JOHANNES KEPLER
UNIVERSITÄT LINZ
Netzwerk für Forschung, Lehre und Praxis

A Multiharmonic Solver for Nonlinear Parabolic Problems

Diplomarbeit

zur Erlangung des akademischen Grades

Diplomingenieur

in der Studienrichtung

Technische Mathematik

Angefertigt am *Institut für Numerische Mathematik*

Betreuung:

O. Univ.-Prof. Dipl.-Ing. Dr. Ulrich Langer

Eingereicht von:

Michael Kolmbauer

Linz, Oktober 2009

Abstract

This diploma thesis is concerned with the numerical solution of nonlinear, scalar potential parabolic partial differential equations in one dimension. Due to a time-harmonic excitation, the resulting periodicity of the solution can be capitalised to switch from the time domain to the frequency domain by approximating the solution by a truncated Fourier series and solving the resulting system of partial differential equations in the Fourier coefficients. Solving this time-independent system by the finite element method yields a large-scale, coupled system of nonlinear equations. This nonlinearity is dealt with by a linearisation in terms of the Newton iteration. Hence, the construction of an appropriate preconditioner for the arising Jacobi system is inevitable for the efficiency of the entire solver. In order to tackle this challenge, we adapt the preconditioned GMRES method, proposed by Cai an Xu (1992) for general non-symmetric and indefinite problems to our Jacobi system.
In addition, we present a spectral analysis for the Cai-Xu preconditioner applied to study the behaviour of the linear problem in order to study the associated convergence rates for two special cases.
We perform numerical tests and report the results, both for the linear and nonlinear case.

Zusammenfassung

Die vorliegende Diplomarbeit befasst sich mit der Entwicklung effizienter numerischer Methoden zur Lösung von nichtlinearen parabolischen partiellen Differentialgleichungen im örtlich eindimensionalen Fall. Eine harmonische Anregung erlaubt es uns, die daraus resultierende Periodizität der Lösung zu nutzen, um vom Zeitbereich in den Frequenzbereich zu wechseln. Durch eine Approximation der periodischen Lösung durch eine abgebrochene Fourierreihe erhalten wir ein System von zeitunabhängigen partiellen Differentialgleichungen in den Fourierkoeffizienten. Die Finite-Elemente-Diskretisierung dieses Systems liefert ein hochdimensionales gekoppeltes nichtlineares Gleichungssystem, dessen Nichtlinearität durch eine Newton Iteration behandelt wird. Um dieses Gleichungssystem effektiv lösen zu können, ist eine geeignete Vorkonditionierungsstrategie für das in jedem Newton Schritt zu lösende Jakobi-System von großer Bedeutung. Hierfür verfolgen wir die Strategie von Cai und Xu, welche einen hierfür passenden GMRES Präkonditionierer entwickelt haben.
Zusätzlich führen wir eine Spektralanalyse des durch die Cai-Xu Strategie präkonditionierten linearen Systems durch, um die Konvergenzeigenschaft der dazugehörigen GMRES Methode quantitativ genau analysieren zu können. Dieses Vorgehen wird für zwei spezielle Parameterkonstellationen durchgeführt.
Abschließend werden die Ergebnisse dieser numerischen Untersuchungen sowohl für den linearen als auch für den nichtlinearen Fall präsentiert.

Contents

Chapter 1

Introduction

This thesis deals with the numerical solution of spatially one-dimensional, nonlinear, parabolic and scalar potential partial differential equations with harmonic excitations. For example, a two-dimensional eddy current problem can be reduced to such a nonlinear parabolic problem. The goal of this work is to develop and analyse new efficient numerical methods for solving such nonlinear time-harmonic problems. In fact, we use a combination of well-known techniques, namely a multiharmonic approach, the finite element method (FEM) and Newton's method for the time approximation, for the space discretisation and for the linearisation respectively. Especially we focus on developing efficient solvers for the large and fully coupled systems of linear finite element equations arising at each step of the Newton iteration.

The mathematical problem

The main steps of our approach are the following: Taking advantage of the periodicity of the solution, we reduce the time-dependent problem to a space-dependent system of partial differential equations by applying a mulitharmonic ansatz. The resulting set of nonlinear equations is then solved by Newton's method, where at each step a large and fully coupled system of linear equations appears. In fact this system is positive definite and non-symmetric, but it can be decomposed into a symmetric and positive definite (SPD) part and a block skew-symmetric part. This set of equations is discretised by the finite element method. The resutling finite element equations are solved by a preconditioned GMRES method. The preconditioner is constructed by the special additive Schwarz technique proposed by Cai and Xu [4].

Review of previous work

The idea of combining a multiharmonic ansatz and the finite element method was pursued for example by Yamada and Bessho in [19]. This approach is called multiharmonic FEM or harmonic-balanced FEM (HBFEM) and has been used by many engineers in different applications, for example, see [8, 9, 12].
In [1, 2] the first rigorous mathematical analysis of the multiharmonic approach to the nonlinear eddy current problem was provided. So, this thesis follows these techniques to obtain similar results for nonlinear parabolic partial differential equations.
The approach of this thesis is strongly aligned with the following two papers.

- In [4], Cai and Xu provided a preconditioning technique for solving non-symmetric or indefinite problems via the GMRES method. This theory is used to solve the resulting system of linear equations.

- In [5], Copeland and Langer applied a multiharmonic approach to a nonlinear parabolic equation, solved the resulting nonlinear problem by Newton's method and constructed an almost optimal solver for the occuring systems of linear equations.

The task of this thesis

In this thesis we want to get an idea of the behaviour of our multiharmonic solver by analysing the one-dimensional problem. This happens in two steps:

- Applying a harmonic approach to the linear parabolic equation and elaborating some analysis for a preconditioned GMRES solver.

- Extending the linear case to the nonlinear one by using the multiharmonic finite element method and providing a solver for the nonlinear system of equations.

The organisation of this thesis

- Chapter 2 - *Problem Formulation and Analysis*:
  Starting from a nonlinear, parabolic scalar potential equation with homogeneous Dirichlet boundary conditions and inhomogeneous initial conditions, we derive the variational formulation and provide some analysis for existence and uniqueness.

- Chapter 3 - *GMRES and Cai-Xu Preconditioning*:
  We present the idea and the main convergence result of the GMRES method for solving non-symmetric or indefinite problems. Moreover a preconditioning technique, developed by Cai an Xu in [4], is summarised.

- Chapter 4 - *Multiharmonic Approach*:
  We apply a multiharmonic ansatz to the original parabolic partial differential equation and end up with a time-independent system of equations.

- Chapter 5 - *The Finite Element Method*:
  After the basic concepts of the finite element method are recapitulated, the finite element method is applied to the linear problem. The resulting linear system of equations is solved by applying the GMRES preconditioner and some analysis is prepared.

  Finally we treat the nonlinear problem and focus on an iterative solution of the nonlinear equations in terms of the Newton iteration.

- Chapter 6 - *Numerical Integration*:
  The integrals appearing from discretisation have to be approximated. Thus some error analysis of appropriate quadrature rules is given.

- Chapter 7 - *Numerical Results*:
  Numerical studies, both for the linear and nonlinear case, manifest the theoretical results achieved in the previous chapters.

- Chapter 8 - *Conclusion and Outlook*:
  The results are summarised and open problems and possible continuation are discussed.

Notations

Concerning notations, we use a subscript h for finite element functions (e.g. u_h) and underline type for real valued vectors (e.g. $\underline{u}_h$). In case of operators and their associated matrices, we do not distinguish in their displaying formats. The individual denotation as real valued matrix or operator can be interpreted from the context.

Abbreviations

FEM	Finite Element Method
GMRES	Generalised Minimal Residual Method
SPD	Symmetric and Positive Definite

Chapter 2

Problem Formulation and Analysis

In this chapter we want to state a one-dimensional nonlinear, parabolic, scalar potential equation as our model problem. After enunciating the classical formulation, we derive the space-time variational formulation, which is the starting point for our discretisation afterwards. As an alternative approach the line variational formulation is given and existence and uniqueness of the solution are proven.

