Technical University of Vienna

Master thesis

Reduced Basis Methods for Low Frequency Electromagnetic Problems

Author:
Christopher Lackner

Supervisor:
Prof. Dr. Joachim Schöberl

2016

Abstract

In this thesis we introduce a new way for adaptively selecting snapshots to construct a reduced basis (RB) subspace for the numerical solution of parabolic differential equations. Our main insterest is in low frequency electromagnetic equations where the displacement currents can be neglected. Constructing the RB subspace by solving shifted stationary problems is a natural and often used attempt, based on the work of Grimme [7]. In [9] the identity of shifts and eigenvalues of the reduced system was derived as a necessary optimality condition. Because the optimal spaces are not nested, Druskin, Lieberman and Zaslavsky proposed the usage of a nested sequence of spaces with adaptively chosen shifts fitted to the eigenvalues in [6]. Using a modified version of the Kolmogorov Smirnow test statistic we derive an algorithm with a better fitting of the shifts to the eigenvalues than in [6]. We present tests on a 2D heat equation example and a 3D electromagnetic one and observe an improved convergence rate with our method.

Acknowledgements

First of all I want to thank Prof. Joachim Schöberl for his inspiring supervision of my work for this thesis and over the last years in general. His comments supported and improved this work greatly.

Many thanks go to my colleague and proofreader Lukas Kogler as well. Lots of discussions and comments from him have improved this thesis.

Finally I also want to thank my girlfriend Alexandra and my family for supporting me during my studies.

Contents

1 Introduction **3**
- 1.1 Electromagnetism . . . 3
- 1.2 Function Spaces . . . 5
- 1.3 Weak Formulation and Discrete Setting . . . 10
- 1.4 Existence and Uniqueness . . . 11
- 1.5 The Fourier Transformed Problem . . . 17
- 1.6 Reduced Basis Method . . . 18
- 1.7 Residuum . . . 19
- 1.8 Time stepping - Runge Kutta Methods . . . 19

2 Optimality Conditions for the Shifts **22**
- 2.1 Rational function approximation . . . 23
- 2.2 Skeleton approximation . . . 24
- 2.3 L^2 - Optimality Conditions . . . 27

3 Nested Reduced Subspaces **31**
- 3.1 The Kolmogorov-Smirnow Test Method . . . 31

4 Additional Stabilization Techniques **33**
- 4.1 Avoiding numerical instabilities by transformation of the source term . . . 33
- 4.2 Orthogonalization to Gradients . . . 34

5 Numerical Examples **37**
- 5.1 Heat Equation . . . 37
- 5.2 Electromagnetic Equation . . . 39

6 Appendix **41**
- 6.1 Construction of orthonormal, divergence free basis . . . 41
- 6.2 Geometrical Distribution of Reduced Eigenvalues . . . 41
- 6.3 Shift - Eigenvalue Correlation . . . 42

1 Introduction

The aim of this thesis is to reduce the order of the system of equations, usually derived by finite elements (FE) or finite differences (FD) methods for a parabolic differential equation. For this a projection onto a good subspace of the FE/FD space is needed. Reduced basis (RB) methods decompose the problem into a slow offline phase, where the subspace is created and a fast online phase, where the equation is solved on the small subspace. This is an additional benefit, if fast online evaluation is needed for known domain.
The discretized parabolic problem

$$M\frac{\partial u(t)}{\partial t} + Au(t) = f(t) = hg(t) \tag{1.1}$$

with $M, A \in \mathbb{R}^{n\times n}$, $h \in \mathbb{R}^n$, $g \in L^2(0,\infty)$ and $u \in L^2([0,\infty), \mathbb{R}^n)$, rises in many physical applications like heat equations or electromagnetic equations with neglectible displacement currents after discretization of the space derivatives. Due to Plancherel's equality the L^2 approximation of problem (1.1) is aquivalent to the approximation of the Fourier transformed problem

$$\tilde{u}(s) = (A + isM)^{-1}h\ \tilde{g}(s) \tag{1.2}$$

for all $s \in \mathbb{R}$. Our aim is to find good subspaces V_r, a basis of V_r, $G \in R^{n\times r}$ and

$$\hat{u}(s) = (\hat{A} + is\hat{M})^{-1}\hat{h}\ g(s) \tag{1.3}$$

with $\hat{A} = G^T AG \in R^{r\times r}$, $\hat{M} = G^T MG \in R^{r\times r}$ $\hat{h} = G^T h \in R^r$, $\hat{u}(t) \in R^r$, such that $\|\tilde{u} - G\hat{u}\|_{L^2(\mathbb{R})}$is sufficiently small.

1.1 Electromagnetism

In this section we want to introduce our main application of the reduced basis method. The equations linking magnetism and electricity were initially developed by Gauss, Ampere and Faraday. James Clerk Maxwell completed the set of equations by introducing the time derivative of the electric field into Ampere's law. The microscopic equations or usually referred to as Maxwells equations in vacuum relate the electric field **E** and the magnetic field **B** to electric charges ρ and electric currents **J**. Displacements of the electric field create a magnetic field and vice versa. The microscopic equations are universally applicable in classical field theory (not incorporating quantuum effects) but include complicated

charges and currents on a microscopic level.

$$\operatorname{div}(\mathbf{E}) = \frac{\rho}{\varepsilon_0} \qquad \text{(Gauss law)} \tag{1.4}$$

$$\operatorname{div}(\mathbf{B}) = 0 \qquad \text{(Gauss law for magnetism)} \tag{1.5}$$

$$\operatorname{curl}(\mathbf{E}) = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{(Faraday's law of induction)} \tag{1.6}$$

$$\operatorname{curl}(\mathbf{B}) = \mu_0 \left(\mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \qquad \text{(Ampere's law)} \tag{1.7}$$

The macroscopic equations average these effects using two additional fields D, the displacement current, and H, the magnetizing field. Now the equations only contain the free charges ρ_f and the free currents $\mathbf{J}_f$.

$$\operatorname{div}(\mathbf{D}) = \frac{\rho_f}{\varepsilon_0} \tag{1.8}$$

$$\operatorname{div}(\mathbf{B}) = 0 \tag{1.9}$$

$$\operatorname{curl}(\mathbf{E}) = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.10}$$

$$\operatorname{curl}(\mathbf{H}) = \mu_0 \left(\mathbf{J}_f + \varepsilon_0 \frac{\partial \mathbf{D}}{\partial t} \right) \tag{1.11}$$

$\mathbf{H}$ and $\mathbf{D}$ are related to $\mathbf{B}$ and $\mathbf{E}$ through an additional set of equations depending on the material, the simplest relation is a material dependent linear one.

$$\mathbf{H} = \frac{1}{\mu_r} \mathbf{B} \tag{1.12}$$

$$\mathbf{D} = \varepsilon_r \mathbf{E}, \tag{1.13}$$

For many materials a linear relation is not appropriate and the coefficient depends nonlinear on the fields or/and the time (the history of the fields). We will use a linear relation, however note that the RB method can also be applied to nonlinear equations, but usually requires expensive operations on the full space for the time stepping. In the linear case the time stepping can be done on the RB subspace only. Using these relations, $\mu = \mu_0 \mu_r$ and $\varepsilon = \varepsilon_0 \varepsilon_r$ we get equations in six variables again.

$$\operatorname{div}(\mathbf{E}) = \frac{\rho_f}{\varepsilon} \tag{1.14}$$

$$\operatorname{div}(\mathbf{B}) = 0 \tag{1.15}$$

$$\operatorname{curl}(\mathbf{E}) = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.16}$$

$$\operatorname{curl}(\mathbf{B}) = \mu \left(\mathbf{J}_f + \varepsilon \frac{\partial \mathbf{E}}{\partial t} \right) \tag{1.17}$$

The free current $\mathbf{J}$ consists of the applied current $\mathbf{J}_A$ and the current from the electric field $\mathbf{J}_E$ which is by Ohm's law

$$\mathbf{J}_E = \sigma \mathbf{E}, \tag{1.18}$$

with σ beeing the conductivity of the material. Since $\mathbf{B}$ is divergence free, and because the kernel of the divergence operator is exactly the range of the curl operator (see de Rham Complex, Theorem 2), $\mathbf{B} = \operatorname{curl}(\mathbf{u})$ for some vector potential $\mathbf{u}$. Using this and (1.18) in (1.17) we get

$$\mu^{-1}\operatorname{curl}(\operatorname{curl}(\mathbf{u})) = \mathbf{J}_A + \sigma\mathbf{E} + \varepsilon\frac{\partial \mathbf{E}}{\partial t}. \tag{1.19}$$

In low frequency computations $\varepsilon\frac{\partial \mathbf{E}}{\partial t} \ll \sigma\mathbf{E}$ and the term $\varepsilon\frac{\partial \mathbf{E}}{\partial t}$ can be neglected. Using (1.16) for the vector potential leads to the equation we want to consider.

$$\sigma\frac{\partial \mathbf{u}}{\partial t} + \mu^{-1}\operatorname{curl}(\operatorname{curl}(\mathbf{u})) = \mathbf{J}_A \tag{1.20}$$

1.2 Function Spaces

For completeness we introduce the function spaces we use, as well as the discrete spaces and some of their basic properties.

Definition 1
Consider the Hilbert spaces

$$L^2(\Omega) = \left\{ v : \int_\Omega |v|^2 < \infty \right\} \tag{1.21}$$

$$H^1(\Omega) = \left\{ v \in L^2(\Omega) : \nabla v \in L^2(\Omega) \right\} \tag{1.22}$$

$$H^{curl}(\Omega) = \left\{ v \in [L^2(\Omega)]^3 : curl(v) \in [L^2(\Omega)]^3 \right\} \tag{1.23}$$

$$H^{div}(\Omega) = \left\{ v \in [L^2(\Omega)]^3 : div(v) \in L^2(\Omega) \right\}. \tag{1.24}$$

with their scalar products, where the derivatives are defined in a weak sense.

$$(u, v)_{L^2(\Omega)} = \int_\Omega \overline{u} v \, dx \tag{1.25}$$

$$(u, v)_{H^1(\Omega)} = \int_\Omega \overline{\nabla u} \cdot \nabla v + \overline{u} v \, dx \tag{1.26}$$

$$(\mathbf{u}, \mathbf{v})_{H^{curl}(\Omega)} = \int_\Omega \overline{curl(\mathbf{u})} \cdot curl(\mathbf{v}) + \overline{\mathbf{u}} \cdot \mathbf{v} \, dx \tag{1.27}$$

$$(\mathbf{u}, \mathbf{v})_{H^{div}(\Omega)} = \int_\Omega \overline{div(\mathbf{u})} div(\mathbf{v}) + \overline{\mathbf{u}} \cdot \mathbf{v} \, dx \tag{1.28}$$

and their corresponding norms $\|u\|_{\mathcal{S}} = \sqrt{(u, u)_{\mathcal{S}}}$.

Definition 2 (Essential boundary conditions)
Define the essential trace operators

$$tr(u) := u|_{\partial\Omega} \tag{1.29}$$

$$tr_\tau(\mathbf{u}) := (\mathbf{u} \cdot \tau)|_{\partial\Omega} \tag{1.30}$$

$$tr_{\mathbf{n}}(\mathbf{u}) := (\mathbf{u} \cdot \mathbf{n})|_{\partial\Omega} \tag{1.31}$$

Theorem 1
The trace operators defined in Definition 2 are well defined bounded operators

$$tr : H^1(\Omega) \to L^2(\partial\Omega) \tag{1.32}$$
$$tr_\tau : H^{curl}(\Omega) \to L^2(\partial\Omega) \tag{1.33}$$
$$tr_{\mathbf{n}} : H^{div}(\Omega) \to L^2(\partial\Omega). \tag{1.34}$$

Using this we can define the spaces with essential boundary conditions

Definition 3 (Spaces with essential boundary conditions)

$$H_0^1(\Omega) := \left\{v \in H^1(\Omega) : tr(v) = 0\right\} \tag{1.35}$$
$$H_0^{curl}(\Omega) := \left\{\mathbf{v} \in H^{curl}(\Omega) : tr_\tau(\mathbf{v}) = 0\right\} \tag{1.36}$$
$$H_0^{div}(\Omega) := \left\{\mathbf{v} \in H^{div}(\Omega) : tr_{\mathbf{n}}(\mathbf{v}) = 0\right\} \tag{1.37}$$

Theorem 2 (De Rham Sequence)
The spaces defined in section 1.2, with their according operators form the so called De Rham sequence

$$\mathbb{R} \xrightarrow{id} H^1 \xrightarrow{\nabla} H^{curl} \xrightarrow{curl} H^{div} \xrightarrow{div} L^2 \xrightarrow{0} 0 \tag{1.38}$$

which is exact in the sense, that the range of each operator corresponds to the kernel of the next operator. In case of essential boundary conditions the sequence

$$\mathbb{R} \xrightarrow{id} H_0^1 \xrightarrow{\nabla} H_0^{curl} \xrightarrow{curl} H_0^{div} \xrightarrow{div} L_0^2 \xrightarrow{0} 0 \tag{1.39}$$

with $L_0^2(\Omega) := \left\{v \in L^2(\Omega) : \int_\Omega v\,dx = 0\right\}$ *is valid and exact.*

We want to use this sequence property, therefore our discrete spaces have to fullfill the De Rham sequence as well. Bossavit [3], with further work of Arnold, Falk and Winther [2] provided methods for constructing piecewise polynomial spaces fullfilling these properties.

