertem Modell:

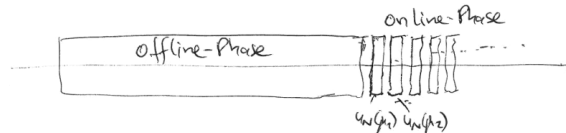


Abbildung 4: Laufzeitvergleich eines detaillierten mit einem reduzierten Modell  
(aus dem Online-Skript von Prof. Dr. Haasdonk zu Reduzierte Basen 2015)

### Zentrale Fragen

- Reduzierte Basis: Wie kann ein möglichst kompakter Teilraum konstruiert werden? Können solche Verfahren *beweisbar* gut sein?
- Reduziertes Modell: Wie kann eine Lösung  $u_N(\mu) \in X_N$  bestimmt werden
- Berechnungs-Effizienz: Wie kann  $u_N(\mu)$  *schnell* berechnet werden?
- Stabilität: Wie kann Stabilität des reduzierten Modells garantiert werden bei wachsendem  $N := \dim X_N$ ?
- Fehlerschätzer: Kann der Fehler des reduzierten zum detaillierten oder analytischen Modells beschränkt werden? Sind die Fehlerschätzer schnell berechenbar?
- Effektivität der Fehlerschätzer: Kann garantiert werden, dass der Schätzer den Fehler nicht zu pessimistisch überschätzt?
- Für welche Problemklassen kann ein RB-Ansatz funktionieren, für welche nicht?

### Vorläufige Gliederung

- 1 Einleitung
- 2 Grundlagen
- 3 RB Verfahren für lineare koerzive Probleme
- 4 Allgemeinere lineare Probleme
- 5 Nichtlineare Probleme
- 6 Instationäre Probleme
- 7 Weiterführende Aspekte

## 2 Grundlagen

Im Folgenden sei  $X$  (oder  $X_1, X_2$ ) stets reeller Hilbertraum mit Skalarprodukt  $\langle \cdot, \cdot \rangle_X$ , Norm  $\| \cdot \|_X$  und Dualraum  $X'$ . Subskript wird weggelassen falls keine Verwechslungsgefahr besteht.

**Definition 2.1** (Parametrische Formen)

Sei  $\mathcal{P} \subset \mathbb{R}^p$  beschränkte Parametermenge. Dann nennen wir

i)  $l : X \times \mathcal{P} \rightarrow \mathbb{R}$  *parametrische stetige Linearform* falls  $\forall \mu \in \mathcal{P}$ :

$$l(\cdot; \mu) \in X'$$

ii)  $a : X_1 \times X_2 \times \mathcal{P} \rightarrow \mathbb{R}$  eine *parametrische stetige* (symmetrische) *Bilinearform*, falls für alle  $\mu \in \mathcal{P}$

$$a(\cdot, \cdot; \mu) : X_1 \times X_2 \rightarrow \mathbb{R} \quad \text{ist bilinear und stetig (symmetrisch)}$$

Wir bezeichnen die Stetigkeitskonstante mit

$$\gamma(\mu) := \sup_{u \in X_1} \sup_{v \in X_2} \frac{a(u, v; \mu)}{\|u\|_{X_1} \|v\|_{X_2}}$$

Falls  $X_1 = X_2 =: X$  und  $a(\cdot, \cdot; \mu)$  ist koerziv für alle  $\mu \in \mathcal{P}$ , so ist  $a(\cdot, \cdot; \cdot)$  *parametrisch koerziv* und wir bezeichnen die Koerzitivitätskonstante mit

$$\alpha(\mu) := \inf_{u \in X} \frac{a(u, u; \mu)}{\|u\|^2}$$

**Bemerkung.** Eine parametrische stetige Bi-/Linearform ist nicht unbedingt stetig bzgl.  $\mu$ . Beispiel:  $X = \mathbb{R}$ ,  $\mathcal{P} = [0, 1]$ ,  $l : X \times \mathcal{P} \rightarrow \mathbb{R}$  definiert durch

$$l(x; \mu) := \begin{cases} x & \text{falls } \mu < \frac{1}{2} \\ \frac{1}{2}x & \text{sonst} \end{cases}$$

**Definition 2.2** (Parametrische Beschränktheit / Lipschitz-Stetigkeit / Koerzitivität)

Wir nennen

i) eine parametrische stetige Linearform  $l$  bzw. Bilinearform  $a$  *gleichmäßig beschränkt* bzgl.  $\mu$  falls ex.  $\bar{\gamma}_l, \bar{\gamma} < \infty$  mit

$$\sup_{\mu \in \mathcal{P}} \|l(\cdot; \mu)\|_{X'} \leq \bar{\gamma}_l \quad \text{bzw.} \quad \sup_{\mu \in \mathcal{P}} \gamma(\mu) \leq \bar{\gamma}$$

ii)  $a$  *gleichmäßig koerziv* bzgl.  $\mu$  falls ex.  $\bar{\alpha} > 0$  mit

$$\inf_{\mu \in \mathcal{P}} \alpha(\mu) \geq \bar{\alpha}$$



iii)  $l$  bzw.  $a$  *Lipschitz-stetig* bzgl.  $\mu$  falls ex.  $L_l$  bzw.  $L_a \in \mathbb{R}^+$ , sodass  $\forall \mu_1, \mu_2 \in \mathcal{P}$  gilt

$$|l(u; \mu_1) - l(u; \mu_2)| \leq L_l \|u\| \|\mu_1 - \mu_2\| \quad \forall u \in X$$

bzw.

$$|a(u, v; \mu_1) - a(u, v; \mu_2)| \leq L_a \|u\| \|v\| \|\mu_1 - \mu_2\| \quad \forall u \in X_1, v \in X_2$$

**Definition 2.3** (Sensitivitätsableitung)

Sei  $\mu_0 \in \mathcal{U} \subset \mathcal{P}$  in Umgebung  $\mathcal{U}$  von  $\mu_0$ . Wir nennen  $f : \mathcal{U} \rightarrow X$  (Frechet)-differenzierbar in  $\mu_0$ , falls ex. ein  $Df(\mu_0) \in L(\mathbb{R}^p, X)$  mit

$$\lim_{h \rightarrow 0} \frac{\|f(\mu_0 + h) - f(\mu_0) - Df(\mu_0)h\|}{\|h\|} = 0$$

Falls  $f$  in jedem  $\mu \in \mathcal{U}$  diffbar, dann existieren insbesondere partielle Ableitungen

$$\frac{\partial}{\partial \mu_i} f(\cdot) := Df(\cdot) e_i : \mathcal{U} \rightarrow X$$

für  $e_i \in \mathbb{R}^p$  Einheitsvektor  $i = 1, \dots, p$ . Falls diese wiederum diffbar in  $\mathcal{U}$  bezeichnet allgemein

$$\partial_\sigma f(\cdot) := \frac{\partial^{|\sigma|}}{\partial \mu_1^{\sigma_1} \dots \partial \mu_p^{\sigma_p}} f(\cdot) : \mathcal{U} \rightarrow X$$

die Sensitivitätsableitung der Ordnung  $|\sigma| := \sum_{i=1}^p \sigma_i$  für Multiindex  $\sigma = (\sigma_i)_{i=1}^p \in \mathbb{N}_0^p$ .

**Bemerkung.** Diese Ableitungen werden später insbesondere bei parameterabhängigen Lösungen  $u(x; \mu)$  verwendet:

$u : \Omega \times \mathcal{P} \rightarrow \mathbb{R}$  mit  $u(\cdot; \mu) \in X$  kann auch als

$u : \mathcal{P} \rightarrow X$  aufgefasst werden mit Sensitivitätsableitungen

$\partial_\sigma u : \mathcal{P} \rightarrow X$ , d.h.  $\partial_\sigma u(\cdot; \mu) \in X \quad \forall \mu \in \mathcal{P}$  und insbesondere

$\partial_\sigma u : \Omega \times \mathcal{P} \rightarrow \mathbb{R}$ , d.h.  $\partial_\sigma$  sind wieder Funktionen auf  $\Omega$