2.1 Classical Formulation

As our model problem we consider a one-dimensional parabolic potential equation with homogeneous Dirichlet boundary and a non-homogeneous initial condition. This nonlinear problem is in classical formulation given as follows: Find $u(x,t) \in C^{2,1}(\Omega_T) \cap C([a,b]\times[0,T))$, such that

$$\begin{aligned} \alpha \frac{\partial u}{\partial t}(x,t) - \frac{\partial}{\partial x}\left(\nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}\right)(x,t) &= f(x,t), \qquad && \forall (x,t) \in \Omega_T, \\ u(a,t) = u(b,t) &= 0, && \forall t \in (0,T), \\ u(x,0) &= u_0(x), && \forall x \in [a,b], \end{aligned} \tag{2.1}$$

where the space-time domain is given by $\Omega_T = (a,b) \times (0,T)$. Furthermore we have the following assumptions imposed on the given input parameters:

- $\nu(\cdot) \in C^1(\mathbb{R}_0^+ \longrightarrow \mathbb{R})$, $0 < \underline{\nu} \leq \nu(s) \leq \overline{\nu}$ for all $s \in \mathbb{R}_0^+$, with given positive constants $\underline{\nu}$ and $\overline{\nu}$. Furthermore $s \mapsto s\nu(s)$ is Lipschitz and monotone for $s \geq 0$,
- $f(x,t) \in C(\Omega_T)$,
- $u_0(x) \in C[a,b]$,
- $\alpha \in \mathbb{R}^+$ is a given positive constant.

2.2 Space-Time Variational Formulation

In order to derive the variational formulation of (2.1), we perform the usual approach. After setting up appropriate spaces, the space-time variational formulation can be derived.

Sobolev Spaces

Thus we need some setting in the Sobolev spaces:

$$H^{1,1}(\Omega_T) := \left\{u \in L^2(\Omega_T) : \partial_t u, \partial_x u \in L^2(\Omega_T)\right\} = H^1(\Omega_T),$$
$$H^{1,0}(\Omega_T) := \left\{u \in L^2(\Omega_T) : \partial_x u \in L^2(\Omega_T)\right\}.$$

These spaces are equipped with the norms

$$\|u\|^2_{H^{1,1}(\Omega_T)} := \|u\|^2_{L^2(\Omega_T)} + \|\partial_t u\|^2_{L^2(\Omega_T)} + \|\partial_x u\|^2_{L^2(\Omega_T)} \quad \text{and}$$
$$\|u\|^2_{H^{1,0}(\Omega_T)} := \|u\|^2_{L^2(\Omega_T)} + \|\partial_x u\|^2_{L^2(\Omega_T)}$$

respectively.

Setting up the Spaces for the Variational Formulation

We choose the following Sobolev spaces as the trial space

$$H^{1,1}_0(\Omega_T) := \left\{u \in H^{1,1}(\Omega_T) : u(a,t) = u(b,t) = 0 \text{ for almost all } t \in (0,T)\right\}$$

and the test space

$$H^{1,0}_0(\Omega_T) := \left\{u \in H^{1,0}(\Omega_T) : u(a,t) = u(b,t) = 0 \text{ for almost all } t \in (0,T)\right\}.$$

Space-Time Variational Formulation

Multiplying (2.1) by a test-function $v(x,t) \in H^{1,0}_0(\Omega_T)$, integrating over the space-time domain and performing integration by parts in the principle term of the space derivative gives the space-time variational setting of (2.1): Find $u \in H^{1,1}_0(\Omega_T)$ such that

$$\begin{aligned}\frac{2}{T}\int_0^T\int_a^b \left[\alpha\frac{\partial u}{\partial t}(x,t)v(x,t) + \nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}(x,t)\frac{\partial v}{\partial x}(x,t)\right] dxdt \\ = \frac{2}{T}\int_0^T\int_a^b f(x,t)v(x,t)dxdt, \quad \forall v(x,t) \in H^{1,0}_0(\Omega_T),\end{aligned} \tag{2.2}$$

and that fulfills the initial condition in a weak sense, namely

$$\|u(\cdot,t) - u_0(\cdot)\|_{L^2(a,b)} \longrightarrow 0 \text{ for } t \longrightarrow 0.$$

The initial condition is meaningfull, since the trace is defined. Now, the given input data fulfill the following weaker conditions:

- $\nu(\cdot) \in C^1(\mathbb{R}^+_0 \longrightarrow \mathbb{R})$, $0 < \underline{\nu} \leq \nu(s) \leq \overline{\nu}$ for all $s \in \mathbb{R}^+_0$, with given positive constants $\underline{\nu}$ and $\overline{\nu}$. Furthermore $s \mapsto s\nu(s)$ is Lipschitz and monotone for $s \geq 0$,
- $f(x,t) \in L^2(\Omega_T)$,
- $u_0(x) \in L^2(a,b)$,
- $\alpha \in \mathbb{R}^+$ is a given positive constant.

The variational formulation (2.2) is the starting point of various discretisation methods in space and time. In this work, we will discretise in time by a multiharmonic approach and in space by the usual finite element approximation. This will be done step by step, but anyhow the discretisation procedure will fit into this framework.

Remark 2.1. *Note, that we just perform partial integration in the space variable. Of course it would also be possible to perform partial integration in the time variable as well, but this will lead to different ansatz- and test-spaces (see, e.g. [11]).*

2.3 Existence and Uniqueness

We now turn to the analysis of nonlinear variational problems obtained by the line variational formulation of (2.1), that is slightly more general than the space-time variational formulation (2.2).

2.3.1 Theoretical Background

Before we can quote the main theorem of existence and uniqueness, we need some basic definitions. Firstly, we need the notion of an evolution triple.

Definition 2.2 (Evolution Triple [20]). *An evolution triple $V \subset H \subset V^*$ fulfills the following properties.*

1. *V is a real, seperable and reflexive Banach space.*
2. *H is a real, seperable Hilbert space.*
3. *The embedding $V \subset H$ is continuous and V is dense in H.*

In order to quote the main theorem of existence and uniqueness, we have to clarify the following definitions.

Definition 2.3. *Let V be a real Banach space with norm $\|\cdot\|$ and let $A : V \longrightarrow V^*$ be an operator from V into its dual V^*. Then*

- *A is called* monotone *iff*

$$\langle A(u) - A(v), u - v \rangle \geq 0, \quad \forall u, v \in V,$$

- *A is called* coercive *iff*

$$\lim_{\|u\| \longrightarrow \infty} \frac{\langle A(u), u \rangle}{\|u\|} = \infty,$$

- *A is called* hemicontinuous *iff the real function*

$$t \longrightarrow \langle A(u + tv), w \rangle$$

is continuous on $[0, 1]$ for all $u, v, w \in V$.

The main theorem of existence and uniqueness of nonlinear parabolic equations reads as follows.

Theorem 2.4 ([21]). *Let $V \subset H \subset V^*$ be an evolution triple and let $A : V \longrightarrow V^*$ be a hemicontinuous, monotone and coercive operator. Suppose that A is bounded, i.e.*

$$\exists c > 0 : \|A(u)\|_{V^*} \leq c \|u\|_V , \quad \forall u \in V.$$

Let $u_0 \in H$ and $F \in L^2((0,T), V^*)$. Then the abstract initial value problem: Find $u \in L^2((0,T), V)$ with a weak derivative $u' \in L^2((0,T), V^*)$, such that

$$\begin{aligned} u'(t) + A(u(t)) &= F(t) \quad \textit{for almost all } t \in (0,T) \textit{ and} \\ u(0) &= u_0 \in H \end{aligned}$$

is uniquely solvable.