Definition 4 (Legendre Polynomials)
The Legendre polynomials are defined by the three-term recurrence

$$l_0(x) := 1, \tag{1.40}$$
$$l_1(x) := x, \tag{1.41}$$
$$(n+1)l_{n+1}(x) := (2n+1)l_n(x)x - nl_{n-1}(x), \qquad n \geq 1 \tag{1.42}$$

The integrated Legendre polynomials of order $n \geq 2$ are defined as

$$L_n(x) := \int_{-1}^{x} l_{n-1}(y)dy \tag{1.43}$$

Lemma 3 (Properties of Legendre polynomials)
The legendre polynomials are $L^2([-1,1])$-orthogonal

$$\int_{-1}^{1} l_i(x)l_j(x)dx = \frac{2}{2i+1}\delta_{ij} \tag{1.44}$$

and span $P^p([-1,1])$. The integrated Legendre polynomials are orthogonal with respect to the H^1 seminorm

$$\int_{-1}^{1} L_i'(x) - L_j'(x)dx = 0, \text{ for } i \neq j, \tag{1.45}$$

they vanish at the interval boundaries and span $P_0^p([-1,1])$. The three term recurrence

$$L_1(x) = x, \tag{1.46}$$

$$L_2(x) = \frac{1}{2}(x^2 - 1), \tag{1.47}$$

$$(n+1)L_{n+1}(x) = (2n-1)xL_n(x) - (n-2)L_{n-1}(x), \quad n \geq 2. \tag{1.48}$$

holds.

Definition 5 (Scaled Legendre Polynomials)
For the tensor product structure on tetrahedral elements we will need scaled versions of the Legendre polynomials

$$l_n^{\mathcal{S}}(x,t) := t^n l_n\left(\frac{x}{t}\right), \quad x \in [-t,t], t \in (0,1] \tag{1.49}$$

$$L_n^{\mathcal{S}}(x,t) := t^n L_n\left(\frac{x}{t}\right), \quad x \in [-t,t], t \in (0,1] \tag{1.50}$$

$$\tag{1.51}$$

Definition 6 ($\boldsymbol{H^1}$ Shape Functions on Tetrahedral Element)
Let λ_i be the hat function in the vertex i then the vertex based basis functions are $\varphi_i^V = \lambda_{v_i}$. For v_1, v_2 being the vertices of the edge E_j, for $0 \leq j \leq p_{E_i} - 2$ the edge based functions of up to order p are

$$\varphi_j^{E_i} = L_{j+2}^{\mathcal{S}}(\lambda_{v_1} - \lambda_{v_2}, \lambda_{v_1} + \lambda_{v_2}). \tag{1.52}$$

For the face $F_i = \{v_1, v_2, v_3\}$, define $\lambda_F := \lambda_{v_1} + \lambda_{v_2} + \lambda_{v_3}$, for $0 \leq j + k \leq p_{F_i} - 3$ the face based functions are

$$\varphi_{j,k}^{F_i} = L_{j+1}^{\mathcal{S}}(\lambda_{v_1} - \lambda_{v_2}, \lambda_{v_1} + \lambda_{v_2})\lambda_{v_3} l_k^{\mathcal{S}}(2\lambda_{v_3} - \lambda_F, \lambda_F) \tag{1.53}$$

For $0 \leq j + k + l \leq p_C - 4$ the cell based functions of the cell $C_i = \{v_1, v_2, v_3, v_4\}$ are

$$\varphi_{j,k,l}^{C} = L_{j+2}^{\mathcal{S}}(\lambda_{v_1} - \lambda_{v_2}, \lambda v_1 + \lambda v_2)\cdot \tag{1.54}$$

$$\cdot\lambda_{v_3} l_k^{\mathcal{S}}(2\lambda_{v_3} - (1 - \lambda_{v_4}), 1 - \lambda_{v_4})\lambda_{v_4} l_l(2\lambda_{v_4} - 1) \tag{1.55}$$

Definition 7 ($H^{\mathbf{curl}}$ Shape Functions on Tetrahedral Element)
Let λ_i be the hat function in the vertex i again, then the H^{curl} conforming shape functions of order p are

- *Edge-based functions of the edge $E_m = [v_1, v_2]$:*
  - *Nédélec functions:*

$$\varphi_m^N = \nabla\lambda_{v_1}\lambda_{v_2} - \lambda_{v_1}\nabla\lambda_{v_2} \tag{1.56}$$

  - *Gradient fields of the H^1 conforming edge functions: for $0 \leq i \leq p_{E_m} - 1$:*

$$\varphi_i^{E_m} = \nabla(L_{i+2}^{\mathcal{S}}(\lambda_{v_1} - \lambda_{v_2}, \lambda_{v_1} + \lambda_{v_2}) \tag{1.57}$$

- *Face-based functions on the face $F_m = [v_1, v_2, v_3]$, define:*

$$u_i := L_{i+2}^{\mathcal{S}}(\lambda_{v_1} - \lambda_{v_2}, \lambda_{v_1} + \lambda_{v_2}), \tag{1.58}$$

$$v_i := \lambda_{v_3} l_i^{\mathcal{S}}(2\lambda_{v_3} - \lambda_F, \lambda_F) \tag{1.59}$$

  for $0 \leq i + j \leq p_{F_m} - 2$:
  - *Type 1 (gradient fields):*

$$\varphi_{(i,j)}^{F_m,1} = \nabla(u_i v_i) \tag{1.60}$$

  - *Type 2:*

$$\varphi_{(i,j)}^{F_m,2} = \nabla u_i v_j - u_i \nabla v_j \tag{1.61}$$

  - *Type 3:*

$$\varphi_{(0,j)}^{F_m,3} = (\nabla\lambda_{v_1}\lambda_{v_2} - \lambda_{v_1}\nabla\lambda_{v_2})v_j \tag{1.62}$$

- *Cell-based functions on the cell $C = [v_1, v_2, v_3, v_4]$, define:*

$$u_i := L_{i+2}^{\mathcal{S}}(\lambda_{v_1} - \lambda_{v_2}, \lambda_{v_1} + \lambda_{v_2}), \tag{1.63}$$

$$v_i := \lambda_{v_3} l_i^{\mathcal{S}}(2\lambda_{v_3} - (1 - \lambda_{v_4}), 1 - \lambda_{v_4}), \tag{1.64}$$

$$w_i := \lambda_{v_4} l_i(2\lambda_{v_4} - 1), \tag{1.65}$$

  for $0 \leq i + j + k \leq p_{C_m} - 3$:

– *Type 1 (gradient fields):*

$$\varphi^{C,1}_{(i,j,k)} = \nabla(u_i v_j w_k) \tag{1.66}$$

– *Type 2:*

$$\varphi^{C,2}_{(i,j,k)} = \nabla u_i v_j w_k - u_i \nabla v_i w_k + u_i v_j \nabla w_k \tag{1.67}$$

$$\varphi^{C,2}_{pC+(i,j,k)} = \nabla u_i v_j w_k + u_i \nabla v_i w_k - u_i v_j \nabla w_k \tag{1.68}$$

– *Type 3:*

$$\varphi^{C,3}_{(0,j,k)} = (\nabla \lambda_{v_1} \lambda_{v_2} - \lambda_{v_1} \nabla \lambda v_2) v_j w_k \tag{1.69}$$

Theorem 4 (Commuting Diagram)
Using the local basis for the H^1 space (Definition 6) for the space V_h and for the H^{curl} space (Definition 7) for W_h. Let Π_V and Π_W be the corresponding interpolation operators, then the left hand side of the commuting diagram

$$\begin{array}{ccccccccc}
\mathbb{R} & \xrightarrow{id} & H^1 & \xrightarrow{\nabla} & H^{curl} & \xrightarrow{curl} & H^{div} & \xrightarrow{div} & L^2 & \xrightarrow{0} & 0 \\
 & & \cup\mathrel{|} & & \cup\mathrel{|} & & \cup\mathrel{|} & & \cup\mathrel{|} & & \\
 & & V & & W & & X & & Y & & \\
 & & \downarrow{\scriptstyle \Pi_V} & & \downarrow{\scriptstyle \Pi_W} & & \downarrow{\scriptstyle \Pi_X} & & \downarrow{\scriptstyle \Pi_Y} & & \\
 & & V_h & \xrightarrow{\nabla} & W_h & \xrightarrow{curl} & X_h & \xrightarrow{div} & Y_h & &
\end{array}$$

holds.

Remark: For the proofs of these properties, other element types and shape functions for the other spaces (H^{div}, L^2) see [12].
Since we want to solve parabolic problems, we have to define functional spaces in time and space. Parabolic equations are of second order in space, but only first order in time, therefore we want to demand different regularity assumptions on the time and space dependence of our solution.

Definition 8
Let H be a Hilbert space,

- *the space $C^k([0,T],H)$ is the set of all functions $u : [0,T] \to H$, which are k-times continuously differentiable. The space is a Banach space with norm*

$$\|u\|_{C^k([0,T],H)} := \sum_{i=0}^{k} \sup_{0 \le t \le T} \left\| u^{(i)}(t) \right\|_H . \tag{1.70}$$

- *the space $L^2((0,T),H)$ is the set of all (aquivalence classes of) measurable functions $u : (0,T) \to H$ with*

$$\|u\|_{L^2((0,T),H)} := \sqrt{\int_0^T \|u(t)\|_H \; dt} < \infty \tag{1.71}$$

$L^2((0,T),H)$ is a Hilbert space with scalar product

$$(u,v)_{L^2((0,T),H)} := \int_0^T (u(t),v(t))_H \; dt \tag{1.72}$$

For our solution to the parabolic equation we need our weak solution to be in the dual space of the solution space only

Definition 9
Define the weak time-space Sobolev space $H^1((0,T),H)$ as

$$H^1((0,T),H) := \left\{u \in L^2((0,T),H) : u_t \in L^2((0,T),H')\right\} \tag{1.73}$$

with H' the dual space of H.

1.3 Weak Formulation and Discrete Setting

We want to solve equation (1.20) using the finite element method for space discretization and a time stepping method to solve the resulting system of ordinary differential equations. For the finite element method we need a weak formulation.

Definition 10 (Weak formulation)
Given $\mathbf{J}_A$, find $\mathbf{u} \in C^1(\mathbb{R}^+, H_0^{curl}(\Omega))$ such that

$$\int_\Omega \sigma \frac{\partial \mathbf{u}}{\partial t} \mathbf{v} dx + \int_\Omega \mu^{-1} curl\,(\mathbf{u})\; curl\,(\mathbf{v})\; dx = \int_\Omega \mathbf{J}_A \mathbf{v}\, dx \tag{1.74}$$

holds for all $\mathbf{v} \in H_0^{curl}(\Omega)$ and all $t \in \mathbb{R}^+$.

Using finite element discretization for equation (1.74) leads to:
Find $u \in C^1(\mathbb{R}^+, \mathbb{R}^n)$ solution to

$$M \frac{\partial u(t)}{\partial t} + Au(t) = f(t) \tag{1.75}$$

with $M \in \mathbb{R}^{n\times n}, A \in \mathbb{R}^{n\times n}$ and $f(t) = g(t) \cdot h$ with $g \in C(\mathbb{R}^+, \mathbb{R}), h \in \mathbb{R}^n$ and

$$M_{ij} = \int_\Omega \sigma \varphi_i \varphi_j dx \tag{1.76}$$

$$A_{ij} = \int_\Omega \mu^{-1} \operatorname{curl}(\varphi_i) \operatorname{curl}(\varphi_j)\, dx \tag{1.77}$$

$$h_i = \int_\Omega \mathbf{J}_A \varphi_i dx, \tag{1.78}$$

where φ_i are the basis functions of the finite element space.

1.4 Existence and Uniqueness

The existence and uniqueness of a solution to problem (1.74) is an essential property to obtain a reliable discrete solution. We will first show these properties on the assumptions of $\sigma, \mu > 0$ and discuss the non conducting ($\sigma = 0$) case later in Theorem 11. We will show the existence and uniqueness not for the space $C^1([0,T], H_0^{\operatorname{curl}}(\Omega))$, but for the weaker space $H^1((0,T), H_0^{\operatorname{curl}}(\Omega))$ from definition 9.