**Definition 2.4** (Separierbare Parameterabhängigkeit)

- i) Eine Funktion  $v : \mathcal{P} \rightarrow X$  nennen wir *separierbar parametrisch*, falls existieren Komponenten  $v^q \in X$  und Koeffizientenfunktionen  $\Theta_v^q : \mathcal{P} \rightarrow \mathbb{R}$  für  $q = 1, \dots, Q_v$  mit

$$v(\mu) = \sum_{q=1}^{Q_v} \Theta_v^q(\mu) v^q$$

- ii) Eine parametrische stetige Linearform  $l : X \times \mathcal{P} \rightarrow \mathbb{R}$  bzw. Bilinearform  $a : X_1 \times X_2 \times \mathcal{P} \rightarrow \mathbb{R}$  ist separierbar parametrisch, falls existieren  $l^q \in X'$  und  $\Theta_l^q : \mathcal{P} \rightarrow \mathbb{R}$  für  $q = 1, \dots, Q_l$  bzw.  $a^q : X_1 \times X_2 \rightarrow \mathbb{R}$  stetig, bilinear und  $\Theta_a^q : \mathcal{P} \rightarrow \mathbb{R}$  für  $q = 1, \dots, Q_a$  mit

$$l(v; \mu) = \sum_{q=1}^{Q_l} \Theta_l^q(\mu) l^q(v) \quad \forall v \in X, \mu \in \mathcal{P}$$

$$a(u, v; \mu) = \sum_{q=1}^{Q_a} \Theta_a^q(\mu) a^q(u, v) \quad \forall u \in X_1, v \in X_2, \mu \in \mathcal{P}$$

**Bemerkung.**

- i) In Literatur auch “affine Annahme” oder “affin parametrisch” verwendet. Wir verwenden jedoch “separierbar”, da  $\Theta_l^q$  auch nichtlinear sein können.
- ii)  $Q_a, Q_l$  sollten möglichst klein sein, weil diese in die Online-Komplexität eingehen, siehe *Abschnitt 3*.

**Satz 2.5** (Energienorm)

Sei  $a : X \times X \times \mathcal{P} \rightarrow \mathbb{R}$  parametrische stetige, koerzive Bilinearform, und  $a_s(u, v; \mu) = \frac{1}{2}(a(u, v; \mu) + a(v, u; \mu))$  der symmetrische Anteil. Dann ist für  $\mu \in \mathcal{P}$

$$\langle u, v \rangle_\mu := a_s(u, v; \mu) \quad \text{bzw.} \quad \|u\|_\mu := \sqrt{\langle u, u \rangle_\mu}$$

das *Energie-Skalarprodukt* bzw. die *Energienorm* bzgl.  $\mu$ . Diese ist äquivalent zu  $\|\cdot\|_X$ :

$$\sqrt{\alpha(\mu)}\|u\| \leq \|u\|_\mu \leq \sqrt{\gamma(\mu)}\|u\|$$

*Beweis.* Skalarprodukt: klar wegen Bilinearität, Stetigkeit und Koerzivität. Normäquivalenz folgt aus Stetigkeit und Koerzivität von  $a_s$ .

$$\alpha(\mu)\|u\|^2 \leq \underbrace{a(u, u; \mu)}_{\leq \|u\|^2 \gamma(\mu)} = a_s(u, u; \mu) = \|u\|_\mu^2$$

□

**Satz 2.6** (Übertragung von Koeffizienten-Eigenschaften)

Seien  $f$  bzw.  $a$  separierbar parametrische stetige Linear- bzw. Bilinearform.

- i) Falls  $\Theta_f^q(\mu)$  bzw.  $\Theta_a^q(\mu)$  beschränkt sind, dann sind  $f$  bzw.  $a$  gleichmäßig beschränkt bzgl.  $\mu$ .
- ii) Falls  $\Theta_a^q(\mu)$  strikt positiv, d.h. ex.  $\bar{\Theta}$  mit  $\Theta_a^q(\mu) \geq \bar{\Theta} > 0 \forall \mu \in \mathcal{P}$  alle Komponenten positiv semidefinit, d.h.  $a^q(v, v) \geq 0 \forall v, q$  und  $a(\cdot, \cdot; \bar{\mu})$  ist koerziv für mindestens ein  $\bar{\mu} \in \mathcal{P}$ , dann ist  $a$  gleichmäßig koerziv bzgl.  $\mu$ .

iii) Falls  $\Theta_f^q, \Theta_a^q$  Lipschitz-stetig, so ist  $f, a$  Lipschitz-stetig bzgl.  $\mu$ .

*Beweis.*

i) Sei  $\bar{\Theta}_f^q \in \mathbb{R}^+$  mit  $|\Theta_f^q(\mu)| \leq \bar{\Theta}_f^q \forall \mu$ . Dann gilt

$$\|f(\cdot; \mu)\| = \left\| \sum_q \Theta_f^q(\mu) f^q \right\| \leq \sum_q |\Theta_f^q(\mu)| \|f^q\| \leq \sum_q \bar{\Theta}_f^q \|f^q\| =: \bar{\gamma}_f < \infty$$

analog für  $a(\cdot, \cdot; \mu)$ .

ii) Für  $u \in X, \mu \in \mathcal{P}$  gilt

$$a(u, u; \mu) = \sum_q \Theta_a^q(\mu) a^q(u, u) = \sum_q \underbrace{\frac{\Theta_a^q(\mu)}{\Theta_a^q(\bar{\mu})}}_{>0} \underbrace{\Theta_a^q(\bar{\mu}) a^q(u, u)}_{\sum(\cdot) = a(u, u; \bar{\mu})} \geq \underbrace{\sum_q \frac{\bar{\Theta}}{\max_{q'} \Theta_a^{q'}(\bar{\mu})} \alpha(\bar{\mu})}_{=: \bar{\alpha} > 0} \|u\|^2$$

iii) Sei  $|\Theta_f^q(\mu_1) - \Theta_f^q(\mu_2)| \leq L_f^q |\mu_1 - \mu_2| \forall \mu_1, \mu_2 \in \mathcal{P}$  mit geeignetem  $L_f^q \in \mathbb{R}$ . Dann gilt

$$\begin{aligned} |f(v; \mu_1) - f(v; \mu_2)| &= \left| \sum_q \Theta_f^q(\mu_1) f^q(v) - \sum_q \Theta_f^q(\mu_2) f^q(v) \right| \\ &\leq \sum_q |\Theta_f^q(\mu_1) - \Theta_f^q(\mu_2)| \|f^q\| \|v\| \\ &\leq \underbrace{\sum_q L_f^q \|f^q\|}_{=: L_f} \|\mu_1 - \mu_2\| \|v\| \end{aligned}$$

analog für  $a(\cdot, \cdot; \mu)$ .

□

**Definition 2.7** (Volles Problem  $(P(\mu))$ )

Seien  $a$  bzw.  $f, l$  parametrische Bilinearform bzw. Linearform und gleichmäßig stetig bzgl.  $\mu$ , sei  $a$  gleichmäßig koerziv bzgl.  $\mu$ . Dann ist für  $\mu \in \mathcal{P}$  gesucht  $u(\mu) \in X$  und  $s(\mu) \in \mathbb{R}$  als Lösung von

$$\begin{aligned} a(u(\mu), v; \mu) &= f(v; \mu) & \forall v \in X \\ s(\mu) &:= l(u(\mu); \mu) \end{aligned}$$

**Bemerkung.**

- Das volle Problem kann also ein analytisches Modell (PDE) oder ein detailliertes Modell (PDE-Diskretisierung) darstellen.
- Symmetrie von  $a$  wird nicht vorausgesetzt.

- In ??, ?? werden Verallgemeinerungen von  $(P(\mu))$  betrachtet.

**Satz 2.8** (Wohlgestelltheit und Stabilität)

Das Problem  $(P(\mu))$  besitzt eine eindeutige Lösung mit

$$\|u(\mu)\| \leq \frac{\|f(\mu)\|_{X'}}{\alpha(\mu)} \leq \frac{\bar{\gamma}_f}{\bar{\alpha}}, \quad |s(\mu)| \leq \|l(\mu)\|_{X'} \|u(\mu)\| \leq \frac{\bar{\gamma}_l \bar{\gamma}_f}{\bar{\alpha}}$$

*Beweis.* Existenz, Eindeutigkeit und Schranke für  $u(\mu)$  folgen mit Lax-Milgram (siehe z.B. Satz 2.5 in Braess'03). Gleichmäßige Stetigkeit und Koerzivität ergeben  $\mu$ -unabhängige Schranke für  $u(\mu)$ . Definition von  $s(\mu)$  ergibt Eindeutigkeit und entsprechende Schranken.  $\square$

**Definition 2.9** (Lösungsmannigfaltigkeit)

Wir definieren

$$\mathcal{M} := \{u(\mu) \in X : \mu \in \mathcal{P} \text{ und } u(\mu) \text{ löst } (P(\mu))\}$$

**Bemerkung.** Wir verwenden den Begriff “Mannigfaltigkeit” nicht im strengen differentialgeometrischen Sinn, weil keine Stetigkeit / Diffbarkeit von  $\mathcal{M}$  gefordert wird.