2.3.2 Application to Parabolic Problems

We show the existence and uniqueness in $V = H_0^1(a,b)$ and therefore the dual space V^* is given by $V^* = H^{-1}(a,b)$ and $H = L^2(a,b)$. Firstly, we fix some time $t \in (0,T)$. For almost everywhere in $(0,T)$, find $u \in V$, such that

$$\left\langle \alpha \frac{\partial u}{\partial t}, v \right\rangle + \langle A(u), v \rangle = \langle F, v \rangle , \quad \forall v \in V, \tag{2.3}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality product $\langle \cdot, \cdot \rangle_{V^* \times V}$. The operators

$$A : V \longrightarrow V^* \quad \text{and} \quad F : V \longrightarrow V^*$$

are defined by the relations

$$\begin{aligned} \langle A(u), v \rangle &:= \int_a^b \nu \left(\left| \frac{\partial u}{\partial x} \right| \right) \frac{\partial u}{\partial x}(x,t) \frac{\partial v}{\partial x}(x) dx \quad \forall u, v \in V, \\ \langle F, v \rangle &:= \int_a^b f(x,t) v(x) dx \quad \forall v \in V. \end{aligned}$$

So almost everywhere in $(0,T)$ the problem, we are concerned with, reads as follows. Find $u \in L^2((0,T), V)$ with a weak derivative $u' \in L^2((0,T), V^*)$, such that

$$\begin{aligned} \alpha u'(t) + A(u(t)) &= F(t) \quad \text{and} \\ u(0) &= u_0. \end{aligned} \tag{2.4}$$

Theorem 2.5. *Let the following assumptions be fulfilled*

1. *$0 < \underline{\nu} \leq \nu(s) \leq \overline{\nu}$ for all $s \in \mathbb{R}_0^+$, with given positive constants $\underline{\nu}$ and $\overline{\nu}$. Furthermore $s \mapsto s\nu(s)$ is Lipschitz and monotone for $s \geq 0$.*
2. *$u_0 \in L^2(a,b)$ and $f \in L^2((0,T), V^*)$ and $\alpha \in \mathbb{R}^+$.*

Then we have a unique solution $u \in L^2((0,T), V)$ with a weak derivative $u' \in L^2((0,T), V^*)$, that solves the initial value problem(2.4)*.*

Proof. We have to verify the assumptions of Theorem 2.4.

1. Evolution Triple: We have the evolution triple $V \subset H \subset V^*$ with

$$V = H_0^1(a,b), \quad H = L^2(a,b) \quad \text{and} \quad V^* = H^{-1}(a,b).$$

2. Hemicontinuous: Hemicontinuity follows directly from the continuity of A.

3. Monotonicity: The operator $A : V \longrightarrow V^*$ is monotone. Since the mapping $s \mapsto s\nu(s)$ is monotone we can conclude, that

$$\langle A(u) - A(v), u - v \rangle \geq 0, \quad \forall u, v \in V.$$

Indeed,

$$\langle A(u) - A(v), u - v \rangle = \int_a^b \left(\nu \left(\left| \frac{\partial u}{\partial x} \right| \right) \frac{\partial u}{\partial x} - \nu \left(\left| \frac{\partial v}{\partial x} \right| \right) \frac{\partial v}{\partial x} \right) \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial x} \right) dx \geq 0.$$

4. Coerciveness: The function ν is positive. Therefore, we obtain

$$\exists c > 0 : \langle A(u), u \rangle \geq c \left\| u \right\|_V^2, \quad \forall u \in V.$$

Indeed,

$$\langle A(u), u \rangle = \int_a^b \nu \left(\left| \frac{\partial u}{\partial x} \right| \right) \frac{\partial u}{\partial x} \frac{\partial u}{\partial x} dx \geq \underline{\nu} \int_a^b \left(\frac{\partial u}{\partial x} \right)^2 dx = \underline{\nu} \left| u \right|_{H^1(a,b)}^2 .$$

With the norm equivalence

$$\left\| u \right\|_{H^1(a,b)} \cong \left| u \right|_{H^1(a,b)}, \quad \forall u \in H_0^1(a,b),$$

the coerciveness of the operator A follows.

5. Boundedness: Since ν is bounded from above, we can conclude

$$\exists c > 0 : \left\| A(u) \right\|_{V^*} \leq c \left\| u \right\|_V, \quad \forall u \in V.$$

Since the given data also satisfy the assumptions of Theorem 2.4, we have a unique solution $u \in L^2((0,T), V)$ with a weak derivative $u' \in L^2((0,T), V^*)$ of the initial value problem (2.4). □

So we gain existence and uniqueness of our initial value problem (2.4).

Chapter 3

GMRES and Cai-Xu Preconditioning

In [14], Saad and Schultz introduced the Generalised Minimal Residual Method (GMRES) that is suited to solve non-symmetric or indefinite systems of equations. The corresponding convergence results of the method can be found in [6]. In this thesis, the method is used to solve the resulting non-symmetric finite element equations. Therefore, the algorithm and the convergence of the GMRES-method is shortly summarised in the first part of this chapter. In the second part, we present a preconditioning technique for the GMRES method, that is based on an additive Schwarz decomposition, which was developed by Cai and Xu in [4].

3.1 GMRES

We want to solve an equation of the form

$$Gu = g, \tag{3.1}$$

where G is a linear operator, defined on a finite-dimensional vector space V, and the right-hand side g is a given vector in V. Let $[\cdot, \cdot]_V$ be an inner product on V that induces the norm $\|\cdot\|_V$.

The general idea of the method is summarised in the following abstract algorithm, for details see [14].

Algorithm 1. The main steps of GMRES algorithm are

1. *Start:* Choose $u_0 \in V$ and compute the initial residual $r_0 = g - Gu_0$.

2. *Iterate:* For $m = 1, \dots$ until convergence, compute

$$z_m = \min_{z \in \mathfrak{K}_m(r_0)} \|g - G(u_0 + z)\|_V$$

   where $\mathfrak{K}_m(r_0)$ is the Krylov subspace given by

$$\mathfrak{K}_m(r_0) = \left\{ r_0, Gr_0, \dots, G^{m-1} r_0 \right\}.$$

3. *Compute the new approximate solution:* $u_m = u_0 + z_m$.

Now, for the case that G is positive definite and real, we have the following convergence theorem.

Theorem 3.1 (Saad, Schultz [14]). *Let us assume that*

$$\underline{c} = \inf_{u \neq 0} \frac{[u, Gu]_V}{[u, u]_V}$$

is positive. Then the GMRES method converges, and, at the m-th iteration, the residual $r_m = g - Gu_m$ *is bounded as*

$$\|r_m\|_V \leq \left(1 - \frac{\underline{c}^2}{\overline{c}^2}\right)^{m/2} \|r_0\|_V \,,$$

where

$$\overline{c} = \sup_{u \neq 0} \frac{\|Gu\|_V}{\|u\|_V}.$$

Proof. For a proof see [6]. □

The constants $\underline{c}$ and $\overline{c}$ can be interpreted in the following way. Let G^{sym} be the symmetric part of G. Then in fact we have

$$\underline{c} = \lambda_{\min}(G^{\text{sym}}) \quad \text{and} \quad \overline{c}^2 = \lambda_{\max}(G^T G).$$

So $\underline{c}$ estimates the minimal eigenvalue of the symmetric part G^{sym} and $\overline{c}$ the norm $\|G\|_V$ of the operator G.

3.2 Preconditioned GMRES for non-SPD problems

In [4] Cai and Xu introduced a preconditioning technique for non-symmetric or indefinite problems, based on an additive Schwarz preconditioner on two grids. The main idea of this technique is to solve a non-symmetric and positive definite part of G only on a coarse subspace of V.

Abstract Settings

Let V be a finite-dimensional Hilbert space with an inner product $(\cdot, \cdot)$ and the corresponding norm $\|\cdot\| = (\cdot, \cdot)^{1/2}$. We want to solve an equation of the form

$$A_M u = f, \tag{3.2}$$

where $A_M = A + M$ is the sum of a symmetric and positive definite operator A and a non-symmetric operator M. It is important to mention, that after choosing a basis of V, the operator equation (3.2) is equivalent to a system of linear equations.

Assumption 1 (A is symmetric and positive definite). The mapping $A : V \to V$ is symmetric and positive definite with respect to $(\cdot, \cdot)$. So, $(A\cdot, \cdot) = (\cdot, \cdot)_A$ defines an inner product on V and induces the norm $\|\cdot\|_A = (A\cdot, \cdot)^{1/2}$, that is sometimes called A-energy norm. Now, there exists a constant $C_1 > 0$, such that

$$\|u\| \leq C_1 \|u\|_A \quad u \in V.$$

Assumption 2 (M is bounded)**.** The mapping $M : V \to V$ satisfies the inquality

$$|(Mu, v)| \leq C_2 \|u\| \|v\|_A \quad \forall u, v \in V$$

for some $C_2 = \text{const} > 0$.