Definition 11
A function $\mathbf{u} \in H^1((0,T), H_0^{curl})$ is a weak solution of equation (1.74) if

$$\int_\Omega \sigma \mathbf{u}' \mathbf{v}\, dx + \int_\Omega \mu^{-1} curl(\mathbf{u})\, curl(\mathbf{v})\, dx = \int_\Omega \mathbf{J}\mathbf{v}\, dx \tag{1.79}$$

holds for all $\mathbf{v} \in H_0^{curl}(\Omega)$, almost all $t \in [0,T]$ and $\mathbf{u}(0) = g$.

We will proof the existence and uniqueness of solutions for this equation with the Galerkin method.

Lemma 5 (Existence of a Galerkin approximant)
Let $\{w_k\}_{k=1}^\infty$ be an orthogonal basis of $H_0^{curl}(\Omega)$ and an orthonormal basis of $[L^2(\Omega)]^3$. There is a unique function

$$\mathbf{u}_m(t) = \sum_{k=1}^m d_m^k(t) w_k \tag{1.80}$$

such that

$$d_m^k(0) = (\mathbf{u}_0, w_k)_{L^2(\Omega)} \tag{1.81}$$

and $\mathbf{u}_m$ solves the Galerkin problem

$$\int_\Omega \sigma \mathbf{u}_m' w_k\, dx + \int_\Omega \mu^{-1} curl(\mathbf{u}_m)\, curl(w_k)\, dx = \int_\Omega \mathbf{J} w_k\, dx \tag{1.82}$$

for all $w_k \in \{w_j\}_{j=1}^m$.

Proof:

Define the matrices M and A and the vectors f and φ with

$$M_{ij} = \int_\Omega \sigma w_i w_j \, dx \tag{1.83}$$

$$A_{ij} = \int_\Omega \mu^{-1} \mathrm{curl}\,(w_i)\,\mathrm{curl}\,(w_j)\; dx \tag{1.84}$$

$$f_i = \int_\Omega \mathbf{J} w_i \, dx \tag{1.85}$$

$$\varphi_i = \int_\Omega \mathbf{u}_0 w_i \, dx. \tag{1.86}$$

Note that M is a positive definite and A a positive semidefinite diagonal matrix. Then $d_m(t) = \begin{pmatrix} d_m^1(t) \\ \vdots \\ d_m^m(t) \end{pmatrix}$ from (1.80) solves

$$M d_m' + A d_m = f \tag{1.87}$$

$$\Leftrightarrow d_m' + M^{-1} A d_m = M^{-1} f \tag{1.88}$$

This equation has the unique solution

$$d_m(t) = e^{-M^{-1}At}\varphi + \int_0^t e^{-M^{-1}A\tau} M^{-1} f(t-\tau)\; d\tau \tag{1.89}$$

with e^B the matrix exponential

$$e^B = \sum_{k=0}^{\infty} \frac{B^k}{k!}. \tag{1.90}$$

Using this solution, the $\mathbf{u}_m$ in (1.80) solves (1.82).

□

For the proof of the uniform continuity of the sequence $\{\mathbf{u}_m\}_{m=1}^\infty$ we need Gronwall's lemma

Lemma 6 (Gronwall's Lemma)
Let $0 \le \xi(t)$ be absolutely continuous function on $[0,T]$ with

$$\xi'(t) \le \phi(t)\xi(t) + \psi(t) \tag{1.91}$$

for almost every $t \in (0,T)$, with $0 \le \phi(t), \psi(t) \in L^1(0,T)$. Then

$$\xi(t) \le e^{\int_0^t \phi(x)\,dx} \left(\xi(0) + \int_0^t \psi(x)\,dx \right) \tag{1.92}$$

for all $0 \le t \le T$.

Lemma 7 (Uniform continuity of the Galerkin approximant)
The sequence $\{\mathbf{u}_m\}$ *is uniformly bounde and the constant* C *in*

$$\max_{0\le t\le T} \|\mathbf{u}_m(t)\|_{L^2(\Omega)} + \|\mathbf{u}_m\|_{L^2((0,T),H_0^{curl}(\Omega))} + \|\mathbf{u}_m'\|_{L^2((0,T),(H_0^{curl})')} \tag{1.93}$$

$$\le C(\|\mathbf{J}\|_{L^2((0,T),L^2(\Omega))} + \|g\|_{L^2(\Omega)}) \tag{1.94}$$

depends only on Ω, T *and the coefficient* μ *and* σ*.*

Proof:
Multiplying equation (1.82) with d_m^k and summing over k gives

$$\int_\Omega \sigma \mathbf{u}_m' \mathbf{u}_m \, dx + \int_\Omega \mu^{-1} \mathrm{curl}\,(\mathbf{u}_m)\,\mathrm{curl}\,(\mathbf{u}_m)\; dx = \int_\Omega \mathbf{J}\mathbf{u}_m \, dx \tag{1.95}$$

Using

$$\int_\Omega \sigma \mathbf{u}_m' \mathbf{u}_m \, dx = \frac{\partial}{\partial t}\left(\frac{1}{2}\int_\Omega \sigma \mathbf{u}_m^2 \, dx\right) \tag{1.96}$$

leads to

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\int_\Omega \sigma \mathbf{u}_m^2 \, dx\right) + \int_\Omega \mu^{-1} \mathrm{curl}\,(\mathbf{u}_m)\,\mathrm{curl}\,(\mathbf{u}_m)\; dx = \int_\Omega \mathbf{J}\mathbf{u}_m \, dx. \tag{1.97}$$

$\sigma > 0$ is only dependent on the space coordinates, therefore

$$\frac{\partial}{\partial t}\|\mathbf{u}_m\|_{L^2(\Omega)} \le C\frac{\partial}{\partial t}\left(\int_\Omega \sigma \mathbf{u}_m^2 \, dx\right) \tag{1.98}$$

Using

$$\int_\Omega \mu^{-1} \mathrm{curl}\,(\mathbf{u}_m)\,\mathrm{curl}\,(\mathbf{u}_m)\; dx \ge 0 \tag{1.99}$$

and

$$\int_\Omega \mathbf{J}\mathbf{u}_m \, dx \le \frac{1}{2}(\|J\|_{L^2(\Omega)}^2 + \|u_m\|_{L^2(\Omega)}^2) \tag{1.100}$$

we get

$$\frac{\partial}{\partial t}\|\mathbf{u}_m\|_{L^2(\Omega)}^2 \le C(\|\mathbf{J}\|_{L^2(\Omega)}^2 + \|\mathbf{u}_m\|_{L^2(\Omega)}^2) \tag{1.101}$$

With

$$\xi(t) := \|\mathbf{u}_m(t)\|_{L^2(\Omega)}^2 \tag{1.102}$$

$$\psi(t) := \|\mathbf{J}\|_{L^2(\Omega)}^2 . \tag{1.103}$$

we can use Gronwall's lemma (Lemma 6) for

$$\xi'(t) \le C(\xi(t) + \psi(t)) \tag{1.104}$$

and get

$$\begin{aligned}\|\mathbf{u}_m(t)\|^2_{L^2(\Omega)} &\leq e^{Ct}\left(\|\mathbf{u}_m(0)\|^2_{L^2(\Omega)} + \int_0^t \|\mathbf{J}(\tau)\|^2_{L^2(\Omega)}\ d\tau\right)\\ &\leq \underbrace{e^{CT}}_{C_1}(\|g\|^2_{L^2(\Omega)} + \|\mathbf{J}\|^2_{L^2((0,T),L^2(\Omega))})\end{aligned} \tag{1.105}$$

Next we want to estimate the $L^2((0,T), H_0^{\mathrm{curl}})$ norm of $\mathbf{u}_m$:

$$\|\mathbf{u}_m\|^2_{L^2((0,T),H_0^{\mathrm{curl}}(\Omega))} = \int_0^T \|\mathbf{u}_m(t)\|^2_{L^2(\Omega)} + \|\mathrm{curl}\,(\mathbf{u})\|^2_{L^2(\Omega)}\ dt \tag{1.106}$$

$$\overset{(1.105)}{\leq} TC_1(\|g\|^2_{L^2(\Omega)} + \|\mathbf{J}\|^2_{L^2((0,T),L^2(\Omega))}) + \int_0^T \|\mathrm{curl}\,(\mathbf{u})\|^2_{L^2(\Omega)}\ dt \tag{1.107}$$

$\mu > 0$, so there holds

$$\|\mathrm{curl}\,(\mathbf{u})\|^2_{L^2(\Omega)} \leq C_2 \int_\Omega \mu^{-1}\mathrm{curl}\,(\mathbf{u}_m)\,\mathrm{curl}\,(\mathbf{u}_m)\ dx \tag{1.108}$$

By integrating equation (1.95) we get an estimate for (1.108).

$$\|\mathbf{u}_m\|^2_{L^2((0,T),H_0^{\mathrm{curl}}(\Omega))} \leq \tag{1.109}$$

$$\leq C_2\left(\|\mathbf{u}_m(0)\|^2_{L^2(\Omega)} + \frac{1}{2}\left(\int_0^T \|\mathbf{u}_m(t)\|^2_{L^2(\Omega)} + \|\mathbf{J}\|^2_{L^2(\Omega)}\ dt\right)\right) \tag{1.110}$$

$$\leq C_2\left(\|g\|^2_{L^2(\Omega)} + \left(\frac{T}{2}C_1 + 1\right)\left(\|g\|^2_{L^2(\Omega)} + \|\mathbf{J}\|^2_{L^2((0,T),L^2(\Omega))}\right)\right) \tag{1.111}$$

$$\leq C_3(\|g\|^2_{L^2(\Omega)} + \|\mathbf{J}\|^2_{L^2((0,T),L^2(\Omega))}) \tag{1.112}$$

Let $v \in H_0^{\mathrm{curl}}(\Omega)$ with $\|v\|_{H_0^{\mathrm{curl}}(\Omega)} \leq 1$. Let $v = v_1 + v_2$ be the decomposition into the Galerkin space and the complement, so $v_1 \in \mathrm{span}\{w_k\}_{k=1}^m$ and $\int_\Omega v_2 w_k\,dx = 0$. Since $\mathbf{u}_m$ is a Galerkin solution, there holds

$$\int_\Omega \sigma\mathbf{u}_m' v_1\,dx + \int_\Omega \mu^{-1}\mathrm{curl}\,(\mathbf{u}_m)\,\mathrm{curl}\,(v_1)\ dx = \int_\Omega \mathbf{J}v_1\,dx \tag{1.113}$$

This implies that

$$\int_\Omega \sigma\mathbf{u}_m' v_1\,dx = \int_\Omega \sigma\mathbf{u}_m' v_1\,dx = \int_\Omega \mathbf{J}v_1\,dx - \int_\Omega \mu^{-1}\mathrm{curl}\,(\mathbf{u}_m)\,\mathrm{curl}\,(v_1)\ dx \tag{1.114}$$

and with Cauchy-Schwarz inequality, $\|v_1\|_{H_0^{\mathrm{curl}}(\Omega)} \leq 1$

$$\|\mathbf{u}_m'\|^2_{(H_0^{\mathrm{curl}})'(\Omega)} \leq C(\|\mathbf{J}\|^2_{L^2(\Omega)} + \|\mathbf{u}_m\|^2_{H_0^{\mathrm{curl}}(\Omega)}) \tag{1.115}$$

Integration in t results in

$$\|\mathbf{u}_m'\|_{L^2((0,T),(H_0^{\mathrm{curl}})'(\Omega))} \leq C(\|\mathbf{J}\|_{L^2((0,T),L^2(\Omega))} + \|\mathbf{u}_m\|_{L^2((0,T),H_0^{\mathrm{curl}}(\Omega))}) \tag{1.116}$$

$$\leq \tilde{C}(\|\mathbf{J}\|_{L^2((0,T),L^2(\Omega))} + \|g\|_{L^2(\Omega)}) \tag{1.117}$$

□

Theorem 8 (Existence of a weak solution)
There exists a weak solution of equation (1.79)

Proof:
According to Lemma 7 $\{\mathbf{u}_m\}$ and $\{\mathbf{u}'_m\}$ are uniformly bounded, so there exist weakly converging subsequences

$$\mathbf{u}_{m_k} \rightharpoonup \mathbf{u} \text{ in } L^2((0,T), H_0^{\text{curl}}(\Omega)) \text{ and} \tag{1.118}$$

$$\mathbf{u}'_{m_k} \rightharpoonup \mathbf{u}' \text{ in } L^2((0,T), (H_0^{\text{curl}})'(\Omega)). \tag{1.119}$$

Next fix an integer n, let $v \in C^1((0,T), H_0^{\text{curl}})$ be of the form

$$v = \sum_{k=1}^{n} \alpha_k(t) w_k \tag{1.120}$$

and $m \geq n$, then

$$\int_0^T \int_\Omega \sigma \mathbf{u}'_m v \, dx + \int_\Omega \mu^{-1} \text{curl}\,(\mathbf{u}_m)\,\text{curl}\,(v)\; dx\, dt = \int_0^T \int_\Omega \mathbf{J} v \, dx\, dt \tag{1.121}$$

and with passing to the weak limit

$$\int_0^T \int_\Omega \sigma \mathbf{u}' v \, dx + \int_\Omega \mu^{-1} \text{curl}\,(\mathbf{u})\,\text{curl}\,(v)\; dx\, dt = \int_0^T \int_\Omega \mathbf{J} v \, dx\, dt \tag{1.122}$$

This holds for all $v \in L^2((0,T), H_0^{\text{curl}})$, since the functions are dense in this space. Therefore

$$\int_\Omega \sigma \mathbf{u}' v \, dx + \int_\Omega \mu^{-1} \text{curl}\,(\mathbf{u})\,\text{curl}\,(v)\; dx = \int_\Omega \mathbf{J} v \, dx \tag{1.123}$$

holds for every $v \in H_0^{\text{curl}}$ and almost every $t \in (0,T)$.