**Beispiel 2.10** (Thermischer Block)

TODO

**Beispiel 2.11** (Matrixgleichung)

- Zu  $\mu \in \mathcal{P}$  suche  $u(\mu) \in \mathbb{R}^H$  als Lösung von

$$A(\mu)u(\mu) = b(\mu)$$

für  $A(\mu) \in \mathbb{R}^{H \times H}$  und  $b(\mu) \in \mathbb{R}^H$ .

- Dies ist Beispiel für  $(P(\mu))$  via

$$X := \mathbb{R}^H, \quad a(u, v; \mu) := v^\top A(\mu)u, \quad f(v; \mu) := v^\top b(\mu)$$

und beliebiger linearer Ausgabe  $l(v; \mu) := \underline{l}^\top v$  für  $\underline{l} \in \mathbb{R}^H$ .

**Beispiel 2.12** ( $Q_a = 1$ )

Falls  $a(\cdot, \cdot; \mu)$ ,  $f(\cdot; \mu)$  separierbar parametrisch mit  $Q_a = 1$  und  $Q_f$  beliebig, so ist  $\mathcal{M}$  enthalten in einem  $Q_f$ -dimensionalen linearen Teilraum von  $X$ :

$$(P(\mu)) \Rightarrow \Theta_a^1(\mu) a^1(u, v) = \sum_q \Theta_f^q(\mu) f^q(v) \quad \forall v \in X$$

$\Theta_a^1(\mu) \neq 0$  wegen  $a$  gleichmäßig koerziv

$$a^1(u, v) = \sum_q \frac{\Theta_f^q(\mu)}{\Theta_a^1(\mu)} f^q(v) \quad \forall v \in X \quad (*)$$

$a^1(\cdot, \cdot)$  ist koerziv,  $f^q$  linear und stetig

$$\begin{aligned} & \xRightarrow{\text{Lax-Milgram}} \text{ex. } u^q, q = 1, \dots, Q_f \text{ mit } a^1(u^q, v) = f^q(v), \quad v \in X \\ & \Rightarrow u := \sum_q \frac{\Theta_f^q(\mu)}{\Theta_a^1(\mu)} u^q \text{ löst } (*) \text{ wegen Linearität} \\ & \Rightarrow u \in \text{span}\{u^q\}_{q=1}^{Q_f} \end{aligned}$$

**Beispiel 2.13** ( $(P(\mu))$  mit vorgegebener Lösung)

Sei  $u : \mathcal{P} \rightarrow X$  beliebig komplizierte Abbildung. Dann existiert ein  $(P(\mu))$  mit  $u(\mu)$  als Lösung via Skalarprodukten:

$$a(v, w; \mu) := \langle w, v \rangle_X, \quad f(v; \mu) := \langle u(\mu), v \rangle_X$$

d.h. Klasse der Probleme  $(P(\mu))$  können beliebig komplizierte, nichtglatte oder sogar unstetige Lösungsmannigfaltigkeit  $\mathcal{M}$  besitzen.

**Bemerkung** (Parameter-Anzahl und Lösungskomplexität). Es gibt (sogar in der Literatur) ein Missverständnis zwischen Parameteranzahl  $p \in \mathbb{N}$  und Komplexität der Lösungsmannigfaltigkeit  $\mathcal{M}$ , denn es kann Redundanz in Parametern vorliegen (siehe Thermischer Block). Extremfall:  $p \in \mathbb{N}$  beliebig, für geeignetes  $a(\cdot, \cdot; \mu)$ ,  $f(\mu)$  hat  $(P(\mu))$  ein  $\mathcal{M}$ , welches in einem 1D-Raum enthalten ist. (Übung) Beispiel 2.13 zeigt andererseits einen anderen Extremfall: Sogar für  $p = 1$  kann bei geeignetem  $(P(\mu))$  das  $\mathcal{M}$  beliebig kompliziert sein (z.B. "Raumfüllende Kurve"). Unter geeigneter Annahmen an  $a(\cdot, \cdot; \mu)$  und  $f(\cdot; \mu)$  können einfache Regularitätseigenschaften von  $u(\mu)$  bzw.  $\mathcal{M}$  geschlossen werden.

**Korollar 2.14** (Beschränktheit von  $\mathcal{M}$ )

Weil  $a(\cdot, \cdot; \mu)$  gleichmäßig koerziv und  $f(\cdot; \mu)$  gleichmäßig beschränkt, so ist  $\mathcal{M}$  beschränkt

$$\mathcal{M} \subseteq B_{\frac{\bar{\gamma}_f}{\bar{\alpha}}}(0)$$

*Beweis.* Klar weil  $\|u(\mu)\| \leq \frac{\bar{\gamma}_f}{\bar{\alpha}}$  nach Satz 2.8. □

**Satz 2.15** (Lipschitz-Stetigkeit)

Falls  $a(\cdot, \cdot; \mu)$ ,  $f(\cdot; \mu)$ ,  $l(\cdot; \mu)$  Lipschitz-stetig bzgl.  $\mu$ , so sind  $u(\mu)$  und  $s(\mu)$  Lipschitz-stetig bzgl.  $\mu$  mit Lipschitz-Konstanten

$$L_u = \frac{L_f}{\bar{\alpha}} + \bar{\gamma}_f \frac{L_a}{\bar{\alpha}^2} \quad \text{und} \quad L_s = L_l \frac{\bar{\gamma}_f}{\bar{\alpha}} + \bar{\gamma}_l L_u$$

*Beweis.* Übung. □

**Satz 2.16** (Diffbarkeit)

Sei  $a(u, \cdot; \mu) \in X'$  Frechet-diffbar in Umgebung von  $(u_0, \mu_0) \subset X \times \mathcal{P}$  und  $f(\cdot; \mu) \in X'$

Frechet-diffbar in Umgebung von  $\mu_0 \in \mathcal{P}$ . Dann ist Lösung  $u(\mu)$  von  $(P(\mu))$  Frechet-diffbar in Umgebung von  $\mu_0 \in \mathcal{P}$  mit

$$D_\mu u(\mu) := - \left( \frac{\partial}{\partial u} F(u, \mu) \right)^{-1} \frac{\partial}{\partial \mu} F(u, \mu)$$

wobei  $F(u, \mu) := a(u, \cdot; \mu) - f(\cdot; \mu) \in X'$ .

*Beweis.* Aus Frechet-Diffbarkeit von  $a(\cdot, \cdot; \cdot)$  und  $f(\cdot; \cdot)$  folgt Frechet-Diffbarkeit von  $F : X \times \mathcal{P} \rightarrow X'$  in Umgebung von  $(u_0, \mu_0)$  mit partiellen Ableitungen

$$\frac{\partial}{\partial \mu} F(u_0, \mu_0) := \frac{\partial}{\partial \mu} a(u_0, \cdot; \mu_0) - \frac{\partial}{\partial \mu} f(\cdot; \mu_0) \in L(\mathbb{R}^p, X')$$

und  $\frac{\partial}{\partial u} F(u_0, \mu_0) \in L(X, X')$  durch

$$\frac{\partial}{\partial u} F(u_0, \mu_0) h_u := a(h_u, \cdot; \mu_0) \in X' \quad \forall h_u \in X$$

Dann erfüllt  $u(\mu)$  als Lösung von  $(P(\mu))$  gerade

$$F(u(\mu), \mu) = 0$$

in Umgebung von  $\mu_0$ . Dann ist (z.B. mit Folgerung 2.15 in Ruzicka: Nichtlineare Funktionalanalysis, Springer 2004) auch  $u(\mu)$  Frechet-diffbar in Umgebung von  $\mu_0$  mit Ableitung

$$D_\mu u(\mu) := - \left( \frac{\partial}{\partial u} F(u, \mu) \right)^{-1} \frac{\partial}{\partial \mu} F(u, \mu)$$

□

### Bemerkung.

- Plausibilität der Ableitungsformel folgt aus formellem Ableiten:

$$\begin{aligned} & D_\mu (F(u(\mu), \mu)) = 0 \\ \Rightarrow & \frac{\partial}{\partial u} F(u(\mu), \mu) D_\mu u(\mu) + \frac{\partial}{\partial \mu} F(u, \mu) = 0 \\ \Rightarrow & \frac{\partial}{\partial u} F(u(\mu), \mu) D_\mu u(\mu) = - \frac{\partial}{\partial \mu} F(u, \mu) \\ \Rightarrow & D_\mu u(\mu) = - \left( \frac{\partial}{\partial u} F(u(\mu), \mu) \right)^{-1} \frac{\partial}{\partial \mu} F(u, \mu) \end{aligned}$$