Assumption 3 (Restriction to subspace V_0)**.** There exists a subspace $V_0 \subset V$, such that $\forall u \in V \quad \exists! u_0 \in V_0 :$

$$(A_M u_0, v_0) = (A_M u, v_0) \quad \forall v_0 \in V_0.$$

This setting defines a projection $P_0 : V \longrightarrow V_0$ such that $P_0 u = u_0$. To this projection some parameter

$$\delta_0 = \sup_{v \in V} \frac{\|(I - P_0)v\|}{\|v\|_A} \tag{3.3}$$

is linked. This parameter δ_0 describes the approximation of the space V by the coarse subspace V_0.

Furthermore, we also need to define the following two operators.
Firstly, we define the restriction $A_0 : V_0 \longrightarrow V_0$ of the linear operator A_M to the coarse subspace V_0 by

$$(A_0 u_0, v_0) = (A_M u_0, v_0), \quad \forall u_0, v_0 \in V_0.$$

Assumption 3 guarantees the invertibility of A_0.
Secondly, we define an ortho-projection $Q_0 : V \longrightarrow V_0$ from the space V to the coarse subspace V_0 by

$$(Q_0 u, v_0) = (u, v_0), \quad \forall u \in V, v_0 \in V_0.$$

Now, based on these definitions, Cai and Xu propose a preconditioner for A_M as follows

$$B_M = A_0^{-1} Q_0 + \beta B, \tag{3.4}$$

where B is a symmetric positive definite preconditioner for A, and β is some scalling parameter depending on the choice of B. The right choice of the balancing paramter β is necessary, since both parts of the preconditioner should be balanced somehow.
In order to compute convergence rates, we need the following theorem to classify the constants of Theorem 3.1.

Theorem 3.2 (Cai, Xu [4])**.** *Let Assumption 1,Assumption 2 and Assumption 3 be valid. Then there exist positive constants ϵ, γ, β and μ, depending on*

$$\lambda_0 = \lambda_{\min}(BA) \quad \textit{and} \quad \lambda_1 = \lambda_{\max}(BA),$$

such that: if $\epsilon \geq \delta_0$,

$$\|B_M A_N M u\|_A \leq \mu \|u\|_A \quad \forall u \in V$$

and

$$(B_M A_M u, u)_A \geq \gamma (u, u)_A \quad \forall u \in V,$$

where the generic constants are given by

$$\gamma = \frac{\lambda_0^2}{8 C_1^2 C_2^2 \lambda_1^2}, \quad \mu = 1 + C_2 \varepsilon + \beta C_1 C_2 \lambda_1,$$

$$\beta = \frac{\lambda_0}{2 C_1^2 C_2^2 \lambda_1^2}, \quad \epsilon = \frac{\lambda_0}{2 \lambda_1 C_2 \sqrt{C_1^2 C_2^2 + 1}}.$$

Proof. For a proof see [4]. □

Next we consider (3.2) and precondition it with (3.4). This yields the following equation of interest:

$$B_M A_M u = B_M f.$$

The choice of the bilinear form as $[\cdot,\cdot]_V = (\cdot,\cdot)_A$ together with Theorem 3.1 and Theorem 3.2 yields the convergence rate estimate

$$\|r_m\|_A \leq \left(1 - \frac{\gamma^2}{\mu^2}\right)^{m/2} \|r_0\|_A \, . \tag{3.5}$$

for the GMRES method.
So summarising the main idea of this section: If the non-symmetric and positive definite part M isn't too large (this is controlled by C_2 in Assumption 2), and if δ_0 is sufficiently small (this is controlled by the coarse space V_0) and if we have a good preconditioner for the symmetric and positive definite part A, we can always construct a good preconditioner for the operator A_M. In fact, there exists a lot of theory to construct efficient preconditioners for A, e.g. by domain decomposition techniques (see [18]) or by multigrid methods (see e.g. [3]) or similar.

Chapter 4

Multiharmonic Approach

The goal of this chapter is to perform a discretisation in time. Instead of using the semi-discretisation strategies like the vertical method of lines with FEM and solving the resulting system of ordinary differential equations by a time-stepping method, we discretise (2.1) in time and solve the resulting system of elliptic partial differential equations. Since we assume a time-harmonic excitation of f of the form

$$f(x,t) = f^c(x)\cos(\omega t) + f^s(x)\sin(\omega t), \tag{4.1}$$

we also have information about the expected solution u. In the linear case (i.e. $\nu = \nu(x)$), the solution u can be expressed in terms of the same base frequency ω. Due to the nonlinearity of ν (i.e. $\nu = \nu(|\nabla u|)$), the solution u depends on higher fonics as well, but still is periodic. However, approximating the solution by a multiharmonic ansatz seems to be quite promising. The arrangement in this chapter is the following: A harmonic ansatz is applied to the linear case and a mulitharmonic ansatz to the nonlinear problem. This approach reduces the original problem (2.1) to a system of coupled ordinary differential equations in the Fourier coefficients. The chapter closes by stating the resulting variational formulation, which is the starting point of a discretisation in space by the finite element method.
Additionally, on the base of the performed computations we build a bridge to the space-time variational problem stated in Chapter 2.

4.1 Steady State Solution

In many practical applications, we aren't really interested in the behaviour of u near the initial time, but in that case, that $t \longrightarrow \infty$. This is called a steady state. So what we really want to gain, is a solution of the original problem (2.1), neglecting the initial condition. So formally, a steady state solution is defined in the following way.

Definition 4.1 (Steady State Solution [1]). *The function $u(x,t)$ is called a periodic steady state solution of equation* (2.1)*, if*

1. *$u(x,t)$ satisfies* (2.1)*, but not necessarily the initial condition,*
2. *$u(x,t)$ is periodic, i.e. there exists a time period T, such that for all $t : u(x,t) = u(x,t+T)$.*

If the steady state solution $u(x,t)$ additionally fulfills the initial condition

$$u(x,0) = \sum_{k=0}^{N} u_k^c(x) = u_0(x),$$

it is also a solution of the initial value problem (2.1).

4.2 The Linear Problem

Before we turn to the general nonlinear problem (2.1), we discuss the features of the corresponding linear problem. So we consider the case where $\nu = \nu(x)$ doesn't depend on the gradient of u. Find $u(x,t) \in C^{2,1}(\Omega_T) \cap C([a,b] \times [0,T))$, such that

$$\begin{aligned} \alpha \frac{\partial u}{\partial t}(x,t) - \frac{\partial}{\partial x}\left(\nu(x)\frac{\partial u}{\partial x}\right)(x,t) = f(x,t) \qquad & (x,t) \in (a,b) \times (0,T) \\ u(a,t) = u(b,t) = 0 \qquad & t \in (0,T) \end{aligned} \tag{4.2}$$

where $\alpha = const > 0$ and $0 < \underline{\nu} \leq \nu(s) \leq \overline{\nu}$. Due to the linearity of ν, (4.2) is refered as the *Linear Problem.*

The Complex Problem

We assume that the right-hand side f is time-harmonic with some frequency ω, i.e.

$$f(x,t) = \hat{f}(x)e^{i\omega t}.$$

This allows us to switch from the time domain to the frequency domain and therefore using the ansatz

$$u(x,t) = \hat{u}(x)e^{i\omega t}$$

with the amplitude $\hat{u}(x)$ and the frequency ω. For the linear problem we obtain the complex time-domain problem

$$-\frac{\partial}{\partial x}\left(\nu(x)\frac{\partial \hat{u}}{\partial x}\right)(x) + i\omega\alpha\hat{u}(x) = \hat{f}(x). \tag{4.3}$$