□

Theorem 9 (Uniqueness of the weak solution)
The solution of (1.79) is unique.

Proof:
Because of the linearity of the problem it is enough to show that $\mathbf{u} = 0$ if $\mathbf{J} = \mathbf{u}_0 = 0$. Equation (1.97) with $\mathbf{J} = 0$ yields

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\left\|\mathbf{u}\right\|^2_{L^2(\Omega)}\right) \leq \alpha\left(\frac{\partial}{\partial t}\left(\frac{1}{2}\int_\Omega \sigma \mathbf{u}\cdot\mathbf{u}\, dx\right) + \int_\Omega \mu^{-1}\text{curl}\,(\mathbf{u})\,\text{curl}\,(\mathbf{u})\; dx\right) = 0 \tag{1.124}$$

and since $\|\mathbf{u}(0)\|_{L^2(\Omega)} = \|\mathbf{u}_0\|_{L^2(\Omega)} = 0$, $\mathbf{u} = 0$ for all times.

□

For the case $\sigma = 0$ on some part of the domain we need another basic analysis result:

Theorem 10 (Helmholtz decomposition)
Let $q \in [L^2(\Omega)]^3$. Then there exists a unique decomposition

$$q = \nabla\phi + curl\,(\psi) \tag{1.125}$$

with $\phi \in H_0^1$ and $\psi \in H^{curl}$.

Proof:
Define ϕ by solving the Dirichlet problem, find $\phi \in H_0^1(\Omega)$ with

$$(\nabla\phi, \nabla v)_{L^2(\Omega)} = (q, \nabla v)_{L^2(\Omega)} \qquad \forall v \in H_0^1(\Omega). \tag{1.126}$$

Then $q - \nabla\phi$ is divergence free and using the De Rham sequence property (Theorem 2) the existence of a vector potential ψ with $\mathrm{curl}\,(\psi) = q - \nabla\phi$ follows.

□

Remark: For $u = \mathrm{curl}\,(\psi)$ there follows

$$\mathrm{div}\,(u) = \mathrm{div}\,(\mathrm{curl}\,(\psi)) = 0 \tag{1.127}$$

from the De Rham sequence (Theorem 2).

Theorem 11
If $\sigma = 0$ on some part Ω_0 of the domain Ω a unique solution u to (1.79) still exists on the additional assumption that

$$div\,(u) = 0 \tag{1.128}$$

Proof:
Assume again, that $J(t,x) = f(x)g(t)$ and define

$$V = \left\{v \in H_0^{\mathrm{curl}} : \mathrm{div}\,(v) = 0\right\} \tag{1.129}$$

and let $u := g(t)u_1 + u_2$ with $u_1 \in V$ the solution of

$$\int_\Omega \mu^{-1}\mathrm{curl}\,(u_1)\,\mathrm{curl}\,(v)\;dx = \int_\Omega f v\,dx \tag{1.130}$$

for all $v \in V$. Because of the Helmholtz decomposition of L^2 (Theorem 10) there exists a unique $u_1 \in V$. Plugging $g(t)u_1 + u_2$ into equation (1.79) leads to

$$\int_\Omega \sigma\frac{\partial}{\partial t}(g(t)u_1 + u_2)v + \mu^{-1}\mathrm{curl}\,(g(t)u_1 + u_2)\,\mathrm{curl}\,(v)\;dx = \int_\Omega g(t) f v\,dx. \tag{1.131}$$

Using (1.130) gives

$$\int_\Omega \sigma \frac{\partial u_2}{\partial t} v + \mu^{-1} \mathrm{curl}\,(u_2)\,\mathrm{curl}\,(v)\; dx = -\int_\Omega \sigma \frac{\partial g(t)}{\partial t} u_1 v \, dx \tag{1.132}$$

On Ω_0, u_2 solves

$$0 = \int_{\Omega_0} \mu^{-1} \mathrm{curl}\,(u_2)\,\mathrm{curl}\,(v)\; dx \tag{1.133}$$

From this we get on Ω_0:

$$0 = \mathrm{curl}\,(\mathrm{curl}\,(u_2)) = \underbrace{\nabla \mathrm{div}\,(u_2)}_{=0} + \Delta u_2, \tag{1.134}$$

so u_2 is only a harmonic extension on Ω_0. Therefore solving u_2 on the entire domain is aquivalent to solving it on $\Omega \setminus \Omega_0$, where $\sigma > 0$ and we get unique solvability from Theorem 8 and Theorem 9.

□

Remark: If $\sigma > 0$ we get $\mathrm{div}\,(u) = 0$ from the formulation, because let $u^\nabla = \nabla\varphi$ be the gradient part of the Helmholtz decomposition of u, then equation (1.79) for the gradient part leads to

$$\int_\Omega \sigma \frac{\partial}{\partial t} u^\nabla v \, dx + \underbrace{\int_\Omega \mu^{-1} \mathrm{curl}\left(u^\nabla\right) \mathrm{curl}\,(v)\; dx}_{=0} = \int_\Omega \mathbf{J} v \, dx \tag{1.135}$$

Because $\mathbf{J}$ and u_0 is divergence free we get

$$\int_\Omega \sigma \frac{\partial}{\partial t} u^\nabla v \, dx = 0 \tag{1.136}$$

for all $v \in H_0^{\mathrm{curl}}$. And therefore

$$u^\nabla \equiv 0 \tag{1.137}$$

and so

$$\mathrm{div}\,(u) = 0. \tag{1.138}$$

As a result, if $\sigma = 0$ we have to put the divergence free condition into the space V to get uniqueness.

1.5 The Fourier Transformed Problem

In this section we transfer the time dependent problem using the fourier transformation. Since this tranformation is a unitary map from $L^2(\mathbb{R}) \to L^2(\mathbb{R})$ we can use it in section 1.6 to construct our basis vectors for the approximation in the RB method.

Definition 12
Define the fourier transform of a function $f \in L^1(\mathbb{R})$ as

$$\mathcal{F}f(s) = \hat{f}(s) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ixs} dx \tag{1.139}$$

with (if existing) the fourier back transform

$$f(x) = \mathcal{F}^{-1}\hat{f}(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(s) e^{ixs} dx \tag{1.140}$$

Theorem 12 (Parseval's Identity)
The fourier transform can be extended to a unitary map from $L^2(\mathbb{R}) \to L^2(\mathbb{R})$, this extension will be called $\mathcal{F}$ as well. For all $f \in L^2(\mathbb{R})$ there holds

$$\|f\|_{L^2(\mathbb{R})} = \|\mathcal{F}f\|_{L^2(\mathbb{R})}. \tag{1.141}$$

The problem (1.75) transfers to

$$isM\mathcal{F}u(s) + A\mathcal{F}u(s) = h\mathcal{F}g(s) \tag{1.142}$$

with the solution

$$\mathcal{F}u(s) = (A + isM)^{-1} h\mathcal{F}g(s). \tag{1.143}$$

Because the L^2 norms of the error and the fourier transform of the error are identical, we need to approximate $(A + isM)^{-1}$ well on our subspace to have a good time dependent solution.

1.6 Reduced Basis Method

The space discretization of problem (1.79) usually leads to a large system of ordinary differential equations in time. Standard time stepping methods solve this system by solving linear equations for each step. Since solving these equations is the most expensive part in terms of computing recources and time, many numerical methods are based on reducing the number of equations one has to solve. Reduced basis methods rely on finding a good subspace for the approximation of the solution and project the system of ODEs onto this subspace. Let

$$M\frac{\partial u(t)}{\partial t} + Au(t) = hg(t) \tag{1.144}$$

be the (i.e. by finite element method in space) discretized problem, with $A, M \in \mathbb{R}^{n\times n}, u(t), h \in \mathbb{R}^n$. We want to find a subspaces V_m and a basis of V_m, $\{b_i\}_{i=1}^m$ such that the solution $u(t)$ can be well approximated in V_m. In section 1.5

we showed that approximating the solution u of (1.144) in the space $L^2(\Omega)$ is equivalent to approximating $\tilde{u}$

$$\tilde{u}(s) = (A + isM)^{-1} h\tilde{g}(s). \tag{1.145}$$

The standard method for creating the subspace V_m is to use (1.145) and build the basis from shifts $s_i \in \mathbb{C}$

$$b_i = (A + s_i M)^{-1} h. \tag{1.146}$$

Using real shifts s_i has the advantage of solving real, symmetric systems of equations in (1.146). In [5] it was shown that when using real shifts, the optimal error is at most twice the error of using shifts in $\mathbb{C}$, so we get the same convergence rate. In section 2 we discuss necessary optimality conditions and error estimates for the shifts s_i, which we use in section 3 to construct nested RB spaces.

1.7 Residuum

The reduced basis method allows us to compute the residuum for each time step in an efficient way. Let G be the projection matrix from the discretized space to the reduced basis subspace. The residuum of our problem is

$$r(t) = M\frac{\partial u(t)}{\partial t} + Au(t) - f(t). \tag{1.147}$$

On the assumption that the error of the solution of the reduced problem is neglectable (since we can compute it with high order) we can insert the RB-solution $u = G^T u_s(t)$. $u_s(t)$ is the solution of the reduced problem on the RB subspace

$$\frac{\partial u_s(t)}{\partial t} = M_s^{-1}(f_s(t) - A_s u_s(t)). \tag{1.148}$$

Inserting this into (1.147) leads to

$$r(t) = MG^T M_s^{-1}(f_s(t) - A_s u_s(t)) + AG^T u_s(t) - f(t) \tag{1.149}$$

using $f_s(t) = Gf(t)$ and $A_s = GAG^T$ we get

$$r(t) = (MG^T M_s^{-1} G - I)(f(t) - AG^T u_s(t)) \tag{1.150}$$

which can be computed for each time step. In the numerical examples in section 5 we show the applicability of this residuum, it converges with the same rate as the error.

1.8 Time stepping - Runge Kutta Methods

Because the reduced problem is small, higher order time stepping methods are feasible. Runge Kutta methods allow arbitrary order, stable methods and are easy to apply.

Definition 13 (Runge Kutta method and Butcher tableau)
A Runge Kutta method of stage number s is defined by

$$y_{n+1} = y_n + h\sum_{i=1}^{s} b_i k_i \tag{1.151}$$

$$k_i = f\left(t_n + c_i h, y_n + h\sum_{j=1}^{s} a_{ij}k_j\right), \qquad i = 1, \dots, s \tag{1.152}$$

where the coefficients are given by the methods Butcher tableau

$$\begin{array}{c|ccc} b_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ b_s & a_{s1} & \dots & a_{ss} \\ \hline & c_1 & \dots & c_s \end{array} \tag{1.153}$$

The concept of stability is very important for our problem because we get rapidly decaying solutions, having a time stepping method which doesn't overshoot is necessary.

Lemma 13 (Stability function)
The stability function of a numerical method is defined as the evolution function $R(z)$ of the method on the equation

$$y'(t) = \lambda y(t) \tag{1.154}$$
$$y(0) = 1 \tag{1.155}$$

where $R(h\lambda)$ is the solution after a numerical step of size h.
The stability function of a Runge Kutta method is

$$R(z) = 1 + zb^T(I - za)^{-1}e \tag{1.156}$$

where e is the vector of ones.

Proof:
See [4], Lemma 6.30

□

Definition 14 (A-Stability)
A numerical method for the equation

$$u'(t) - Au(t) = f(t) \tag{1.157}$$

is called A-stable if $|R(z)| \leq 1$ for all $h > 0$ with $z = h\lambda$, $\lambda \in \sigma(A)$

Definition 15 (L-Stability)
A numerical method is called L-stable if it is A-stable and $R(z) \to 0$ *for* $h \to \infty$.