- Man kann zeigen, dass die Sensitivitäts-Ableitungen  $\partial_{\mu_i} u(\mu) \in X$  für  $i = 1, \dots, p$  erfüllen das sogenannte *Sensitivitätsproblem*

$$a(\partial_{\mu_i} u(\mu), v; \mu) = \tilde{f}_i(v; u(\mu), \mu)$$

mit rechter Seite  $\tilde{f}_i(\cdot; u(\mu), \mu) \in X'$  gegeben durch

$$\tilde{f}_i(\cdot; w, \mu) := \partial_{\mu_i} f(\cdot; \mu) - \partial_{\mu_i} a(w, \cdot; \mu)$$

d.h. das Problem  $(P(\mu))$  mit modifizierter rechter Seite, in welcher insbesondere  $u(\mu)$  eingeht. (Übung)

- Hinreichend für die Diffbarkeit von  $a, f$  in Satz 2.16 sind z.B. im Fall von separierbarer Parameterabhängigkeit die Diffbarkeit der Koeffizienten  $\Theta_a^q(\mu), \Theta_f^q(\mu)$ ,  $q = 1, \dots$  (Übung)
- Ähnliche Aussagen / Sensitivitätsprobleme gelten für Ableitungen höherer Ordnung. Also überträgt sich Glattheit der Koeffizientenfunktionen auf Glattheit der Lösung / Mannigfaltigkeit.

### 3 RB-Methoden für lineare koerzive Probleme

#### 3.1 Primales RB-Problem

**Definition 3.1** (Reduzierte Basis, RB-Räume)

Sei  $S_N = \{\mu_1, \dots, \mu_N\} \subset \mathcal{P}$  Menge von Parametern mit (o.B.d.A.) linear unabhängigen Lösungen  $\{u(\mu_i)\}_{i=1}^N$  von  $(P(\mu_i))$ . Dann ist  $X_N := \text{span} \{u(\mu_i)\}_{i=1}^N$  ein sog. *Lagrange-RB-Raum*.

Sei  $\mu^0 \in \mathcal{P}$  und  $u(\mu)$  Lösung von  $(P(\mu^0))$   $k$ -mal diffbar in Umgebung von  $\mu^0$ . Dann ist

$$X_{k,\mu^0} := \text{span} \{ \partial_\sigma u(\mu^0) : \sigma \in \mathbb{N}_0^p, |\sigma| \leq k \}$$

ein *Taylor-RB-Raum*. Eine Basis  $\Phi_N = \{\varphi_1, \dots, \varphi_N\} \subseteq X$  eines RB-Raums ist eine *reduzierte Basis*.

**Bemerkung.**

- $\Phi_N$  kann direkt aus Snapshots  $u(\mu^i)$  oder, für numerische Stabilität (siehe ??), auch orthonormiert sein.
- Wahl der Parameter  $\{\mu^i\}$  ist entscheidend für Güte des RB-Modells:  
Hier: zufällige oder äquidistante Menge ausreichend  
Später: intelligente Wahl durch a-priori Analysis oder Greedy-Verfahren
- Es ex. auch andere Arten von RB-Räumen (Hermite, POD). Gemeinsam ist diesen die Konstruktion aus Snapshots von  $u$  bzw.  $\partial_\sigma u$ .
- Andere MOR-Techniken:  $\Phi_N$  kann auch komplett unabhängig von Snapshots auf andere Weise konstruiert werden: Balanced Truncation, Krylov-Räume, etc. (siehe z.B. Antoulas: Approximation of large scale dynamical systems, SIAM 2004)

**Definition 3.2** (Reduziertes Problem  $(P_N(\mu))$ )

Sei eine Instanz von  $(P(\mu))$  gegeben und  $X_N \subseteq X$  ein RB-Raum. Zu  $\mu \in \mathcal{P}$  ist die RB-Lösung  $u_N(\mu) \in X_N$  und Ausgabe  $s_N(\mu) \in \mathbb{R}$  gesucht mit

$$\begin{aligned} a(u_N(\mu), v; \mu) &= f(v; \mu) & \forall v \in X_N \\ s_N(\mu) &= l(u_N; \mu) \end{aligned}$$

**Bemerkung.**

- Wir nennen obiges “primal” weil im Fall  $f \neq l$  oder  $a$  asymmetrisch, kann mit Hilfe eines geeigneten dualen Problems bessere Schätzung für  $s$  erreicht werden.
- Obiges ist “Ritz-Galerkin”-Projektion im Gegensatz zu “Petrov-Galerkin”-Projektion, welches für nicht-koerzive Probleme notwendig ist.  $\rightsquigarrow$  ??

**Satz 3.3** (Galerkin-Projektion, Galerkin-Orthogonalität)

Sei  $P_\mu : X \rightarrow X_N$  die orthogonale Projektion bzgl. Energieskalarprodukt  $\langle \cdot, \cdot \rangle_\mu$ , sei  $a$  symmetrisch und  $u(\mu)$ ,  $u_N(\mu)$  Lösung von  $(P(\mu))$  bzw.  $(P_N(\mu))$ . Dann:



- i)  $u_N(\mu) = P_\mu u(\mu)$  “Galerkin-Projektion”  
 ii)  $\langle e(\mu), v \rangle_\mu = 0 \quad \forall v \in X_N$ , wobei  $e(\mu) := u(\mu) - u_N(\mu)$

*Beweis.* Nach Aufgabe 1/Blatt 1 ist  $P_\mu$  wohldefiniert, denn  $(X, \langle \cdot, \cdot \rangle_\mu)$  ist Hilbertraum und  $X_N \subseteq X$  abgeschlossen weil endlichdimensional. Orthogonale Projektion des Fehlers ergibt

$$\begin{aligned} & \langle P_\mu u(\mu) - u(\mu), \varphi_i \rangle_\mu = 0 & \forall i = 1, \dots, N \\ \Leftrightarrow & a(P_\mu u(\mu) - u(\mu), \varphi_i; \mu) = 0 & \forall i = 1, \dots, N \\ \Leftrightarrow & a(P_\mu u(\mu), \varphi_i; \mu) = a(u(\mu), \varphi_i; \mu) = f(\varphi_i; \mu) & \forall i = 1, \dots, N \end{aligned}$$

- i) also ist  $P_\mu u(\mu)$  Lösung von  $(P_N(\mu))$   
 ii)  $e(\mu)$  ist also Projektions-Fehler, orthogonal nach Aufgabe 1/Blatt 1

□

**Bemerkung.** Für  $a$  nichtsymmetrisch gilt immer noch folgende “Galerkin-Orthogonalität”

$$a(u - u_N, v; \mu) = 0 \quad \forall v \in X_N$$

(auch wenn  $a$  kein Skalarprodukt)

**Satz 3.4** (Existenz und Eideutigkeit für  $(P_N(\mu))$ )

Zu  $\mu \in \mathcal{P}$  ex. eindeutige Lösung  $u_N(\mu) \in X_N$  und RB-Ausgabe  $s_N(\mu) \in \mathbb{R}$  von  $(P_N(\mu))$ . Diese sind beschränkt

$$\begin{aligned} \|u_N(\mu)\| &\leq \frac{\|f(\cdot; \mu)\|_{X'}}{\alpha(\mu)} \leq \frac{\bar{\gamma}_f}{\bar{\alpha}} \\ \|s_N(\mu)\| &\leq \|l(\cdot; \mu)\| \|u_N(\mu)\| \leq \frac{\bar{\gamma}_l \bar{\gamma}_f}{\bar{\alpha}} \end{aligned}$$

*Beweis.* Weil  $X_N \subset X$  ist  $a(\cdot, \cdot; \mu)$  stetig und koerziv auf  $X_N$ .

$$\begin{aligned} \alpha_N(\mu) &:= \inf_{v \in X_N} \frac{a(v, v; \mu)}{\|v\|^2} \geq \inf_{v \in X} \frac{a(v, v; \mu)}{\|v\|^2} = \alpha(\mu) > 0 \\ \gamma_N(\mu) &:= \sup_{u, v \in X_N} \frac{a(u, v; \mu)}{\|u\| \|v\|} \leq \sup_{u, v \in X} \frac{a(u, v; \mu)}{\|u\| \|v\|} = \gamma(\mu) < \infty \end{aligned}$$

analog  $f, l$  stetig auf  $X_N$ . Existenz, Eindeutigkeit und Schranken folgen also mit Lax-Milgram analog zu 2.8. □

**Korollar 3.5** (Lipschitz-Stetigkeit)

Seien  $f, l$  gleichmäßig beschränkt und  $a, f, l$  Lipschitz-stetig bzgl.  $\mu$ , dann sind auch  $u_N(\mu), s_N(\mu)$  Lipschitz-stetig bzgl.  $\mu$  mit  $L_u, L_s$  wie in 2.15.