In the rest of the thesis, we don't work with this complex notations. However we give the real reformulation of the complex and symmetric problem (4.3) and end up with a non-symmetric real problem.
In fact the complex problem is coupled with the real one by

$$\hat{f}(x) = f^c(x) - if^s(x) \text{ and } \hat{u}(x) = u^c(x) - iu^s(x).$$

Real Formulation

We assume that the right-hand side f is harmonic with freqency ω, which gives a time period $T = \frac{2\pi}{\omega}$, and leads consequently to the form

$$f(x,t) = f^c(x)\cos(\omega t) + f^s(x)\sin(\omega t).$$

Therefore, we can use the ansatz for the solution as

$$u(x,t) = u^c(x)\cos(\omega t) + u^s(x)\sin(\omega t),$$

with the same frequency ω. Inserting this ansatz into (4.2), we obtain a time-independent equation for the Fourier coefficients $u^c(x)$ and $u^s(x)$:

$$\left(-\frac{\partial}{\partial x}\left(\nu(x)\frac{\partial u^c}{\partial x}\right) + \alpha\omega u^s\right)\cos(\omega t) + \left(-\alpha\omega u^c - \frac{\partial}{\partial x}\left(\nu(x)\frac{\partial u^s}{\partial x}\right)\right)\sin(\omega t) = \\ = f^c\cos(\omega t) + f^s\sin(\omega t).$$

According to Chapter 2 we derive the variational formulation in time. Therefore we set up the ansatz- and testspace, that are identical indeed. Using the fact, that the trigonometric functions $\sin(\omega t)$ and $\cos(\omega t)$ are contained in $L_2(0,T)$, we state

$$L_2^{\text{harm}}(0,T) := \left\{u(x,\cdot) : u(x,\cdot) = u^c(x)\cos(\omega\cdot) + u^s(x)\sin(\omega\cdot)\right\}.$$

Since $L_2^{\text{harm}}(0,T) = \text{span}\left\{\cos(\omega t), \sin(\omega t)\right\}$, it is enough to test (4.2) with $\cos(\omega t)$ and $\sin(\omega t)$. Integrating over the time period T and taking advantage of the following orthogonality arguments

$$\frac{2}{T}\int_0^T \cos(\omega t)\cos(\omega t)dt = 1$$
$$\frac{2}{T}\int_0^T \cos(\omega t)\sin(\omega t)dt = 0$$
$$\frac{2}{T}\int_0^T \sin(\omega t)\sin(\omega t)dt = 1$$

we arrive at the system in classical formulation: Find $u = (u^c, u^s) \in \left(C^2(a,b)\right)^2 \cap \left(C[a,b]\right)^2$, such that

$$\begin{aligned} -\begin{pmatrix} \frac{d}{dx}\left(\nu(x)\frac{du^c}{dx}\right) \\ \frac{d}{dx}\left(\nu(x)\frac{du^s}{dx}\right)\end{pmatrix} + \alpha\omega\begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}\begin{pmatrix} u^c \\ u^s\end{pmatrix} &= \begin{pmatrix} f^c \\ f^s\end{pmatrix}, \\ u^c(a) = u^c(b) &= 0, \\ u^s(a) = u^s(b) &= 0. \end{aligned} \tag{4.4}$$

So the original linear problem (4.2) has been reduced to a system of equations for the Fourier coefficients $u^c(x)$ and $u^s(x)$ that only depend on the space coordinate. Equation (4.4) is the starting point of the discretisation in space afterwards.

Multiplying (4.4) by a testfunction $(v^c, v^s) \in \left(H_0^1(a,b)\right)^2$ and integrating over the domain (a,b) gives the following variational formulation: Find $u \in \left(H_0^1(a,b)\right)^2$, such that

$$\int_a^b \nu\left(\frac{dv^c}{dx}, \frac{dv^c}{dx}\right)\begin{pmatrix} \frac{du^c}{dx} \\ \frac{du^s}{dx}\end{pmatrix} + \alpha\omega\left(v^c, v^s\right)\begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}\begin{pmatrix} u^c \\ u^s\end{pmatrix} dx = \int_a^b \left(v^c, v^s\right)\begin{pmatrix} f^c \\ f^s\end{pmatrix} dx, \tag{4.5}$$

for all $v \in \left(H_0^1(a,b)\right)^2$.

4.3 The Nonlinear Problem

Now we want to consider the nonlinear case, where ν depends on the gradient of u, i.e. $\nu = \nu\left(\left|\frac{\partial u}{\partial x}\right|\right)$. In the linear case, the solution u can be expressed in terms of the same frequency as the excitation f, but this isn't true in the nonlinear case anymore. Due to the nonlinearity, the solution also depends on higher harmonics. However, on the account of the periodicity of the solution, it can consequently be expanded in a Fourier series and, therefore, approximated by a multiharmonic approach.

Real Formulation

We assume that the right-hand side $f(x,t)$ is given by a harmonic excitation with frequency ω, and therefore $f(x,t)$ has the form

$$f(x,t) = f^c(x)\cos(\omega t) + f^s(x)\sin(\omega t) \tag{4.6}$$

Consequently the solution u is periodic and we can use a Fourier series as an ansatz

$$u(x,t) = \sum_{k=0}^{\infty} u_k^c(x)\cos(k\omega t) + u_k^s(x)\sin(k\omega t), \tag{4.7}$$

where the Fourier coefficients are given by

$$u_k^c(x) = \frac{2}{T}\int_0^T u(x,t)\cos(k\omega t)dt \quad \text{and} \quad u_k^s(x) = \frac{2}{T}\int_0^T u(x,t)\sin(k\omega t)dt.$$

Here the time period is $T = \frac{2\pi}{\omega}$. Next we define the potential

$$\psi[u](x,t) := \nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}(x,t).$$

Hence, the potential $\psi[u](x,t)$ is periodic as well, and therefore, it can be expanded into a Fourier series:

$$\psi[u](x,t) = \sum_{k=0}^{\infty} \psi_k^c[u](x)\cos(k\omega t) + \psi_k^s[u](x)\sin(k\omega t), \tag{4.8}$$

where the Fourier coefficients are given by

$$\psi_k^c[u](x) = \frac{2}{T}\int_0^T \psi[u](x,t)\cos(k\omega t)dt \quad \text{and} \quad \psi_k^s[u](x) = \frac{2}{T}\int_0^T \psi[u](x,t)\sin(k\omega t)dt.$$

Remark 4.2. *Due to the harmonic excitation* (4.6)*, we have to take all harmonics in* (4.7) *and* (4.8) *into account. In [1] Bachinger handles the case of a pure cosine excitation of the form $f(x,t) = f^c(x)\cos(\omega t)$. In fact he shows, that in that case a reduction to odd harmonics both in u and ψ can be performed. Therefore the solution u can be entirely represented by odd harmonics, i.e.*

$$u(x,t) = \sum_{k=0}^{\infty} u_{2k+1}^c(x)\cos((2k+1)\omega t) + u_{2k+1}^s(x)\sin((2k+1)\omega t), \tag{4.9}$$

In fact we are only interested in harmonic excitations. The next remark gives some additionally results in case of periodic excitations.

Remark 4.3. *Beyond that, the analysis gives us also information about periodic excitation. In that case, the right-hand side can also be expressed in terms of a Fourier series, i.e.*

$$f(x,t) = \sum_{k=0}^{\infty} f_k^c(x)\cos(k\omega t) + f_k^s(x)\sin(k\omega t),$$

where the Fourier coefficients are given by

$$f_k^c(x) = \frac{2}{T}\int_0^T f(x,t)\cos(k\omega t)dt \quad \text{and} \quad f_k^s(x) = \frac{2}{T}\int_0^T f(x,t)\sin(k\omega t)dt.$$

Since we have more exciting oscillations involved, we expect more required addends in the Fourier series of the solution u.

The aim for numerical calculations is, not taking all harmonics into account. So we truncate the Fourier series, that give us the finite sums

$$u(x,t) \approx \sum_{k=0}^{N} u_k^c(x)\cos(k\omega t) + u_k^s(x)\sin(k\omega t) \tag{4.10}$$

respectively

$$\psi[u](x,t) \approx \sum_{k=0}^{N} \psi_k^c[u](x)\cos(k\omega t) + \psi_k^s[u](x)\sin(k\omega t). \tag{4.11}$$

Notation. For simplicity, from now on the truncated Fourier series of u is denoted by u as well. Similarly the truncated Fourier series of ψ is denoted by the same letter ψ.