The concept of A- and L- stability can be explained on the example of the definition of the stability function

$$y'(t) = \lambda y(t) \tag{1.158}$$
$$y(0) = 1 \tag{1.159}$$

with the solution $y(t) = e^{\lambda t}$. L-stable methods are usually required if the eigenvalue $\lambda < 0$. Then the norm of the solution is exponentially decaying, the A-stability claims, that the numerical solution is not growing and the L-stability assumtion results in our solution to tend to 0 as $t \to \infty$. In our case the eigenvalues of A are all negative, so we want L-stability to prevent growing numerical solutions. In our examples in section 5 we will use the third order L-stable Gauss Radau RK-method for the solution of the reduced problem.

Theorem 14
The 3 stage Gauss Radau Runge-Kutta method with the butcher tableau

$$\begin{array}{c|c} b & a \\ \hline & c \end{array} = \begin{array}{c|ccc} \frac{16-\sqrt{6}}{36} & \frac{88-7\cdot\sqrt{6}}{360} & \frac{296-169\cdot\sqrt{6}}{1800} & \frac{-2+3\cdot\sqrt{6}}{225} \\ \frac{16+\sqrt{6}}{36} & \frac{296+169\sqrt{6}}{1800} & \frac{88+7\sqrt{6}}{360} & \frac{-2-3\sqrt{6}}{225} \\ \frac{1}{9} & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \\ \hline & \frac{4-\sqrt{6}}{10} & \frac{4+\sqrt{6}}{10} & 1 \end{array} \tag{1.160}$$

is L-stable and of order 3.

Proof:
See [4]

□

2 Optimality Conditions for the Shifts

We want to consider two approaches for construcing good shifts, first we follow [6] to see when using the skeleton approximation, finding optimal shifts relates to solving the third Zolotarev problem in the complex plane. The adaptive algorithm of this paper uses this problem to get good snapshots. Secondly, in section 2.3, we derive the equivalence of reduced eigenvalues and shifts as a necessary optimality condition and use this condition to construct an adaptive method for the next snapshots.
Consider the discretized parabolic problem

$$Au(t) + M\frac{\partial u(t)}{\partial t} = f(t) \tag{2.1}$$

and let $f(t) = h \cdot g(t)$ with $h \in \mathbb{R}^n$, $g(t) \in L^2(\mathbb{R}^+)$. In this section we will follow the steps in [5] to reduce our subspace approximation problem to a rational function approximation problem. The fourier transfered problem states

$$A\mathcal{F}u(s) + isM\mathcal{F}u(s) = h\mathcal{F}g(s) \tag{2.2}$$

with solution

$$\mathcal{F}u(s) = (A + isM)^{-1}h\mathcal{F}g(s). \tag{2.3}$$

Let $u_m(t)$ be the reduced basis solution to equation (2.1),

$$A_m u_m(t) + M_m\frac{\partial u_m(t)}{\partial t} = h_m g(t) \tag{2.4}$$

and $V \in \mathbb{R}^{m\times n}$ the orthonormal projection matrix from $\mathbb{R}^n$ to the reduced basis subspace, with $A_m = VAV^T, M_m = VMV^T \in \mathbb{R}^{m\times m}$ and $h_m = Vh \in \mathbb{R}^m$ the projections of the matrices and vectors onto the reduced basis subspace. Then

$$\mathcal{F}u_m(s) = (A_m + isM_m)^{-1}h_m\mathcal{F}g(s) = V(A + isM)^{-1}V^TVh\mathcal{F}g(s). \tag{2.5}$$

Let $\Omega^T := \Omega \times \mathbb{R}^+$ and $\Omega^F := \Omega \times \mathbb{R}$ the corresponding frequency domain, with Parseval's identity, the L^2-error of the RB method is

$$\|u - u_m\|^2_{L^2(\Omega^T)} = \|\mathcal{F}u - \mathcal{F}u_m\|^2_{L^2(\Omega^F)} = \tag{2.6}$$

$$\int_{\mathbb{R}} \left\|(A + isM)^{-1}h\mathcal{F}g(s) - V(a + isM)^{-1}V^TVh\mathcal{F}g(s)\right\|^2_{L^2(\Omega)} ds \tag{2.7}$$

2.1 Rational function approximation

Lemma 15
Let $q(\lambda) := \prod_{l=1}^{m}(\lambda+s_l)$, with s_l being the shifts of the RB-space with projection matrix V. Then for all polynomials p with degree less than m there holds

$$\frac{p(A)}{q(A)}\varphi = V\frac{p(a)}{q(a)}V^T\varphi. \tag{2.8}$$

Proof:
Let

$$\mathcal{P} := \{v \in \mathbb{R}^n : v = \frac{p(A)}{q(A)}\varphi \text{ with } \dim p \leq m-1\} \tag{2.9}$$

$$\mathcal{Q} := \{v \in \mathbb{R}^n : v = V\frac{p(a)}{q(a)}V^T\varphi \text{ with } \dim p \leq m-1\} \tag{2.10}$$

Note that $\dim \mathcal{P} = \dim \mathcal{Q} = m$. Let $v_i = (A+s_iM)^{-1}\varphi$, $\tilde{v}_i = V(a+s_iM)^{-1}V^T\varphi$. Since $v_i \in \mathcal{P}$ and $\tilde{v}_i \in \mathcal{Q}$ these vectors span the spaces $\mathcal{P}$ and $\mathcal{Q}$. So to proof the lemma we have to show that $v_i = \tilde{v}_i$. This is fulfilled since $\tilde{v}_i$ is the Galerkin approximation of v_i and $v_i \in \mathcal{V}$.

□

Theorem 16
The error of the approximation $\hat{f} = VfV^T$ of a function f onto the RB subspace fullfilles: For any p with degree less than m and $\varphi \in L^2(\Omega)$ there holds

$$\left\|f(A)\varphi - Vf(a)V^T\varphi\right\|_{L^2(\Omega)} \leq 2 \max_{\lambda\in[\lambda_{min},\lambda_{max}]}\left|f(\lambda) - \frac{p(\lambda)}{q(\lambda)}\right| \|\varphi\|_{L^2(\Omega)}. \tag{2.11}$$

with $q(\lambda)$ defined in Lemma 15 and $[\lambda_{min}, \lambda_{max}]$ the spectral interval of A.

Proof:
Using Lemma 15, the triangle inequality and the eigenvalue decomposition of the symmetric matrices A and $a = VAV^T$ there holds

$$\left\|f(A)\varphi - Vf(a)V^T\varphi\right\|_{L^2(\Omega)} \overset{\text{L } 15}{=} \left\|f(A)\varphi - \frac{p(A)}{q(A)}\varphi + V\frac{p(a)}{q(a)}V^T\varphi - Vf(a)V^T\varphi\right\|_{L^2(\Omega)} \tag{2.12}$$

$$\leq \left\|f(A)\varphi - \frac{p(A)}{q(A)}\varphi\right\|_{L^2(\Omega)} + \left\|Vf(a)V^T\varphi - V\frac{p(a)}{q(a)}V^T\varphi\right\|_{L^2(\Omega)} \tag{2.13}$$

$$= \left\|\sum_{i=1}^{n}\left(f(\lambda_i) - \frac{p(\lambda_i)}{q(\lambda_i)}\right)(z_i,\varphi)\right\|_{L^2(\Omega)} + \left\|\sum_{i=1}^{m}\left(f(\theta_i) - \frac{p(\theta_i)}{q(\theta_i)}\right)(\phi_i,\varphi)\right\|_{L^2\Omega)} \tag{2.14}$$

With λ_i and θ_i the respective eigenvalues and z_i and ϕ_i the corresponding normalized eigenvectors of $M^{-1}A$ and $m^{-1}a$. Because a and m are the projection matrix of A and M on a subspace there holds $\lambda_{\min} \leq \theta_i \leq \lambda_{\max}$ for all $1 \leq i \leq m$ and

$$\left\| f(A)\varphi - Vf(a)V^T\varphi \right\|_{L^2(\Omega)} \leq 2 \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} \left| f(\lambda) - \frac{p(\lambda)}{q(\lambda)} \right| \|\varphi\|_{L^2(\Omega)} . \quad (2.15)$$

□

Theorem 16 reduces the problem of finding the optimal shifts to a rational approximation problem for the function $\frac{1}{\lambda+is}$ for λ in the spectrum of $M^{-1}A$ and $s \in \operatorname{supp}(\mathcal{F}g)$.

Corollary 17
With $q(\lambda) = \prod_{i=1}^{m}(\lambda + s_i)$ and the subspace $\mathcal{V}$ chosen as in (6.1), the RKSM error is bound by

$$\|u - u_m\|_{L^2(\Omega^T)} \leq \quad (2.16)$$

$$\sqrt{\int_{\mathbb{R}} \min_{p \in \mathbb{P}^{m-1}(\mathbb{R})} \max_{\lambda \in [\lambda_{min}, \lambda_{max}]} \left| \frac{1}{\lambda + is} - \frac{p(\lambda)}{q(\lambda)} \right|^2 |\mathcal{F}g(s)|^2 ds} \, \|\varphi\|_{L^2(\Omega)} \quad (2.17)$$

Proof:
Use theorem 16 with $f(\lambda) = (\lambda + is)^{-1}$ in (2.7) and take the minimal polynomial.

□

2.2 Skeleton approximation

The skeleton approximation of functions in two variables was introduced in [11]. In the following section we follow the investigation of this approximation for the function $\frac{1}{x+y}$ in [10]. It is defined as

$$f_{\text{skel}}(\lambda, s) = \begin{pmatrix} \frac{1}{\lambda+s_1} & \cdots & \frac{1}{\lambda+s_m} \end{pmatrix} \tilde{M}^{-1} \begin{pmatrix} \frac{1}{s+\lambda_1} \\ \vdots \\ \frac{1}{s+\lambda_m} \end{pmatrix} \quad (2.18)$$

with

$$\tilde{M} = \begin{pmatrix} \frac{1}{\lambda_1+s_1} & \cdots & \frac{1}{\lambda_1+s_m} \\ \vdots & \ddots & \vdots \\ \frac{1}{\lambda_m+s_1} & \cdots & \frac{1}{\lambda_m+s_m} \end{pmatrix} \quad (2.19)$$

Lemma 18
The skeleton approximation is a decomposition of a seperable function of rank m. For the seperable function

$$g(x,y) = \sum_{k=1}^{m} u_k(x)v_k(y) \tag{2.20}$$

there holds

$$g(x,y) = g_{skel}(x,y) \tag{2.21}$$

Sketch of Proof:
This problem relates to the skeleton decomposition for matrices. Let $A \in \mathbb{R}^{n\times n}$ be of rank $r < n$, choose $C \in \mathbb{R}^{n\times r}$ as r rows of A and $R \in \mathbb{R}^{r\times n}$ as r columns of A and let $\hat{A} \in \mathbb{R}^{r\times r}$ be their intersection matrix. If $\hat{A}$ is nonsingular then $A = C\hat{A}^{-1}R$. The Lemma states the same for the continous setting. For a rigorous proof see [8].

.

The more interesting property of the skeleton approximation is how good it is if g is not seperable. Lemma 18 suggests that the best approximation in the space of seperable functions of rank m can be achieved. The following theorem and corollary show that the problem of finding good interpolation points leads to the Zolotarev problem in the complex plane.

Theorem 19
For $f(\lambda, s) = \frac{1}{\lambda+s}$ the relative error is

$$(\lambda + s)(f(\lambda,s) - f_{skel}(\lambda,s)) = \prod_{i=1}^{m} \frac{\lambda-\lambda_i}{\lambda+s_i}\frac{s-s_i}{s+\lambda_i} = \frac{r(\lambda)}{r(-s)} \tag{2.22}$$

with

$$r(x) = \prod_{i=1}^{m} \frac{x-\lambda_i}{x+s_i}. \tag{2.23}$$

Proof:
Define $\tilde{f}_{\text{skel}}$ as

$$\tilde{f}_{\text{skel}}(\lambda,s) := \frac{1}{\lambda+s}\left(1 + \prod_{i=1}^{m} \frac{\lambda-\lambda_i}{\lambda+s_i}\frac{s-s_i}{s+\lambda_i}\right) \tag{2.24}$$

To proof the theorem we need to show that $f_{\text{skel}} = \tilde{f}_{\text{skel}}$. First note that f_{skel} and $\tilde{f}_{\text{skel}}$ interpolate $\frac{1}{\lambda+s}$ in $(\lambda_i)_{i=1}^m$ and $(s_i)_{i=1}^m$. So if we can show that $\tilde{f}_{\text{skel}}$

has seperable rank m, we can use Lemma 18 to proof the result.

$$\tilde{f}_{\text{skel}} = \frac{1}{\prod_{i=1}^{m}(\lambda+s_i)(s+\lambda_i)} \left(\underbrace{\frac{1}{\lambda+s}\left(\prod_{i=1}^{m}(\lambda+s_i)(s+\lambda_i) + \prod_{i=1}^{m}(\lambda-\lambda_i)(s-s_i)\right)}_{:=p} \right) \tag{2.25}$$

p is a rational function which vanishes in its singularity $\lambda = -s$ and is therefore a polynomial and p can be rewritten as

$$p(\lambda, s) = \sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\lambda^i s^j \tag{2.26}$$

From that follows that f_{skel} is of seperable rank m since

$$p(\lambda, s) = \sum_{i=1}^{m} \lambda^i \sum_{j=1}^{m} p_{ij} s^j \tag{2.27}$$

and the result follows from Lemma 18.