*Beweis.* Analog zu 2.15 / Übung. □

**Satz 3.6** (Diskrete RB-Probleme)

Sei  $\Phi_N = \{\varphi_1, \dots, \varphi_N\}$  eine reduzierte Basis für  $X_N$ . Für  $\mu \in \mathcal{P}$ ,

$$\begin{aligned} A_N(\mu) &:= (a(\varphi_j, \varphi_i; \mu))_{i,j=1}^N && \in \mathbb{R}^{N \times N} \\ \underline{l}_N(\mu) &:= (l(\varphi_i; \mu))_{i=1}^N && \in \mathbb{R}^N \\ \underline{f}_N(\mu) &:= (f(\varphi_i; \mu))_{i=1}^N && \in \mathbb{R}^N \end{aligned}$$

und  $\underline{u}_N = (u_{N,i})_{i=1}^N \in \mathbb{R}^N$  als Lösung von

$$A_N(\mu) \underline{u}_N = \underline{f}_N(\mu) \quad (3.1)$$

Dann ist  $u_N(\mu) := \sum_{i=1}^N u_{N,i} \varphi_i$  und  $s_N(\mu) := \underline{l}_N^\top(\mu) \underline{u}_N$ .

*Beweis.* Einsetzen und Linearität zeigt, dass

$$a\left(\sum u_{N,j} \varphi_j, \varphi_i; \mu\right) = (A_N(\mu) \underline{u}_N)_i = (\underline{f}_N)_i = f(\varphi_i; \mu)$$

□

**Satz 3.7** (Kondition bei ONB und Symmetrie)

Falls  $a(\cdot, \cdot; \mu)$  symmetrisch und  $\Phi_N$  ist ONB, so ist Kondition von (3.1) unabhängig von  $N$  beschränkt

$$\text{cond}_2(A_N) := \|A_N\|_2 \|A_N^{-1}\|_2 \leq \frac{\gamma(\mu)}{\alpha(\mu)}$$

*Beweis.* Wegen Symmetrie gilt

$$\text{cond}_2(A_N) = \frac{|\lambda_{\max}|}{|\lambda_{\min}|} \quad (3.2)$$

mit betragsmäßig größtem/kleinstem Eigenwert  $\lambda_{\max}/\lambda_{\min}$  von  $A_N(\mu)$ . Sei  $\underline{u}_{\max} = (u_i)_{i=1}^N \in \mathbb{R}^N$  Eigenvektor zu  $\lambda_{\max}$  und

$$\underline{u}_{\max} := \sum_{i=1}^N u_i \varphi_i \in X_N$$

Dann gilt

$$\begin{aligned} \lambda_{\max} \|\underline{u}_{\max}\|^2 &= \lambda_{\max} \underline{u}_{\max}^\top \underline{u}_{\max} = \underline{u}_{\max}^\top A_N \underline{u}_{\max} \\ &= \sum_{i,j=1}^N u_i u_j a(\varphi_j, \varphi_i; \mu) = a\left(\sum_j u_j \varphi_j, \sum_i u_i \varphi_i; \mu\right) \\ &= a(\underline{u}_{\max}, \underline{u}_{\max}; \mu) \leq \gamma(\mu) \|\underline{u}_{\max}\|^2 \end{aligned}$$

Wegen

$$\|\underline{u}_{\max}\|^2 = \left\langle \sum u_i \varphi_i, \sum u_j \varphi_j \right\rangle = \sum u_i u_j \langle \varphi_i, \varphi_j \rangle = \sum u_i^2 = \|\underline{u}_{\max}\|^2$$

folgt  $|\lambda_{\max}| \leq \gamma(\mu)$ . Analog zeigt man  $|\lambda_{\min}| \geq \alpha(\mu)$  also folgt mit (3.2) die Behauptung. □

**Bemerkung** (Unterschied FEM zu RB). Es bezeichne  $A_h(\mu) \in \mathbb{R}^{H \times H}$  die FEM Matrix (oder FV/FD).

- i) Die RB-Matrix  $A_N(\mu) \in \mathbb{R}^{H \times H}$  ist klein aber typischerweise vollbesetzt im Gegensatz zur großen aber dünnbesetzten Matrix  $A_h$ .
- ii) Die Kondition von  $A_N$  verschlechtert sich nicht mit wachsendem  $N$  (falls eine ONB verwendet wird), während die Konditionszahl von  $A_h$  typischerweise polynomiell in  $H$  wächst, also schlechter wird.

**Satz 3.8** (Reproduktion von Lösungen)

Seien  $u(\mu)$ ,  $u_N(\mu)$  Lösungen von  $(P(\mu))$  bzw.  $(P_N(\mu))$ ,  $e_i \in \mathbb{R}^n$   $i$ -ter Einheitsvektor

- i) Falls  $u(\mu) \in X_N \Rightarrow u_N(\mu) = u(\mu)$
- ii) Falls  $u(\mu) = \varphi_i \in \Phi_N \Rightarrow u_N(\mu) = e_i \in \mathbb{R}^N$

*Beweis.*

- i) Mit  $u(\mu), u_N(\mu) \in X_N \Rightarrow e := u(\mu) - u_N(\mu) \in X_N$ . Wegen Galerkin-Orthogonalität ( $a(e, v; \mu) = 0 \forall v \in X_N$ ) und Koerzivität folgt:

$$0 = a(e, e; \mu) \geq \underbrace{\alpha(\mu)}_{>0} \underbrace{\|e\|^2}_{\geq 0} \Rightarrow \|e\| = 0 \Rightarrow e = 0 \Rightarrow u = u_N$$

- ii)  $u_N(\mu) = \varphi_i$ , nach i). Mit Eindeutigkeit der Basisexpansion folgt die Behauptung. □

**Bemerkung.**

- Reproduktion von Lösungen ist grundlegende Konsistenzeigenschaft. Es gilt trivialerweise falls/sobald Fehlerschranken vorliegen, aber für komplexe RB-Probleme ohne Fehlerschranken ist obiges ein guter Test.
- Validierung für Programmcode: Wähle Basis aus Snapshots  $\varphi_i = u(\mu^i)$ ,  $i = 1, \dots, N$ , ohne Orthonormierung, dann muss  $u_N(\mu^i) = e_i \in \mathbb{R}^N$  ein Einheitsvektor sein.

## 3.2 Fehleranalyse

**Satz 3.9** (Céa, Beziehung zur Bestapproximation)

Für alle  $\mu \in \mathcal{P}$  gilt

$$\|u(\mu) - u_N(\mu)\| \leq \frac{\gamma(\mu)}{\alpha(\mu)} \inf_{v \in X} \|u - v\|$$

*Beweis.*  $\forall v \in X_N$  mit Stetigkeit und Koerzivitat

$$\begin{aligned} \alpha \|u - u_N\|^2 &\leq a(u - u_N, u - u_N) = a(u - u_N, u - v) + \underbrace{a(u - u_N, v - u_N)}_{=0 \text{ (Galerkin-Orth.)}} \\ &\leq \gamma(\mu) \|u - u_N\| \|u - v\| \end{aligned}$$

Division durch  $\alpha$ ,  $\|u - u_N\|$  liefert

$$\|u - u_N\| \leq \frac{\gamma}{\alpha} \|u - v\|$$

also Behauptung durch Infimum-Bildung.  $\square$

**Bemerkung.**

- i) hnliche Bestapproximationsaussagen gelten auch fur andere Interpolationstechniken, aber die zugehorige Lebesgue-Konstante divergiert meist mit wachsender Dimension  $N$ . Obiges ist konzeptioneller Vorteil von Galerkin-Projektion uber anderen Interpolationstechniken, da  $\frac{\gamma}{\alpha}$  unabhangig von  $N$  beschrankt bleibt. “Quasi-Optimalitat” der Galerkin-Projektion/des RB-Ansatzes.
- ii) Falls  $a(\cdot, \cdot; \mu)$  zusatzlich symmetrisch ist, kann um eine “Wurzel” verbessert werden mittels Normaquivalenz 2.5 und Bestapproximation der orthogonalen Projektion (Aufg. 1/Blatt 1)

$$\begin{aligned} \sqrt{\alpha} \|u - u_N\| &\stackrel{2.5}{\leq} \|u - u_N\|_\mu = \|u - P_\mu u\|_\mu = \inf_{v \in X_N} \|u - v\|_\mu \stackrel{2.5}{\leq} \sqrt{\gamma} \inf_{v \in X_N} \|u - v\| \\ \Rightarrow \|u - u_N\| &\leq \sqrt{\frac{\gamma}{\alpha}} \inf_{v \in X_N} \|u - v\| \end{aligned}$$

- iii) Implikation von 3.9: Wahle guten Approximationsraum  $X_N$ , so wird Galerkin-Projektion/RB-Approximation auch garantiert gut sein.