In order to stay general, the truncated multiharmonic ansatz also considers non time-dependent parts of the excitation and consequently the solution by allowing the running index k to be 0. In fact we have

$$u(x,t) = u_0^c(x) + \sum_{k=1}^{N} u_k^c(x)\cos(k\omega t) + u_k^s(x)\sin(k\omega t).$$

Inserting this ansatz in our model problem (2.1), we obtain a time-independent equation for the Fourier coefficients $u_0^c(x), u_1^c(x), u_1^s(x), \cdots, u_N^c(x), u_N^s(x)$

$$\alpha\omega \sum_{k=0}^{N} k\left(u_k^c(x)\cos(k\omega t) - u_k^s(x)\sin(k\omega t)\right) -$$
$$- \frac{\partial}{\partial x}\left(\sum_{k=0}^{N} \psi_k^c\,[u]\,(x)\cos(k\omega t) + \psi_k^s\,[u]\,(x)\sin(k\omega t)\right) = \sum_{k=0}^{N} f_k^c(x)\cos(k\omega t) + f_k^s(x)\sin(k\omega t).$$

According to Chapter 2 we derive the variational formulation in time. Therefore we set up the ansatz- and testspace, that are identical indeed. Using the fact, that the trigonometric functions $\sin(k\omega t)$ and $\cos(k\omega t)$ are contained in $L^2(0,T)$, we state

$$L_2^{\text{mharm}}(0,T) := \left\{ u(x,\cdot) : u(x,\cdot) = \sum_{k=0}^{N} u_k^c(x)\cos(k\omega\cdot) + u_k^s(x)\sin(k\omega\cdot) \right\}.$$

Since $L_2^{\text{mharm}}(0,T) = \text{span}\left\{\cos(k\omega t)\right\}_{k=0}^{N} \cup \left\{\sin(k\omega t)\right\}_{k=1}^{N}$, it is enough to test (4.2) with these basis functions. Be aware, that, since $\sin(0) = 0$ this basis function can be neglected. Next we integrate over the time period T and take advantage of the following orthogonality arguments

$$\begin{aligned}
\frac{2}{T}\int_0^T \cos(k\omega t)\cos(l\omega t)dt &= \delta_{kl} && k \neq 0 \vee l \neq 0\\
\frac{2}{T}\int_0^T \cos(k\omega t)\sin(l\omega t)dt &= 0 && l \neq 0\\
\frac{2}{T}\int_0^T \sin(k\omega t)\sin(l\omega t)dt &= \delta_{kl} && l \neq 0\\
\frac{1}{T}\int_0^T \cos(0)\cos(0)dt &= 1 && k = 0 \wedge l = 0.
\end{aligned} \tag{4.12}$$

So we arrive at a system of nonlinear equations: Find $u(x) = (u_0^c(x), u_1^c(x), u_1^s(x), \cdots, u_N^c(x), u_N^s(x))^T \in \left(C^2(a,b)\right)^{2N+1}$, such that

$$\begin{gathered}
\alpha\omega \underbrace{\begin{pmatrix} 0 & & & & & & \\ & 0 & 1 & & & & \\ & -1 & 0 & & & & \\ & & & \ddots & & & \\ & & & & & 0 & N \\ & & & & & -N & 0 \end{pmatrix}}_{=:D_N} \begin{pmatrix} u_0^c(x) \\ u_1^c(x) \\ u_1^s(x) \\ \vdots \\ u_N^c(x) \\ u_N^s(x) \end{pmatrix} - \nabla \cdot \begin{pmatrix} \psi_0^c\left[u\right](x) \\ \psi_1^c\left[u\right](x) \\ \psi_1^s\left[u\right](x) \\ \vdots \\ \psi_N^c\left[u\right](x) \\ \psi_N^s\left[u\right](x) \end{pmatrix} = \begin{pmatrix} f_0^c(x) \\ f_1^c(x) \\ f_1^s(x) \\ \vdots \\ f_N^c(x) \\ f_N^s(x) \end{pmatrix} \\
\begin{pmatrix} u_0^c(a) \\ u_1^c(a) \\ u_1^s(a) \\ \vdots \\ u_N^c(a) \\ u_N^s(a) \end{pmatrix} = \begin{pmatrix} u_0^c(b) \\ u_1^c(b) \\ u_1^s(b) \\ \vdots \\ u_N^c(b) \\ u_N^s(b) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}
\end{gathered} \tag{4.13}$$

Rewriting in a compact form gives the following time-independent, vector valued equation

$$\begin{aligned}
\alpha\omega D_N u(x) - \nabla \cdot \psi[u](x) = f(x) \quad x \in (a,b),\\
u(a) = u(b) = 0.
\end{aligned}$$

In case of the harmonic excitation (4.6), the Fourier coefficients $f_i^c(x)$ and $f_i^s(x)$ of the right-hand side f are zero for $i \neq 1$.

Note that in order to derive $\psi[u](x)$ we need some integration in time. The integrals

$$\begin{aligned}
\psi_k^c\left[u\right](x) &= \frac{2}{T}\int_0^T \nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}(x,t)\cos(k\omega t)dt\\
\psi_k^s\left[u\right](x) &= \frac{2}{T}\int_0^T \nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}(x,t)\sin(k\omega t)dt
\end{aligned}$$

have to be evaluated by some suitable quadrature rule. Since this expression still depends on u given by (4.10) we obtain a fully coupled system of nonlinear equations.

Multiplying by a testfunction $v = (v_0^c(x), v_1^c(x), v_1^s(x), \cdots, v_N^c(x), v_N^s(x))^T \in \left(H_0^1(a,b)\right)^{2N+1}$ and integrating over the domain gives the following variational formulation. Find $u = (u_0^c(x), u_1^c(x), u_1^s(x), \cdots, u_N^c(x), u_N^s(x))^T \in \left(H_0^1(a,b)\right)^{2N+1}$, such that

$$\int_a^b \alpha\omega D_N u(x) \cdot v(x) + \psi[u](x) \cdot \nabla v(x) dx = \int_a^b f(x) \cdot v(x) dx, \quad \forall v \in \left(H_0^1(a,b)\right)^{2N+1}. \tag{4.14}$$

Link of the multiharmonic approach to the Space-Time Variational Formulation

Finally, we want to demonstrate the link between the space-time variational formulation of Chapter 2 and the multiharmonic approach of this chapter. In fact, inserting the truncated Fourier series of u (4.10) and ψ (4.11) into the space-time variational formulation (2.2) and testing with testfunctions of the same kind yields the same system of nonlinear equations as achieved in (4.13). By defining the spaces

$$L_2^{mharm}(0,T) = \left\{ u(x,t) : u(x,t) = \sum_{k=0}^{N} u_k^c(x)\cos(k\omega t) + u_k^s(x)\sin(k\omega t) : u_k^c, u_k^s \in H_0^1(a,b) \right\}$$
$$\subset H_0^{1,1}(\Omega_T)$$

and

$$V_0 = H_0^{1,1}(\Omega_T),$$

the variational formulation can be stated as follows: Find $u \in L_2^{mharm}(0,T)$ such that

$$\frac{2}{T}\int_0^T \int_a^b \alpha \frac{\partial u}{\partial t}(x,t) v(x,t) + \nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}(x,t)\frac{\partial v}{\partial x}(x,t) dx dt$$
$$= \frac{2}{T}\int_0^T \int_a^b f(x,t) v(x,t) dx dt, \quad \forall v(x,t) \in V_0.$$

Inserting this ansatz and evaluating the integrals by taking advantage of the cited orthogonality arguments yields (4.14).