□

Corollary 20
With $r(x) = \frac{x-\lambda_i}{x+s_i}$ the RKSM error on the subspace $\mathcal{V}$ is bounded by

$$\|u - u_m\|_{L^2(\Omega^T)} \leq \frac{1}{\lambda_{min}} \|g\|_{L^2(\mathbb{R})} \|\varphi\|_{L^2(\Omega)} \frac{\max_{\lambda\in[\lambda_{min},\lambda_{max}]} |r(\lambda)|}{\min_{s\in i\mathbb{R}} |r(s)|} \tag{2.28}$$

Proof:
With theorem 19 we get

$$\|u - u_m\|_{L^2(\Omega^T)} \leq \max_{\lambda\in[\lambda_{\min},\lambda_{\max}]} |r(\lambda)| \sqrt{\int_{\mathbb{R}} \left|\frac{1}{\lambda_{\min}+is}\right|^2 \left|\frac{\mathcal{F}g(s)}{r(-is)}\right|^2 ds} \|\varphi\|_{L^2(\Omega)}. \tag{2.29}$$

Taking the minimum on the denominators and then using Parseval's identity gives us the stated result.

□

Remark: Using the skeleton approximation may cost us the sharpness of the bound, since from Corollary 17 it follows that the best approximation to

$$\frac{1}{\lambda+s} \simeq f_{\text{approx}}(\lambda, s) = \frac{\tilde{f}(\lambda, s)}{q(\lambda)} \tag{2.30}$$

with $\tilde{f}(\cdot, s) \in \mathbb{P}^{m-1}$ for each $s \in i\mathbb{R}$ is optimal. However the skeleton approximation leads to an easy way to compute good snapshots for the RKSM subspace.

Remark: Here we neglect a possible usage of the function g. If we want to simulate just some specific frequencies, a better estimate would be i.e. to have the minima in the denominator only on the interval of the desired frequencies. So this estimate can be viewed as a general reduced basis subspace independent of g.

2.3 L^2 - Optimality Conditions

We want to use the results of Gugercin, Antoulas and Beattie in [9], where they analyse $\mathcal{H}^2$ optimality conditions for the transfer functions of linear SISO (single input single output) dynamical systems. Their necessary optimality condition for the selection of the snapshots in the reduced basis model will be used later in our adaptive algorithm. For this result we will need a basic complex analysis result, the residuum theorem.

Definition 16 (Residuum)
Let U be a simply connected domain, μ a point and $f : U \setminus \{\mu\} \to \mathbb{C}$ holomorphic. The residuum of f in μ is

$$res[f, \mu] := \frac{1}{2\pi i} \oint_{\partial U} f(z)dz \tag{2.31}$$

Theorem 21 (Residuum Theorem)
Let U be a simply connected open subset of $\mathbb{C}$, $\mu_1, ..., \mu_n$ points of U and f a function which is holomorphic on $U \setminus \{\mu_1, ..., \mu_n\}$. If γ is a positive orientated simple closed curve with the points μ_i in it's interior, then

$$\frac{1}{2\pi i} \int_\gamma f(z)dz = \sum_{k=1}^{n} res[f, \mu_k] \tag{2.32}$$

Lemma 22
Let μ_i be a simple pole of f, then

$$res[f, \mu_i] = \lim_{s \to \mu_i} (s - \mu_i) f(s) \tag{2.33}$$

Definition 17 (Transfer Function)
Considering the frequency domain problem (1.142), the transfer function $G \in L^2(\mathbb{C}, \mathbb{C}^n) =: \mathcal{H}$ *in the frequency domain is defined as*

$$\mathcal{F}u(s) = G(s)\mathcal{F}g(s) \tag{2.34}$$

$$G(is) = (A + isM)^{-1}h \tag{2.35}$$

G Hilbert space with with scalar product

$$\langle G, H\rangle_{\mathcal{H}} := \frac{1}{2\pi}\int_{\mathbb{R}} \overline{G(is)}M_0 H(is)ds \tag{2.36}$$

where M is defined by

$$\int_\Omega \varphi_i\varphi_j\,dx, \tag{2.37}$$

with $\{\varphi_i\}_{i=1}^n$ the basis functions of our discrete space.
In [1], Antoulas optained an expression for $\|G\|_{\mathcal{H}_2}$, that can be used for our space $\mathcal{H}$ as well.

Lemma 23
Suppose that $G(s)$ *has singularities at* $\lambda_1, ..., \lambda_n$ *and* $H(s)$ *at* $\mu_1, ..., \mu_m$*, both sets in the open right half plane. Then*

$$\langle G, H\rangle_{\mathcal{H}} = \sum_{k=1}^{m} res[G(-s)MH(s), \mu_k] \tag{2.38}$$

furthermore if μ_i *is a simple pole then*

$$[res[G(-s)MH(s), \mu_i] = G(-\mu_i)Mres[H(s), \mu_i] \tag{2.39}$$

Proof:
For $R \in \mathbb{R}^+$ define the right half circle

$$C_R := \{z \in \mathbb{C} : z = iw, w \in [-R, R]\} \cup \left\{z \in \mathbb{C} : z = Re^{i\xi}, \xi \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]\right\}. \tag{2.40}$$

For R large enough, all μ_i are enclosed by C_R. Using the Residue Theorem (Theorem 21) we get

$$\langle G, H\rangle_{\mathcal{H}} = \frac{1}{2\pi}\int_{\mathbb{R}} G(-is)MH(is)ds = \tag{2.41}$$

$$= \lim_{R\to\infty}\frac{1}{2\pi i}\int_{C_R} G(-s)MH(s)ds \tag{2.42}$$

$$= \sum_{k=1}^{n} \text{res}[G(-s)MH(s), \mu_k] \tag{2.43}$$

If μ_i is a simple pole of $H(s)$ then μ_i is a simple pole of $G(-s)MH(s)$ as well, therefore

$$\begin{aligned} \text{res}[G(-s)MH(s), \mu_i] &= \lim_{s \to \mu_i} (s - \mu_i) G(-s)MH(s) && (2.44) \\ &= G(-\mu_i)M \lim_{s \to \mu_i} (s - \mu_i) H(s) && (2.45) \\ &= G(-\mu_i)M\text{res}[H(s), \mu_i]. && (2.46) \end{aligned}$$

□

Corollary 24
If $G(s)$ has simple poles at $\lambda_1, ...\lambda_n \in \mathbb{R}^+$ then

$$\|G\|_{\mathcal{H}} = \sqrt{\sum_{k=1}^{n} G(-\lambda_k)M res[G(s), \lambda_k]} \tag{2.47}$$

Lemma 25
Define the space $\mathcal{M}(\mu), \mu \in \mathbb{R}^r$ of all transfer functions with simple poles exactly at $\mu_1, ..., \mu_r \in \mathbb{R}^+$. $G_r \in \mathcal{M}(\mu)$ fulfilles necessary local optimality conditions if

$$\langle G - G_r, H \rangle_{\mathcal{H}} = 0 \tag{2.48}$$

for all $H \in \mathcal{M}(\mu)$.

Proof:
This follows directly from $\mathcal{M}(\mu)$ beeing a Hilbert space, since $\mathcal{M}(\mu)$ is a closed subspace of $\mathcal{H}$.

□

This lets us introduce our necessary optimality condition for the reduced order transfer function

Theorem 26
A necessary local optimality condition for G_r beeing a reduced order transfer function of rank r of G with simple poles at $\mu_1, ..., \mu_r$ is that

$$G_r(-\mu_i) = G(-\mu_i) \tag{2.49}$$

for all $i = 1, ..., r$.

Proof:
For all $H \in \mathcal{M}(\mu)$, G_r fullfilles

$$0 = \langle G - G_r, H \rangle_{\mathcal{H}} = \sum_{i=1}^{r} (G(-\mu_i) - G_r(-\mu_i))\text{res}[H(s), \mu_i]. \tag{2.50}$$

Since res$[H(s), \mu_i]$ is arbitrary the result follows.

□

Corollary 27
For the RB subspace to fullfill necessary optimality conditions, the shifts $\{s_i\}_{i=1}^m$ *must match the eigenvalues* $\{\lambda_i\}_{i=1}^m$ *of* $M_m^{-1}A_m$.

Proof:
From definition 17 there follows, that the poles $\mu_1, ..., \mu_r$ of G_r are the negative eigenvalues of the reduced system

$$\mu_i = -\lambda_i \tag{2.51}$$

Using theorem 26 we get that our transfer function has to fullfill

$$(A + \lambda_i M)^{-1}h = (A_m + \lambda_i M_m)^{-1}h \tag{2.52}$$

This is exactly the case if $(A + \lambda_i M)^{-1}h \in \mathcal{V}$. So

$$s_i = \lambda_i \tag{2.53}$$

□

Remark: Note that $\{\mu_i\}_{i=1}^r$ are not initially known. One way to create the subspace is to start with an arbitrary one, i.e. with geometrically distributed shifts and iteratively compute new subspaces which interpolate G in the negative poles of the old G_r. This technique was analysed in [9]. It has significant disadvantages. First, for one order r the subspace needs to be computed multiple times with expensive large sparse matrices operations and secondly, the lower order subspaces are not included in higher order ones, so you cannot iteratively extend the space dimension if the accuracy is not good enough. The second disadvantage is shared by all optimal subspaces, therefore we want to consider not optimal, but asymptotic optimal spaces which can be computed iteratively.

Remark: Theorem 26 states that $\lambda_i = s_i$ is a necessary optimality condition. Using this in Corollary 20 gives $|r(s)| = 1$ for $s \in i\mathbb{R}$ and there holds

$$\|u - u_m\|_{L^2(\Omega^T)} \leq \frac{1}{\lambda_{\min}} \|g\|_{L^2(\mathbb{R})} \|\varphi\|_{L^2(\Omega)} \max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |r(\lambda)|. \tag{2.54}$$

So as an additional condition, one can minimize

$$\max_{\lambda \in [\lambda_{\min}, \lambda_{\max}]} |r(\lambda)| \tag{2.55}$$

3 Nested Reduced Subspaces

The main problem with optimal reduced order subspaces is that they are not nested. Because the shifts of an n dimensional optimal subspace are usually not included in the shifts of the optimal $n+1$ dimensional space, a lot of expensive big matrix operations have to be done for each dimension (even if we can somehow get the optimal shifts without any expensive operations). Since we do not know the needed dimension of our reduced order space we want to extend it iteratively until the error is small enough. The two approaches we want to compare are derived from the two results of section 2. First the new shift s_{n+1} is obtained by solving

$$s_{j+1} = \max_{s\in[\lambda_{\min},\lambda_{\max}]} \left|\frac{1}{r_n(-s)}\right| \text{ with } r(s) = \prod_{j=1}^{m} \frac{s-\lambda_j}{s+s_j} \tag{3.1}$$

proposed in [6], which is derived from the RKSM residual of Corollary 20, with λ_j and s_j the eigenvalues and shifts of the reduced m dimensional system. Second, we use the necessary optimality condition of Theorem 26 to derive a method based on the statistical Kolmogorov Smirnow test to match the distribution of the shifts to the distribution of the eigenvalues.

3.1 The Kolmogorov-Smirnow Test Method

The Kolmogorov-Smirnow test is a statistical test on equal distribution of two data sets. We want to minimize the statistic of a slightly modified version of this test to fit the shifts to the eigenvalues.

Definition 18
The empirical distribution function of a sample $\{\alpha_i\}_{i=1}^m$ is defined as

$$F_\alpha(x) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{\alpha_i \le x}, \tag{3.2}$$

or in words, the number of sample values lower than x divided by the sample size.