**Satz 3.10** (Ausgabe und Bestapproximation)

- i) Fur alle  $\mu \in \mathcal{P}$  gilt

$$|s(\mu) - s_N(\mu)| \leq \|l(\cdot; \mu)\|_{X'} \frac{\gamma(\mu)}{\alpha(\mu)} \inf_{v \in X_N} \|u - v\|$$

- ii) Fur den sog. “compliant” Fall (d.h.  $a(\cdot, \cdot; \mu)$  symmetrisch und  $l = f$ ) gilt sogar

$$\begin{aligned} 0 \leq s(\mu) - s_N(\mu) &= \|u - u_N\|_\mu^2 \\ &= \inf_{v \in X_N} \|u - v\|_\mu^2 \\ &\leq \gamma(\mu) \inf_{v \in X_N} \|u - v\|^2 \end{aligned}$$

*Beweis.*

i) Klar mit Céa, Bestapproximation und Linearität

$$|s(\mu) - s_N(\mu)| = |l(u) - l(u_N)| = |l(u - u_N)| \leq \|l\| \|u - u_N\| \leq \|l\| \frac{\gamma}{\alpha} \inf_{v \in X_N} \|u - v\|$$

ii) Wegen  $a(\cdot, \cdot; \mu)$  symmetrisch gilt wie in voriger Bemerkung

$$\|u - u_N\|_\mu = \|u - P_\mu u\|_\mu = \inf_{v \in X_N} \|u - v\| \quad (3.3)$$

Damit

$$\begin{aligned} s(\mu) - s_N(\mu) &= l(u) - l(u_N) \stackrel{f=l}{=} f(u) - f(u_N) = f(u - u_N) \\ &= a(u, u - u_N) - \underbrace{a(u_N, u - u_N)}_{=0 \text{ (Gal.-Orth./Symm.)}} = \|u - u_N\|_\mu^2 \\ &\stackrel{3.3}{=} \inf_{v \in X_N} \|u - v\|_\mu^2 \\ &\stackrel{2.5}{\leq} \gamma \inf_{v \in X_N} \|u - v\|^2 \end{aligned}$$

Also insbesondere  $s - s_N = \|u - u_N\|_\mu^2 \geq 0$ .

□

**Bemerkung.**

- Im “compliant” Fall ist der Ausgabefehler i.A. sehr klein, da das Quadrat des RB-Fehlers eingeht.
- Im “nicht-compliant” Fall geht der RB-Fehler nur linear in die Schranke ein, das wird später durch primal-duale Technik verbessert.
- Aus ii) folgt nicht nur Fehlerschranke, sondern sogar Vorzeichen-Information,  $s_N(\mu)$  ist untere Schranke für  $s$ .

**Korollar 3.11** (Monotoner Fehlerabfall in Energienorm)

Falls  $a(\cdot, \cdot; \mu)$  symmetrisch,  $(X_N)_{N=1}^{N_{\max}}$  Folge von RB-Räumen, mit  $X_N \subseteq X_{N'}, \forall N \leq N'$  (“hierarchische Räume”) und für  $\mu \in \mathcal{P}$  setze  $e_{u,N} := u(\mu) - u_N(\mu)$ ,  $e_{s,N} := s(\mu) - s_N(\mu)$ .

i) Dann ist  $(\|e_{u,N}\|_\mu)_{N=1}^{N_{\max}}$  monoton fallend.

ii) Falls  $l = f$  (also “compliant” Fall) ist  $e_{s,N}$  monoton fallend.

*Beweis.*

i) Mit (3.3) gilt für  $N \leq N'$

$$\|e_{u,N}\|_\mu = \|u - u_N\|_\mu \stackrel{(3.3)}{=} \inf_{v \in X_N} \|u - v\|_\mu \geq \inf_{v \in X_{N'}} \|u - v\|_\mu \stackrel{(3.3)}{=} \|e_{u,N'}\|_\mu$$

ii) Mit Satz 3.10 ii) gilt

$$e_{s,N} = \|e_{u,N}\|_\mu^2, \text{ also Behauptung folgt mit i)}$$

□

**Bemerkung.**

- “Worst-case” ist Stagnation des Fehlers (unrealistisch, jeder neue Basisvektor müsste orthogonal zum Fehler  $e_N(\mu)$  sein). In Praxis ist bei geschickter Basiswahl und “glatten” Problemen exponentielle Konvergenz zu erwarten, siehe Basisgenerierung, §??.
- Monotonie gilt nicht notwendigerweise bezüglich anderen Normen trotz Normäquivalenz

$$c\|e_{u,N}\|_\mu \leq \|e_{u,N}\| \leq C\|e_{u,N}\|_\mu, \text{ mit } c, C \text{ unabhängig von } N$$

Fehlernorm  $\|e_{u,N}\|$  kann gelegentlich anwachsen, bleibt aber in einem “Korridor”, welcher monoton fällt.

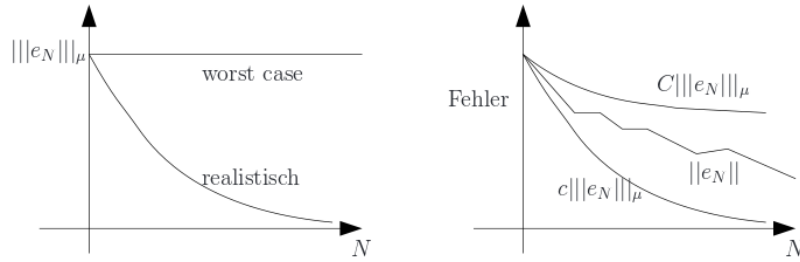


Abbildung 5: Fehlerabfall mit wachsender reduzierter Dimension.  
(aus B. Haasdonk, Reduzierte-Basis-Methoden, Skript zur Vorlesung SS 2011, Universität Stuttgart, IANS-Report 4/11, 2011.)

**Bemerkung** (Gleichmäßige Konvergenz von Lagrange RB-Ansatz).

- Sei  $\mathcal{P}$  kompakt und  $S_N := \{\mu^1, \dots, \mu^N\} \subset \mathcal{P}$ ,  $N \in \mathbb{N}$ , sodass die sog. Füll-Distanz (fill-distance)  $h_N$  gegen 0 geht:

$$h_N := \sup_{\mu \in \mathcal{P}} \text{dist}(\mu, S_N), \quad \text{dist}(\mu, S_N) := \min_{\mu' \in S_N} \|\mu - \mu'\|$$

$$\lim_{N \rightarrow \infty} h_N = 0$$

- Falls  $u(\mu)$ ,  $u_N(\mu)$  Lipschitz-stetig mit Lipschitz-Konstante  $L_u$  unabhängig von  $N$ , so folgt für alle  $N$ ,  $\mu$  und “nächstes”  $\mu^* := \arg \min_{\mu' \in S_N} \|\mu - \mu'\|$ :

$$\begin{aligned} \|u(\mu) - u_N(\mu)\| &\leq \|u(\mu) - u(\mu^*)\| + \|u(\mu^*) - u_N(\mu^*)\| + \|u_N(\mu^*) - u_N(\mu)\| \\ &\leq L_u \underbrace{\|\mu - \mu^*\|}_{\leq h_N} + 0 + L_u \underbrace{\|\mu - \mu^*\|}_{\leq h_N} \leq 2L_u h_N \end{aligned}$$

- Also folgt uniforme Konvergenz

$$\lim_{N \rightarrow \infty} \sup_{\mu \in \mathcal{P}} \|u(\mu) - u_N(\mu)\| = 0$$

- Jedoch Konvergenzrate linear in  $h_N$  ist nicht praktisch bedeutsam, weil  $h_N$  sehr langsam mit  $N$  abfällt, also muss  $N$  sehr groß sein, um kleinen Fehler zu garantieren.
- Wir werden sehen, dass bei gleichmäßig koerziven Problemen und geschickter Wahl der  $\mu^i$  sogar exponentielle Konvergenz erreicht wird.