Incorporating the Initial Condition

Since we are interested in a steady state solution, we don't take the initial condition into account. If we are not interested in an excited state, but in the first period of the harmonic, the Fourier coefficient corresponding to $k = 0$ can be used to fulfill the initial condition $u(x,0) = u_0(x)$. The coefficient $u_0^c(x)$ can be computed via the initial setting at time $t = 0$:

$$u_0^c(x) = u(x,0) - \sum_{k=1}^{N} u_k^c(x) = u_0(x) - \sum_{k=1}^{N} u_k^c(x).$$

Therefore $u_0^c(x)$ can be eliminated from the system (4.13). In this case, the condition $u(x,0) = u_0(x)$ is an essential one. Due to the periodicity of $u(x,t)$, namely

$$u(x,T) = \sum_{k=0}^{N} u_k^c(x) = u(x,0) = u_0(x),$$

the value at $t = T$ has to be fixed also. Therefore, the trial- and test-spaces are stated as

$$\tilde{L}_2^{mharm}(0,T) = \Big\{u(x,t) \in L_2^{mharm}(0,T) : u(x,0) = u(x,T) = u_0(x)\Big\} \subset H_0^{1,1}(\Omega_T)$$

and

$$V_{00} = \Big\{u(x,t) \in L_2^{mharm}(0,T) : u(x,0) = u(x,T) = 0\Big\} = H_{0,0}^{1,1}(\Omega_T).$$

Now, the variational formulation reads as follows: Find $u \in \tilde{L}_2^{mharm}(0,T)$, such that

$$\begin{aligned}\frac{2}{T}\int_0^T \int_a^b \alpha \frac{\partial u}{\partial t}(x,t)v(x,t) + \nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}(x,t)\frac{\partial v}{\partial x}(x,t)dxdt \\ = \frac{2}{T}\int_0^T \int_a^b f(x,t)v(x,t)dxdt, \quad \forall v(x,t) \in V_{00}.\end{aligned}$$

It is important to mention, that in this case, the consideration of the initial condition enforces a modification of the right-hand side f.

Chapter 5

The Finite Element Method

The aim of this chapter is to perform the discretisation in space by using the finite element method. So, after giving a short introduction to the finite element method, we apply FEM to our linear problem. The resulting system of linear equations is solved by the preconditioned GMRES iteration proposed by Cai and Xu. In order to get a feeling of the behaviour of the convergence rates of the associated GMRES iteration, a spectral analysis is performed for the case that $\nu(x) = \nu =$ const. We provide a complete Fourier analysis for two special cases. Firstly we investigate the case, where the coarse grid isn't taking into account and we only use the symmetric part as a preconditioner. This case is formally related to $\beta = \infty$. In the second case, we get the coarse grid by doubling the meshsize of the fine grid, i.e. $H = 2h$ Concerning the nonlinear problem, the main issue is to calculate the Frechet derivative of the nonlinear operator, that we need in further procedure for applying Newton's iteration.

5.1 Introduction to the Finite Element Method

Various introduction into the finite element method (FEM) at least for 1D problems can be found in [3, 10, 15, 17, 23]. We give a very short outline for the most important features in order to clarify subsequent concepts and notations.

Variational Formulation

The starting point of the finite element discretisation in space is the variational formulation. The latter can always be written in the following form. Find $u \in V$, such that

$$a(u,v) = \langle F, v\rangle \quad ,\forall v \in V. \tag{5.1}$$

where $a(\cdot,\cdot) : V \times V \longrightarrow \mathbb{R}$ is a bilinear form and $\langle F, v\rangle\, V \longrightarrow \mathbb{R}$ is a linear form on the Hilbertspace V. The next theorem gives complete information of the existence and uniqueness of a weak solution of (5.1).

Theorem 5.1 (Lax-Milgram Theorem [7]). *Let V be a Hilbert space with scalar product $(\cdot,\cdot)$ and the corresponding norm $\|\cdot\|$, $F : V \longrightarrow \mathbb{R}$ a (V-)bounded linear functional ($F \in V^*$) and $a(\cdot,\cdot) : V \times V \longrightarrow \mathbb{R}$ a bilinear form with the following properties:*

1. *$a(\cdot,\cdot)$ is V-elliptic, i.e. there exists a constant $\mu_1 > 0$, such that*

$$\mu_1 \|v\|^2 \leq a(v,v), \quad \forall v \in V,$$

2. $a(\cdot,\cdot)$ is V-bounded, i.e. there exists a constant $\mu_2 > 0$, such that

$$|a(u,v)| \leq \mu_1 \|u\| \|v\|, \quad \forall u,v \in V.$$

Then there exists a unique solution $u^* \in V$ of the variational problem (5.1)*.*

Galerkin Method

Let us assume, that we have finite-dimensional subspaces $V_h \subset V$, and let us associate to the variational problem (5.1) the following discrete problems Find $u_h \in V_h \subset V$, such that

$$a(u_h, v_h) = \langle F, v_h \rangle \quad , \forall v_h \in V_h. \tag{5.2}$$

Since V_h is a subspace of V, Theorem 5.1 guarantees existence and uniqueness of a solution $u_h^* \in V_h$. In fact we want to achieve, that for $h \longrightarrow 0$, the discrete solution u_h^* converges to the exact solution u^*. The next lemma provides information about the estimation of the discretisation errror by the approximation error.

Lemma 5.2 (Cea's Lemma [17]). *Let V be a Hilbert space, $F \in V^*$ and $a(\cdot,\cdot) : V \times V \longrightarrow \mathbb{R}$ an elliptic and bounded bilinear form (i.e. $a(\cdot,\cdot)$ fulfills the assumptions of Theorem 5.1). Let V_h be a finite-dimensional subspace of V. Assuming, that $u^* \in V$ is the unique solution of* (5.1) *and $u_h^* \in V_h$ is the unique solution of* (5.2)*, then we have*

$$\|u^* - u_h^*\| \leq \frac{\mu_2}{\mu_1} \inf_{w_h \in V_h} \|u^* - w_h\|.$$

Having guaranteed the existence and uniqueness of a discrete solution in V_h, the next step is to choose appropriate basis functions of V_h. Once the basis $\{p_i(x)\}_{i=1}^{N_h}$ is chosen, i.e. $V_h = \operatorname{span}\{p_i(x)\}_{i=1}^{N_h}$, we can represent u_h and can switch from (5.2) to the corresponding matrix representation. Using the representation

$$u_h(x) = \sum_{i=1}^{N_h} u_h^{(i)} p_i(x)$$

we can switch to the Galerkin system. Find $\underline{u}_h = \left(u_h^{(i)}\right)_{i=1}^{N_h} \in \mathbb{R}^{N_h}$, such that

$$K_h \underline{u}_h = \underline{f}_h, \tag{5.3}$$

where

$$(K_h)_{i,j} = a(p^{(i)}, p^{(j)}) \quad \text{and} \quad \left(\underline{f}_h\right)_i = \left\langle F, p^{(i)} \right\rangle.$$

In fact the linear system of equations (5.3) is equivalent to the original discrete problem (5.2). Consequently this defines the so called Ritz-Isomorphism

$$u_h \in V_h \longleftrightarrow \underline{u}_h \in \mathbb{R}^{N_h}.$$

5.1.1 Finite Element Spaces

Generally speaking, the finite element method is characterised by a special choice of V_h and its basis functions. In this section we want to define these finite element spaces for our one-dimensional problem. Since we want to solve our problem on two grids in the end, we define a fine and a coarse finite element space.

Setting up the mesh

Coarse Mesh: We discretise the interval $[a, b]$ into N_H subintervals by introducing nodes y_i, $i = 0, \ldots, N_H + 1$, that fulfill

$$a = y_0 < y_1 < \ldots < y_{N_H - 1} < y_{N_H} < y_{N_h + 1} = b.$$

This gives us a triangulation T_H as a set of subintervals $T_k = [y_{k-1}, y_k], k = 1, \ldots, N_H + 1$. For simplicity, we use equidistant spaced nodes, i.e. $y_k - y_{k-1} = H$, that gives us the meshsize of H.

Fine Mesh: We discretise the interval $[a, b]$ into $N_h + 1$ subintervals by introducing nodes x_i, $i = 0, \ldots, N_h + 1$, that fulfill

$$a = x_0 < x_1 < \ldots < x_{N_h - 1} < x_{N_h} < x_{N_h + 1} = b.$$

This gives us a triangulation T_h as a set of subintervals $T_k = [x_{k-1}, x_k], k = 1, \ldots, N_h + 1$. For simplicity, we use equidistant spaced nodes, i.e. $x_k - x_{k-1} = h$, that gives us the meshsize of h.