Definition 19 (Two Sample Kolmogorov-Smirnow-Test)
The Kolmogorov Smirnow statistic $D_{\alpha\beta}$ for two (not necessarily equal sized) samples $\{\alpha_{i=1}^n\}$ and $\{\beta_{i=1}^m\}$ is

$$D_{\alpha\beta} = \sup_x |F_\alpha(x) - F_\beta(x)| \tag{3.3}$$

A method for finding the next shift s_{m+1} is to find the minimum

$$\min_{s_{m+1}} = D_{s,\lambda} \tag{3.4}$$

with $s_1, ..., s_m$ being the old shifts, $s = \{s_1, ..., s_{m+1}\}$ and $\lambda = \{\lambda_1, ...\lambda_m\}$ the eigenvalues of the reduced system.
This method has some problems. Firstly, it doesn't give a unique result. Consider the example $s = \{1, 3\}$ and $\lambda = \{2, 4\}$. The minimum will be $D_{s\lambda} = 1$, but this minimum is aquired for any $s \in \mathbb{R}$, so we get no information about the next shift for that setup. The second problem is that the eigenvalue of the next step can differ from the eigenvalues of the previous one. Let us consider some eigenvalues which are close together, $\lambda = \{101, 102, 103\}$ which we got from using the same shifts $s = \{101, 102, 103\}$. In the next step the eigenvalues change slightly and a new small eigenvalue is added, $\lambda = \{1, 98, 99, 100\}$. Then the small eigenvalue should be more important, but the supremum is smallest at $D_{s,\lambda} = 3$ for $s_4 \in [100, \infty)$. Another problem arises, if there is a concentration of eigenvalues at a point x_1 and a single exposed value at x_2. If the number of eigenvalues at x_1 is larger than the number of shifts at x_1 plus one, the new shift will be at x_1. In Appendix 6.3 we explain that it is possible that other eigenvalues at x_2 are missed, because having big shifts will lead to too many big eigenvalues and vice versa.
This is why we will consider the L^2 norm of the difference of the empirical distribution functions, instead of the C^∞ norm. This gives exposed eigenvalues more weight and makes the choice for the new shift (almost) unique.

Definition 20
The L^2-KS statistic $E_{\alpha\beta}$ for two (not necessary equally sized) samples $\{\alpha_{i=1}^n\}$ and $\{\beta_{i=1}^m\}$ is

$$E_{\alpha\beta} = \sqrt{\int_\Omega |F_\alpha(x) - F_\beta(x)|^2 \, dx} \tag{3.5}$$

Definition 21 (Modified KS-Method)
Our modified KS-method is:
Find $s_{n+1} \in \mathbb{R}^+$ minimizer of

$$\min_{s_{n+1} \in \mathbb{R}^+} E_{s \cup \{s_{n+1}\}, \lambda}. \tag{3.6}$$

4 Additional Stabilization Techniques

In this chapter we want to discuss some additional techniques which are useful for solving the parabolic electromagnetic equation

$$\frac{\partial \sigma u}{\partial t} + \operatorname{curl}\left(\mu^{-1}\operatorname{curl}(u)\right) = f \tag{4.1}$$

4.1 Avoiding numerical instabilities by transformation of the source term

Let the source term of $\mathbf{J}$ be in the non conducting part of the domain. Define the matrices A and M as in section 1.3:

$$A_{ij} := \int_\Omega \mu^{-1}\operatorname{curl}(\varphi_i)\operatorname{curl}(\varphi_j)\, dx \tag{4.2}$$

$$M_{ij} := \int_\Omega \sigma\varphi_i\varphi_j\, dx \tag{4.3}$$

with $\mu > 0$, $\sigma \geq 0$. The right hand side h is defined as

$$h_i := \int_\Omega \mathbf{J}\varphi_i\, dx \tag{4.4}$$

For stability of the parabolic problem, h must be bounded in the M^{-1}-norm. This is only true if $\mathbf{J}$ has support in the conducting ($\sigma > 0$) regions of our domain only. To avoid the instability we transform the problem.
Let $u(t) = g(t)\cdot u_0(\mathbf{x}) + u_1(t,\mathbf{x})$ with u_0 being the solution of

$$\int_\Omega \mu^{-1}\operatorname{curl}(u_0)\operatorname{curl}(v) + \frac{\sigma}{\tau}u_0 v\, dx = \int_\Omega hv\, dx \tag{4.5}$$

with τ being a characteristic time length. The mass term is introduced to obtain a divergence free u_0. The equation for u_1 is

$$\int_\Omega \sigma\frac{\partial u_0}{\partial t}v + \sigma\frac{\partial u_1}{\partial t}v\, dx + \int_\Omega \mu^{-1}\operatorname{curl}(u_0)\operatorname{curl}(v) + \mu^{-1}\operatorname{curl}(u_1)\operatorname{curl}(v)\, dx \tag{4.6}$$

$$= \int_\Omega f(t)v\, dx \tag{4.7}$$

Using equation (4.5) the red terms cancel by adding the green term and the new equation for u_1 is

$$\int_\Omega \sigma \frac{\partial u_1}{\partial t} v \, dx + \int_\Omega \mu^{-1} \mathrm{curl}\,(u_1)\,\mathrm{curl}\,(v)\; dx = \int_\Omega (-g'(t) + \frac{1}{\tau} g(t)) \sigma u_0 v \, dx. \quad (4.8)$$

Now the divergence free source term only has support in the conducting domains. In table 4.1 the eigenvalues of a 10 and a 30 dimensional reduced subspace are compared. Note that the condition number of the small problem gets huge, which makes the problem numerically instable. For the transformed problem the condition number stays about the same.

20 dim no trans	30 dim no trans	20 dim with trans	30 dim with trans
1.95947289e-01	2.10047073e+01	1.95695396e-01	1.95644509e-01
7.82974279e-01	2.73784021e+02	3.89371266e-01	2.38667095e-01
1.93787152e+00	5.30874470e+02	7.80847242e-01	4.67867657e-01
4.62113296e+00	8.36620953e+02	1.78840639e+00	7.84457914e-01
1.02480189e+01	9.73680483e+02	3.22387721e+00	1.78427235e+00
2.24998426e+01	1.94997973e+03	5.79450818e+00	2.04529485e+00
5.01608197e+01	3.14153248e+03	1.09141669e+01	3.21436691e+00
8.69875120e+01	4.99464561e+03	2.04747803e+01	5.56740766e+00
2.11404023e+02	8.70629788e+03	4.02332591e+01	9.99115652e+00
4.47746366e+02	1.72952052e+04	7.86688876e+01	1.72027581e+01
8.91462639e+02	3.31938884e+04	1.42849203e+02	2.53247391e+01
1.78805940e+03	5.05666805e+04	2.77333361e+02	4.12533297e+01
3.44004485e+03	5.59727432e+04	5.46430523e+02	7.60508859e+01
6.47510040e+03	6.10054565e+04	1.06125099e+03	9.83109436e+01
1.16620345e+04	6.55373127e+04	2.14542666e+03	1.67811541e+02
2.71980266e+04	9.47971734e+04	4.40168454e+03	2.95340558e+02
5.00148351e+04	1.48644118e+05	9.06320070e+03	5.08341879e+02
6.98020688e+04	1.64982606e+05	2.57185639e+04	7.47692466e+02
1.90010001e+05	4.65132124e+05	5.66114319e+04	1.21401175e+03
3.80993390e+12	3.16121792e+05	1.05612187e+05	1.69815635e+03
	1.25410044e+06		2.20313305e+03
	2.08922750e+06		3.58597804e+03
	3.17310051e+06		5.18029895e+03
	5.73427778e+06		8.67824545e+03
	1.14229067e+07		1.68474955e+04
	2.36064745e+07		3.13138055e+04
	5.75226236e+07		5.92777995e+04
	1.12919589e+08		8.60327481e+04
	2.02206711e+08		1.99720601e+05
	1.40882134e+21		1.75594147e+06

Table 4.1: Eigenvalues of 10-dim and 30-dim reduced basis subspace without and with transformation of the source term.

4.2 Orthogonalization to Gradients

As mentioned in section 4.1 it is important to have a divergence free source term. When approximating prescribed (divergence free) currents on the finite

element space, the geometry can cause some numerical errors and lead to a non divergence free term. Divergence free is equivalent to L^2-orthogonal to gradients as can be seen by applying Gauss divergence theorem. Because the high order basis functions on our finite element contain by construction the gradients of the H^1 functions (see section 1.2), which can be removed from the finite element space, only an orthogonalization to the lowest order Nedelec basis functions has to be done. We will need the following lemma to construct the orthogonalization scheme.

Lemma 28
Let λ_i be the vertex basis functions (hat functions) and φ_{ij} the Nedelec edge basis functions. Let $\mathcal{N}(i)$ be the vertex patch of i, so all neighboring vertices of vertex i. There holds

$$\nabla\lambda_i = \sum_{j\in\mathcal{N}(i)} \varphi_{ij}. \tag{4.9}$$

Proof:
For all i there holds

$$\sum_{j\in\mathcal{N}(i)} \varphi_{ij} = \sum_{j\in\mathcal{N}(i)} (\nabla\lambda_i\lambda_j - \nabla\lambda_j\lambda_i) = \tag{4.10}$$

$$\nabla\lambda_i \underbrace{\left(\sum_{j\in\mathcal{N}(i)} \lambda_j\right)}_{=1-\lambda_i} - \underbrace{\left(\sum_{j\in\mathcal{N}(i)} \nabla\lambda_j\right)}_{=\nabla(1-\lambda_i)=-\nabla\lambda_i} \lambda_i = \tag{4.11}$$

$$\nabla\lambda_i(1-\lambda_i) + \nabla\lambda_i\lambda_i = \nabla\lambda_i \tag{4.12}$$

□

Let G be the mapping of Nedelec basis functions to the corresponding vertex basis functions and u a basis function of our reduced order subspace of H^{curl}. We want to find $\varphi \in H_0^1(\Omega)$ with $\nabla\varphi\cdot\nabla w = u\cdot\nabla w$ for all $w \in H_0^1(\Omega)$. Then

$$\int_\Omega u\cdot\nabla w_k dx = (GMu)_k \tag{4.13}$$

with M beeing the mass matrix on Ω. So the equation

$$\int_\Omega \nabla\varphi\cdot\nabla w dx = \int_\Omega u\cdot\nabla w dx \tag{4.14}$$

for all $w \in H_0^1(\Omega)$ is

$$A_L\varphi = GMu \tag{4.15}$$

with A_L beeing the laplacian matrix on Ω. $\nabla\varphi$ transformed to H^{curl} basis functions is $G^T\varphi$. Therefore the orthogonalized $\tilde{u}$ is

$$\tilde{u} = u - G^T A_L^{-1} GMu. \tag{4.16}$$

We will use the orthogonalized basis functions for our subspace to eliminate zero eigenvalues in the reduced order curl-curl matrix.

Remark: For a right hand side the orthogonalization would be

$$\tilde{h} = h - MG^T A_L^{-1} Gh. \tag{4.17}$$

Instead of orthogonalizing the basis functions one could also orthogonalize the numerically disturbed right hand side.

5 Numerical Examples

5.1 Heat Equation

The first example is a basis reduction for a dynamic heat equation in 2 dimensions. There is a small region with high conductivity and the heat source inside the bigger enclosing region.
We want to solve

$$\frac{\partial u}{\partial t} - \alpha \Delta u = f, \text{ in } \Omega \tag{5.1}$$

$$u = 0, \text{ on } \partial\Omega \tag{5.2}$$

on $\Omega = [-1, 1]^2$, with $f = \sin(50 \cdot 2\pi t)$ in $[0, 0.5]^2$ and 0 elsewhere, and $\alpha = 100$ on $[0, 0.5]^2$ and 1 elsewhere. Figure 5.1 contains the L^2 norm of the error and the relative error of the n-dimensional subspace reduced basis solution against a solution on the big FEM space. In the left bottom the eigenvalue distribution and the shift distribution are plotted on a logarithmic scale for the two methods. One can see that the distribution of the eigenvalues matches the distribution of the shifts better with the KS-method than with the method using the minimizing function (3.1), which is called "Druskin" in the plots. The error convergence is slightly better when using the KS-method as well. In the right bottom we see that the residuum derived in section 1.7 is a good error indicator. The "true" solution on the big FEM space was computed with an relative error of about 10^{-6}, the KS-method reaches this error with a subspace dimension of 30, method (3.1) with a space dimension of 35. A hierarchical geometrical distribution on the spectral range, independent of the distrubution of the eigenvalues reaches the desired accuracy for a 42 dimensional subspace. The residuum improves further with increasing subspace dimension, so further convergence can be expected. The most expensive part in the computation of this problem is the solution of linear systems. To obtain this accuracy with an implicit euler method we had to solve the problem with a time step size of $2 \cdot 10^{-8}$, meaning that at a frequency of 50 hertz the simulation of one wavelength needs the solution of one million linear systems. With our KS-RB-method we can compute arbitrary time intervals with only 30 solutions of linear problems of the same size.