**Lemma 3.12** (Fehler-Residuums-Beziehung)

Für  $\mu \in \mathcal{P}$  definieren wir mittels der RB-Lösung  $u_N$  das Residuum  $r(\cdot; \mu) \in X'$  bzw. seinen Riesz-Repräsentanten  $v_r(\mu) \in X$

$$\langle v_r(\mu), v \rangle_X := r(v; \mu) := f(v; \mu) - a(u_N(\mu), v; \mu) \quad \forall v \in X$$

Dann erfüllt der Fehler  $e(\mu) := u(\mu) - u_N(\mu)$

$$a(e(\mu), v; \mu) = r(v; \mu) \quad \forall v \in X$$

*Beweis.*  $a(e(\mu), v; \mu) = \underbrace{a(u, v)}_{f(v)} - a(u_N, v) = r(v)$

□

**Bemerkung.**

- Fehler erfüllt “ $(P(\mu))$  mit Residuum als rechte Seite”
- Insbesondere ist  $r(v; \mu) = 0 \quad \forall v \in X_N$  (wegen Galerkin-Orthogonalität)
- $r(\cdot; \mu) = 0 \quad \Rightarrow \quad e = 0$

**Satz 3.13** (A-posteriori Fehlerschätzer, absoluter Fehler)

Sei  $\mu \in \mathcal{P}$ ,  $u(\mu)$  bzw.  $u_N(\mu)$  Lösung von  $(P(\mu))$ ,  $(P_N(\mu))$  und  $e = u - u_N$ . Sei  $\alpha_{LB}(\mu)$  eine untere Schranke für  $\alpha(\mu)$  und  $v_r \in X$  Riesz-Repräsentant von  $r(\cdot; \mu)$  aus Lemma 3.12. Dann gelten folgende Schranken

i) Fehler in Energienorm

$$\|e(\mu)\|_\mu \leq \Delta_N^{en}(\mu) := \frac{\|v_r\|}{\sqrt{\alpha_{LB}(\mu)}}$$

ii) Fehler in  $X$ -Norm  $\|\cdot\|$

$$\|e(\mu)\| \leq \Delta_N(\mu) := \frac{\|v_r\|}{\alpha_{LB}(\mu)}$$

iii) Ausgabefehler

$$|s(\mu) - s_N(\mu)| \leq \Delta_{N,s}(\mu) := \|l(\cdot; \mu)\| \Delta_N(\mu)$$

*Beweis.*

i) Normäquivalenz 2.5 impliziert

$$\|e\| \leq \frac{\|e\|_\mu}{\sqrt{\alpha_{LB}(\mu)}}$$

Damit folgt

$$\|e\|_\mu^2 = a_s(e, e) = a(e, e) = r(v) = \langle v_r, e \rangle \leq \|v_r\| \|e\| \leq \frac{\|v_r\|}{\sqrt{\alpha_{LB}(\mu)}}$$

Division durch  $\|e\|_\mu$  liefert die Behauptung i).

ii) Koerzivität liefert

$$\alpha_{LB}(\mu) \|e\|^2 \leq a(e, e) = r(e) = \langle v_r, e \rangle \leq \|v_r\| \|e\|$$

Division durch  $\alpha_{LB}$  und  $\|e\|$  liefert ii).

iii) Stetigkeit von  $l$  liefert

$$|s(\mu) - s_N(\mu)| = |l(u - u_N; \mu)| \leq \|l\| \|u - u_N\| \stackrel{ii)}{\leq} \|l\| \Delta_N$$

□

### Bemerkung.

- $\alpha_{LB}(\mu)$  soll eine *schnell berechenbare* untere Schranke an  $\alpha(\mu)$  sein, z.B.  $\alpha_{LB}(\mu) := \bar{\alpha}$  falls  $\bar{\alpha}$  bekannt, andere Möglichkeiten folgen später (“min  $\Theta$ ”, “SCM”).
- $\Delta_N$  ist also immer um Faktor  $\sqrt{\alpha_{LB}(\mu)}$  schlechter.
- Beschränkung des Fehlers durch Residuums-Norm ist bekannte Technik aus FEM, um FEM-Lösung  $u_h$  gegen Sobolev-Raum Lösung  $u$  abzuschätzen. In diesem Fall ist  $X$   $\infty$ -dimensional und Residuums-Norm algorithmisch nicht berechenbar. In RB-Methoden wird  $\|v_r\|$  eine *berechenbare* Größe sobald  $X$  endlich-dimensional, z.B. FEM-Raum, ist. Für Residuum ist  $u_N(\mu)$  erforderlich, daher sind Schranken “*a posteriori*”.
- Allgemeines Vorgehen (und alternative Begründung für ii)) zur Herleitung von Fehlerschranken: Zeige, dass Fehler  $e$  erfüllt  $(P(\mu))$  mit rechter Seite, genannt  $r$  (Residuum), wende a-priori Stabilitätsaussage an:

$$\|e\| \leq \frac{\|r\|}{\alpha(\mu)} \quad \text{z.B. Lax-Milgram}$$

und erhalte berechenbare Größe durch Wahl  $X = X_{FEM}$  und untere Schranke  $\alpha_{LB}(\mu) \leq \alpha(\mu)$ .



- Weil die Schranken beweisbare obere Schranken an Fehler darstellen, nennt man sie “rigorose” Fehlerschranken (vgl. “zuverlässige” Schätzer in FEM, bei denen jedoch die Konstante unbekannt ist).
- Fehlerschranken liefern eine Absicherung für RB-Methoden, “certified” RB-Methode, im Gegensatz zu vielen anderen Reduktionsmethoden (z.B. Krylov-Raum-Methoden).
- Ausgabefehler ist grob, indem  $\Delta_N$  nur linear eingeht. Verbesserungen können für den “compliant” Fall oder mit primal-dual Techniken erreicht werden. ( $\rightsquigarrow$  §??)

**Korollar 3.14** (Verschwindende Fehlerschranke)

Falls  $u(\mu) = u_N(\mu)$  dann ist  $\Delta_N(\mu) = \Delta_N^{en}(\mu) = \Delta_{N,s}(\mu) = 0$

*Beweis.*

$$\begin{aligned} 0 &= a(0, v; \mu) = a(e, v; \mu) = r(v; \mu) \\ \Rightarrow r &\equiv 0 \Rightarrow \|v_r\| = 0 \Rightarrow \Delta_N = \Delta_N^{en} = \Delta_{N,s} = 0 \end{aligned}$$

□

**Bemerkung.**

- Dies ist initialer Wunsch an eine Fehlerschranke: diese soll verschwinden falls exakte Approximation vorliegt. Dies ist Grundlage dafür, dass der Faktor der Überschätzung endlich ist.
- Aussage ist trivial für *effektive* Fehlerschätzer (sehen wir bald), aber in komplexen Problemen kann 3.14 schon das maximal erreichbare sein.
- 3.14 ist wieder sinnvoll um Programmcode zu validieren.

**Satz 3.15** (A-posteriori Fehlerschranken, relative Fehler)

Mit Bezeichnungen/Voraussetzungen aus 3.13 und unter Annahme, dass alle Brüche im Folgenden wohldefiniert sind, gilt:

i) Für den relativen Fehler gilt in Energienorm:

$$\frac{\|e(\mu)\|_\mu}{\|u(\mu)\|_\mu} \leq \Delta_N^{en,rel}(\mu) := 2 \frac{\|v_r\|}{\sqrt{\alpha_{LB}(\mu)}} \cdot \frac{1}{\|u_N(\mu)\|_\mu} \quad \text{falls} \quad \Delta_N^{en,rel} \leq 1$$

ii) Für den relativen Fehler gilt in  $X$ -Norm:

$$\frac{\|e(\mu)\|}{\|u(\mu)\|} \leq \Delta_N^{rel}(\mu) := 2 \frac{\|v_r\|}{\alpha_{LB}(\mu)} \cdot \frac{1}{\|u_N(\mu)\|} \quad \text{falls} \quad \Delta_N^{rel} \leq 1$$

*Beweis.*

i) Falls  $\Delta_N^{en,rel}(\mu) \leq 1$ , so ist

$$\begin{aligned} \left| \frac{\|u\|_\mu - \|u_N\|_\mu}{\|u_N\|_\mu} \right| &\stackrel{\Delta\text{-Ungl.}}{\leq} \frac{\|u - u_N\|_\mu}{\|u_N\|_\mu} = \frac{\|e\|_\mu}{\|u_N\|_\mu} \stackrel{3.13 \text{ i)}}{\leq} \frac{\|v_r\|}{\sqrt{\alpha_{LB}(\mu)} \|u_N\|_\mu} \\ &= \frac{1}{2} \Delta_N^{en,rel}(\mu) \leq \frac{1}{2} \end{aligned}$$

Falls  $\|u_N\|_\mu > \|u\|_\mu$  gilt  $\|u_N\|_\mu - \|u\|_\mu \leq \frac{1}{2} \|u_N\|_\mu$  also

$$\frac{1}{2} \|u_N\|_\mu \leq \|u\|_\mu \quad (*)$$

Falls  $\|u\|_\mu \geq \|u_N\|_\mu$ , so ist  $(*)$  klar. Damit folgt

$$\frac{\|e\|_\mu}{\|u\|_\mu} \stackrel{3.13 \text{ i)}}{\leq} \frac{\|v_r\|}{\sqrt{\alpha_{LB}}} \cdot \frac{1}{\|u\|_\mu} \stackrel{(*)}{\leq} \frac{\|v_r\|}{\sqrt{\alpha_{LB}}} \cdot \frac{1}{\|u_N\|_\mu} \cdot 2 = \Delta_N^{en,rel}(\mu)$$

ii) analog zu i).