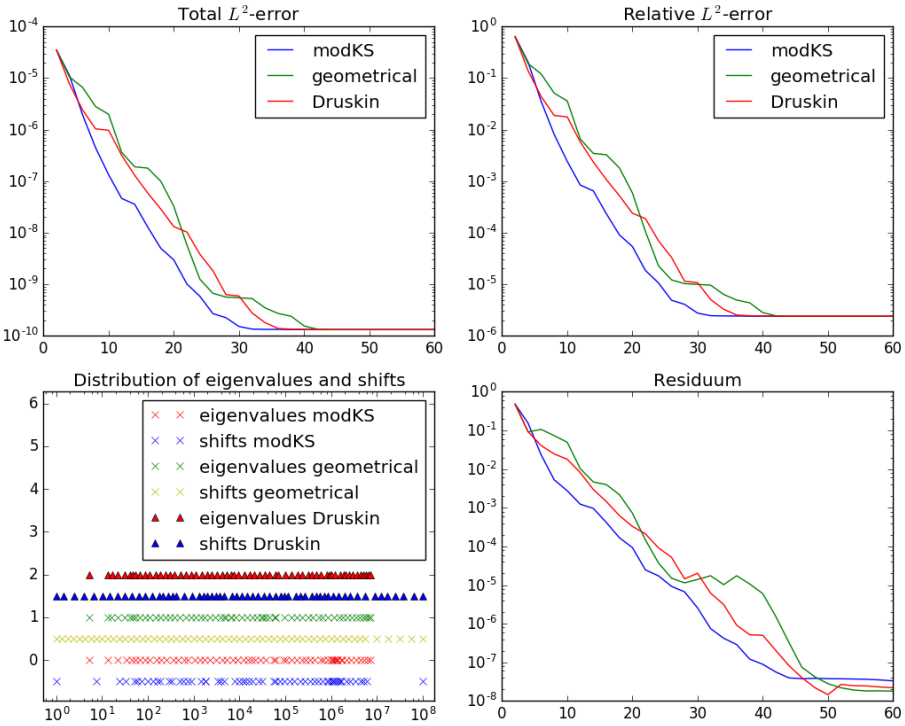

Figure 5.1: KS-method vs minimization of (3.1)(labeled Druskin) for the 2D heat equation example

5.2 Electromagnetic Equation

The main problem of this thesis is the parabolic electromagnetic equation. As a test example we want to calculate the electric fluxes arising in a conducting plate when exposed to a dynamic magnetic field. In figure 5.2 the geometry is shown. A magnet with changing magnetic field is put above a conducting plate. The equations for this problem are

$$\sigma \frac{\partial \mathbf{u}}{\partial t} + \operatorname{curl}\left(\mu^{-1}\operatorname{curl}\left(\mathbf{u}\right)\right) = f, \text{ in } \Omega \tag{5.3}$$

$$tr_\tau \mathbf{u} = 0, \text{ on } \partial\Omega \tag{5.4}$$

With $\mu = \mu_0\mu_r$, $\mu_0 = 1.257 \cdot 10^{-6}$, $\mu_r = 2000$ in the plate, else $\mu_r = 1$. $\sigma = 2 \cdot 10^6$ in the plate, else 10^{-6} (regularization factor). $f = \begin{pmatrix} 0 \\ 0 \\ -\sin(50 \cdot 2\pi t) \end{pmatrix}$ in the magnet, else $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Figure 5.3 displays the results. The better fitting of the shifts to the eigenvalues of the KS-method is again represented in a faster convergence. Our reference solution on the big space is of relative accuracy of 10^{-4}. The KS-method obtains this accuracy at a subspace dimension of 18, method (3.1) at a dimension of 22. For this accuracy we needed a step size of 10^{-7} with an implicit euler method on the full space , so at a frequency of 50 hertz, 200.000 solutions to large linear problems for one wavelength. With the RB method we could reduce that number to 18 independent of the time interval.

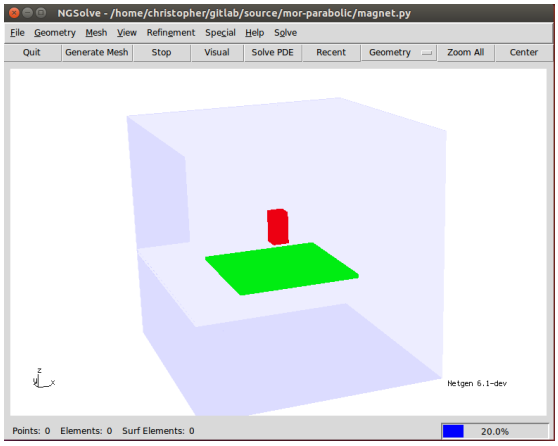

Figure 5.2: Geometry for problem (5.3)

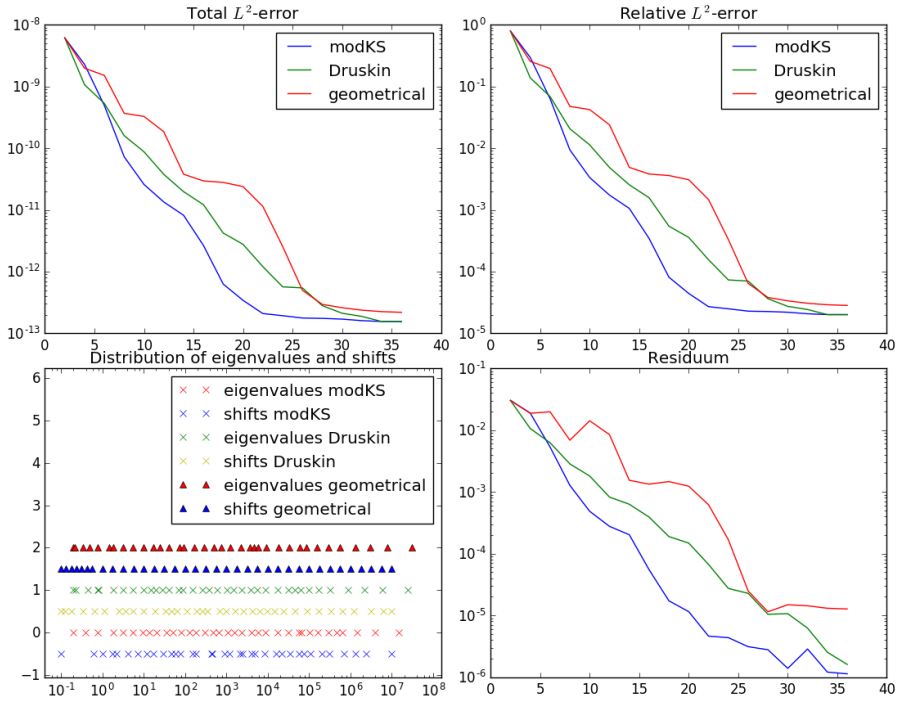

Figure 5.3: KS-method vs minimization of (3.1)(labeled Druskin) for the 3D electromagnetic example

6 Appendix

6.1 Construction of orthonormal, divergence free basis

Let V_m be the reduced basis subspace

$$\mathcal{V}_m = \text{span } \{((A + s_1 M)^{-1} h), ..., ((A + s_m M)^{-1} h)\} \tag{6.1}$$

An orthonormal basis (with respect to M) V of $\mathcal{V}_m$ can be constructed with the following algorithm, the orthogonalization to gradients is explained in section 4.2.

Algorithm 1 Reduced subspace basis

for $i = 0$ **to** $i < m$ **do**
  $v_i = (A + s_i M)^{-1} h$
  $v_i \mathrel{-}= G^T A_L^{-1} G M v_i$ {orthogonalization to gradients}
  for $j = 0$ **to** $j < i$ **do**
    $v_i \mathrel{-}= (v_j^T M v_i) v_j$ {Gram Schmidt orthogonalization}
  end for
  $v_i \mathrel{*}= 1/(v_i^T M v_i)$ {normalization}
end for

$$V = \begin{pmatrix} v_1^T \\ \vdots \\ v_m^T \end{pmatrix}$$

6.2 Geometrical Distribution of Reduced Eigenvalues

Consider the time dependent initial problem

$$M \frac{\partial \mathbf{u}}{\partial t} + A\mathbf{u} = 0 \text{ in } \Omega \times (0, T) \tag{6.2}$$

$$\mathbf{u}(0) = \mathbf{u}_0 \tag{6.3}$$

$$\mathbf{u} = 0 \text{ on } \partial\Omega \times [0, T). \tag{6.4}$$

Let $\{\lambda_i\}_{i=1}^n$ the eigenvalues and $\{\mathbf{v}_i\}_{i=1}^n$ the eigenvectors of $M^{-1}A$. The solution to this problem is

$$\mathbf{u}(t) = \sum_{i=1}^{n} e^{-\lambda_i t}(\mathbf{u}_0, \mathbf{v}_i)_M \tag{6.5}$$

If $(\mathbf{u}_0, \mathbf{v}_i)_M$ is not neglectable on all parts of the spectrum $[\lambda_{\min}\lambda_{\max}]$ then the optimal reduced eigenvalues would be geometrically distributed, such that $\{e^{\tilde{\lambda}_i}\}$ is uniformly distributed. This is true for the problem with a source term as well.

6.3 Shift - Eigenvalue Correlation

Our reduced subspace basis $\{b_i\}$ is defined by

$$b_i = (A + s_i M)^{-1} h = \sum_{i=1}^{n} \frac{1}{\lambda_i + s_i}(f, e_i)e_i \tag{6.6}$$

with λ_i and e_i the eigenvalues and eigenvectors of A with respect to M. If s_i is small then for λ_1 small and λ_2 large there holds

$$\frac{1}{\lambda_1 + s_i} \approx \frac{1}{2\lambda_1} \tag{6.7}$$

$$\frac{1}{\lambda_2 + s_i} \approx \frac{1}{\lambda_2}. \tag{6.8}$$

Because $\frac{1}{2\lambda_1} \gg \frac{1}{\lambda_2}$ mainly the small eigenvectors are put into the reduced system. For s_i large there holds

$$\frac{1}{\lambda_1 + s_i} \approx \frac{1}{s_i} \tag{6.9}$$

$$\frac{1}{\lambda_2 + s_i} \approx \frac{1}{2s_i}. \tag{6.10}$$

Now $\frac{1}{s_i} \approx \frac{1}{2s_i}$, so the small and the large eigenvectors are both put into the reduced system with the same weight, so a uniform distribution of reduced eigenvalues can be expected if only large shifts are used. In section 6.2 we discuss why a geometric distribution is optimal for arbitrary right hand side, so there are too many large eigenvectors in the reduced system then. As a result, using only small shifts leads to a reduced system with too many small shifts and vice versa.

List of Tables

4.1 Eigenvalues of 10-dim and 30-dim reduced basis subspace without and with transformation of the source term. 34

List of Figures

5.1 KS-method vs minimization of (3.1)(labeled Druskin) for the 2D heat equation example 38
5.2 Geometry for problem (5.3) 40
5.3 KS-method vs minimization of (3.1)(labeled Druskin) for the 3D electromagnetic example 40

Bibliography

[1] A.C. Antoulas, *Approximation of large-scale dynamical systems*, Advances in Design and Control DC-06, SIAM, Philadelphia, 2005

[2] D. N. Arnold, R. S. Falk, R. Winther, *Differential Complexes and Stability of Finite Element Methods I. The De Rham Complex*, Volume 142, IMA Volumes in Mathematics and its Applications pp 23-46

[3] A. Bossavit *Mixed Finite Elements and the Complex of Whitney Forms*, The Mathematics of Finite Elements and Applications, Ac. Press (London), 1988

[4] Peter Deuflhard and Folkmar Bornemann, *Numerische Mathematik 2: Gewohnliche Differentialgleichungen*, Walter de Gruyter Verlag, Berlin, 3. Auflage, Edition, 2008.

[5] V. Druskin, L. Knizhnerman, M. Zaslavsky, *Solution of large scale evolutionary problems using rational Krylov subspaces with optimized shifts*, SIAM J. SCI COMPUT., Vol. 31, No. 5, pp. 3760-3780 (2009)

[6] V. Druskin, C. Lieberman, M. Zaslavsky, *On adaptive choice of shifts in rational Krylov subspace reduction of evolutionary problems*, SIAM J. SCI COMPUT., Vol. 32, No. 5, pp. 2485-2496 (2010)

[7] E. Grimme, *Krylov Projection Methods for Model Reduction*, Ph.D. thesis, University of Illinois at Urbana-Champain, Urbana, IL, 1997

[8] A. Gorodetsky, S. Karaman, Y. Marzouk, *Function-Train: A continuous analogue of the tensor-train decomposition*, arXiv:1510.09088 [math.NA]

[9] S. Gugercin, A.C. Antoulas, C. Beattie, $\mathcal{H}_2$ *model reduction for large-scale linear dynamical systems*, SIAM J. Matrix Anal. Appl

[10] I. V. Oseledets, *Lower bounds for seperable approximations of the Hilbert kernel*, Mat. Sb., 198 (2007), pp. 425-432 (translated into English)

[11] E. E. Tyrtyshnikov, *Mosaic-skeleton approximations*, Calcolo, 33 (1996), pp. 47-57

[12] S.Zaglmayr, *High Order Finite Element Methods for Electromagnetic Field Computation*, PhD thesis Johannes Kepler University Linz, 2006