□

### Bemerkung.

- Analog folgt auch relativer Ausgabefehlerschätzer

$$\frac{|s(\mu) - s_N(\mu)|}{|s(\mu)|} \leq \Delta_{N,s}^{rel}(\mu) := \frac{\|l(\cdot; \mu)\| \cdot \Delta_N}{|s_N(\mu)|} \cdot 2 \quad \text{falls} \quad \Delta_{N,s}^{rel}(\mu) \leq 1$$

- Relative Fehlerschranken sind nur mit Zusatzbedingung  $(\Delta_*^{rel} \leq 1)$  gültig. Diese Bedingung ist jedoch konkret überprüfbar. Falls  $\Delta_N^{rel}(\mu) > 1$ , sollte der RB-Raum verbessert werden.

### Satz 3.16 (Effektivität der Fehlerschranken)

Mit Bezeichnungen aus 3.13 sei  $u(\mu) \neq u_N(\mu)$  und  $\gamma_{UB}(\mu) < \infty$  eine obere Schranke an  $\gamma(\mu)$ . Dann sind die *Effektivitäten*  $\eta_N^{en}(\mu)$  und  $\eta_N(\mu)$  definiert und beschränkt durch

i)

$$\eta_N^{en}(\mu) := \frac{\Delta_N^{en}(\mu)}{\|e\|_\mu} \leq \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}$$

Falls  $a(\cdot, \cdot; \mu)$  symmetrisch, gilt sogar  $\eta_N^{en}(\mu) \leq \sqrt{\frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}}$

ii)

$$\eta_N(\mu) := \frac{\Delta_N(\mu)}{\|e\|_\mu} \leq \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)}$$

*Beweis.*

$$\begin{aligned} \text{ii) } \|v_r\|^2 &= \langle v_r, v_r \rangle = r(v_r) = a(e, v_r) \leq \gamma_{UP}(\mu) \|e\| \|v_r\| \\ \|v_r\| &\leq \gamma_{UB}(\mu) \|e\| \end{aligned} \quad (3.4)$$

Damit

$$\frac{\Delta_N(\mu)}{\|e\|} = \frac{\|v_r\|}{\alpha_{LB}} \cdot \frac{1}{\|e\|} \stackrel{(3.4)}{\leq} \frac{\gamma_{UB}}{\alpha_{LB}} \cdot \frac{\|e\|}{\|e\|}$$

i)

$$\frac{\Delta_N^{en}(\mu)}{\|e\|_\mu} = \frac{\|v_r\|}{\sqrt{\alpha_{LB}}} \cdot \frac{1}{\underbrace{\|e\|_\mu}_{\geq \sqrt{\alpha_{LB}} \cdot \|e\|}} \leq \frac{\|v_r\|}{\alpha_{LB}} \cdot \frac{1}{\|e\|} \stackrel{\text{ii)}}{\leq} \frac{\gamma_{UB}}{\alpha_{LB}}$$

Falls  $a(\cdot, \cdot)$  symmetrisch, gilt wegen Normäquivalenz

$$\|v_r\|_\mu \leq \sqrt{\gamma_{UB}} \|v_r\|$$

und

$$\|v_r\|^2 = a(e, v_r) \stackrel{\text{CS}}{\leq} \|e\|_\mu \|v_r\|_\mu \Rightarrow \|v_r\| \leq \|e\|_\mu \cdot \sqrt{\gamma_{UB}}$$

Damit

$$\frac{\Delta_N^{en}(\mu)}{\|e\|_\mu} = \frac{\|v_r\|}{\sqrt{\alpha_{LB}}} \cdot \frac{1}{\|e\|_\mu} \leq \frac{\|e\|_\mu \cdot \sqrt{\gamma_{UB}}}{\sqrt{\alpha_{LB}} \cdot \|e\|_\mu}$$

□

### Bemerkung.

- Wir nennen  $\Delta_N, \Delta_N^{en}$  daher “effektive” Fehlerschranken weil Faktor der Überschätzung höchstens  $\frac{\gamma_{UB}}{\alpha_{LB}}$  beträgt.
- “Rigorousität” also äquivalent mit  $\eta_N(\mu) \geq 1$ .
- Für den Ausgabefehler  $\Delta_{N,s}(\mu)$  ohne weitere Annahmen keine Effektivität beweisbar. Tatsächlich kann  $\frac{\Delta_{N,s}}{|s-s_N|}$  beliebig groß oder nicht definiert sein, falls  $\Delta_{N,s} \neq 0$ , aber  $s(\mu) = s_N(\mu)$ :

Wähle  $X_N$  und  $\mu$  so dass  $u(\mu) \neq u_N(\mu)$ , wird erreicht durch  $u(\mu) \notin X_N$

$$\Rightarrow e(\mu) \neq 0 \Rightarrow \Delta_N \neq 0, \Delta_{N,s} \neq 0 \quad \text{falls } l \neq 0$$

Wähle  $l(\cdot; \mu) \neq 0$ , so dass  $l(u - u_N; \mu) = 0$

$$\Rightarrow s(\mu) - s_N(\mu) = l(u - u_N; \mu) = 0$$

- Wir nennen die Fehlerschranken auch *Fehlerschätzer* weil sie äquivalent zum Fehler sind.

$$\|e\| \leq \Delta_N \leq \eta_N \|e\|$$

**Satz 3.17** (Effektivität, relative Fehlerschätzer)

Für  $\Delta_N^{rel}(\mu)$  aus 3.15 ist Effektivität definiert und beschränkt durch

$$\eta_N^{rel}(\mu) := \frac{\Delta_N^{rel}(\mu)}{\frac{\|e\|}{\|u\|}} \leq 3 \frac{\gamma_{UB}(\mu)}{\alpha_{LB}(\mu)} \quad \text{falls} \quad \Delta_N^{rel}(\mu) \leq 1$$

*Beweis.* Wie in Beweis zu 3.15 impliziert  $\Delta_N^{rel} \leq 1$ :

$$\left| \frac{\|u\| - \|u_N\|}{\|u\|} \right| \leq \frac{1}{2}$$

Falls  $\|u_N\| \leq \|u\|$  so gilt  $\|u\| - \|u_N\| \leq \frac{1}{2}\|u_N\|$  also

$$\|u\| \leq \frac{3}{2}\|u_N\|$$

Falls  $\|u_N\| > \|u\|$ , so ist (\*) klar. Dann gilt

$$\eta_N^{rel}(\mu) = \underbrace{\frac{2\|v_r\|}{\alpha_{LB}(\mu)\|u_N\|}}_{\Delta_N^{rel}} \cdot \frac{1}{\frac{\|e\|}{\|u\|}} \stackrel{(3.4)}{\leq} 2 \frac{\gamma_{UB}\|e\|}{\alpha_{LB}\|e\|} \cdot \frac{\|u\|}{\|u_N\|} \stackrel{(*)}{\leq} 3 \frac{\gamma_{UB}}{\alpha_{LB}}$$

□

**Bemerkung.**

- Ähnlich für  $\Delta_N^{en,rel}$
- Verbesserung von Schranken und Effektivität durch Normwechsel.

Wähle  $\bar{\mu} \in \mathcal{P}$  und  $\|u\| := \|u\|_{\bar{\mu}}$  als neue Norm auf  $X$ . Dann gilt für symmetrisches  $a$ :  $\alpha(\bar{\mu}) = 1 = \gamma(\bar{\mu})$  also Effektivitäten  $\eta_N, \eta_N^{en} = 1$ , Schätzer sind genau der echte Fehler. Dies lässt  $u_N$  unberührt, liefert aber bessere Fehlerschätzung. Im Fall von Stetigkeit bzgl.  $\mu$  kann auch in Umgebung von  $\bar{\mu}$  gute Effektivität erwartet werden.