The simplest method of constructing the fine mesh is to regard the midpoints of the individual subintervals T_k of the coarse mesh as addional nodes, see Figure 5.1.

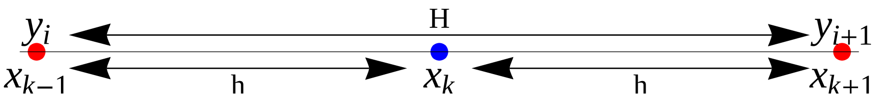

Figure 5.1: Construction of the fine mesh

Setting up the finite element spaces

As the set of trial- and test-functions we use the space of piecewise affine-linear continuous functions on $[a, b]$. By denoting P_1 the set of all polynomials of degree less equal one, we obtain for the two different grides respectively

$$V_h = \{u_h \in C[a, b] : v|_T \in P_1, \forall T \in T_h\} \quad \text{and}$$
$$V_H = \{u_H \in C[a, b] : v|_T \in P_1, \forall T \in T_H\}.$$

The next step is to choose appropriate basis functions. In our case, the choice is the usual nodal basis as defined below.

Choosing the basis function

Coarse Mesh: We choose a uniform triangulation of (a, b) with meshsize H and discretise the space $H_0^1(a, b)$ by piecewise linear functions. Therefore we choose the usual nodal basis

$$\Phi_i(x) = \begin{cases} \frac{x - y_{i-1}}{H} & \text{for } y_{i-1} < x \leq y_i \\ \frac{y_{i+1} - x}{H} & \text{for } y_i < x \leq y_{i+1} \\ 0 & \text{else} \end{cases}$$

and approximating $\left(H_0^1(a, b)\right)^2$ by $S_H^1 = \left(\text{span}\,\{\Phi_i\}_{i=1}^{N_H}\right)^2$.

Fine Mesh: We choose a uniform triangulation of (a,b) with meshsize h and discretise the space $H_0^1(a,b)$ by piecewise linear functions. Therefore we choose the usual nodal basis

$$\varphi_i(x) = \begin{cases} \frac{x-x_{i-1}}{h} & \text{for } x_{i-1} < x \leq x_i \\ \frac{x_{i+1}-x}{h} & \text{for } x_i < x \leq x_{i+1} \\ 0 & \text{else} \end{cases}$$

and approximating $\left(H_0^1(a,b)\right)^2$ by $S_h^1 = \left(\text{span}\,\{\varphi_i\}_{i=1}^{N_h}\right)^2$.

Obviously the coarse space S_H^1 is contained in the fine space S_h^1, i.e. S_H^1 is a subspace of S_h^1.

5.2 The Linear Problem

On the following pages we use the finite element method described above to solve the linear time-harmonic problem. The Cai-Xu preconditioner is applied to the resulting system of linear equations and the convergence behaviour of the corresponding GMRES method is explored by developing some spectral analysis.

5.2.1 Variational Formulation

As our starting point we use the variational formulation (4.5) derived in Chapter 4.

$$\int_a^b \nu \left(\frac{dv^c}{dx}, \frac{dv^c}{dx}\right) \begin{pmatrix} \frac{du^c}{dx} \\ \frac{du^s}{dx} \end{pmatrix} + \alpha\omega\,(v^c, v^s) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u^c \\ u^s \end{pmatrix} dx = \int_a^b (v^c, v^s) \begin{pmatrix} f^c \\ f^s \end{pmatrix} dx.$$

Let us now define the bilinear form

$$a(u,v) := \int_a^b \nu \frac{du^c}{dx}\frac{dv^c}{dx}dx + \int_a^b \nu \frac{du^s}{dx}\frac{dv^s}{dx}dx + \alpha\omega \int_a^b u^c v^s dx - \alpha\omega \int_a^b u^s v^c dx$$

and the linear form

$$\langle F, v\rangle := \int_a^b f^c v^c + f^s v^s dx.$$

For simplicity we also introduce the following vector valued abbrevations $v = (v^c, v^s)$ and $u = (u^c, u^s)$. Now we can state (4.4) in variational form. Find $u \in \left(H_0^1(a,b)\right)^2$, such that

$$a(u,v) = \langle F, v\rangle \qquad \forall v \in \left(H_0^1(a,b)\right)^2.$$

For the analysis, we also want to define the following bilinear forms

$$(L\varphi,\psi) := \int_a^b \nu \frac{d\varphi}{dx}\frac{d\psi}{dx}dx \quad \text{and} \quad (D\varphi,\psi) := \alpha\omega \int_a^b \varphi\psi dx.$$

Furthermore, we introduce

$$(Au,v) := (Lu^c, v^c) + (Lu^s, v^s) \quad \text{and} \quad (Mu,v) := (Du^c, v^s) - (Du^s, v^c).$$

Now our bilinear form $a(\cdot,\cdot)$ reads as

$$a\,(u,v) = (Au,v) + (Mu,v) \quad \forall u,v \in \left(H_0^1(a,b)\right)^2.$$

Existence and Uniqueness

In order to show existence and uniqueness, we have to verify the assumptions of Lax-Milgram (Theorem 5.1).

- $a(\cdot,\cdot)$ is positive definite. Indeed, since we have the identity

$$(Mu, u) = -(u, Mu) = -(Mu, u) \Leftrightarrow (Mu, u) = 0$$

the following estimates are valid:

$$\begin{aligned}
a(u,u) &= \int_a^b \nu(x) \left(\frac{du^c}{dx}\right)^2 dx + \int_a^b \nu(x) \left(\frac{du^s}{dx}\right)^2 dx \\
&\geq \underline{\nu} \int_a^b \left(\frac{du^c}{dx}\right)^2 dx + \underline{\nu} \int_a^b \left(\frac{du^s}{dx}\right)^2 dx \\
&\geq \frac{\underline{\nu}}{C_F^2} \left(\|u^c\|^2_{L^2(a,b)} + \|u^s\|^2_{L^2(a,b)} \right) \geq \frac{\underline{\nu}}{C_F^2 + 1} \left(\|u^c\|^2_{H^1(a,b)} + \|u^s\|^2_{H^1(a,b)} \right) \\
&= \frac{\underline{\nu}}{C_F^2 + 1} \|u\|^2_{(H^1(a,b))^2} = \mu_1 \|u\|^2_{(H^1(a,b))^2},
\end{aligned}$$

with $\mu_1 = \frac{\underline{\nu}}{C_F^2+1}$. Here we have used the Friedrich's inequality

$$\|u\|_{L^2(a,b)} \leq C_F \, |u|_{H^1(a,b)}, \quad \forall u \in H_0^1(a,b) \tag{5.4}$$

with the Friedrichs constant $C_F^2 = \frac{(b-a)^2}{2}$.

- $a(\cdot,\cdot)$ is bounded: First of all we will show, that we have boundedness of each component $(A\cdot,\cdot)$ and $(M\cdot,\cdot)$. Using triangle inequality and Chauchy's inequality, we get

$$\begin{aligned}
|(Au, v)| &= \left| \int_a^b \nu(x) \frac{du^c}{dx}\frac{dv^c}{dx} dx + \int_a^b \nu(x) \frac{du^s}{dx}\frac{dv^s}{dx} dx \right| \\
&\leq \overline{\nu} \int_a^b \left|\frac{du^c}{dx}\right| \left|\frac{dv^c}{dx}\right| dx + \overline{\nu} \int_a^b \left|\frac{du^s}{dx}\right| \left|\frac{dv^s}{dx}\right| dx \\
&\leq \overline{\nu} \, |u^c|_{H^1(a,b)} |v^c|_{H^1(a,b)} + \overline{\nu} \, |u^s|_{H^1(a,b)} |v^s|_{H^1(a,b)} \\
&\leq \overline{\nu} \left(|u^c|^2_{H^1(a,b)} + |u^s|^2_{H^1(a,b)} \right)^{1/2} \left(|v^c|^2_{H^1(a,b)} + |v^s|^2_{H^1(a,b)} \right)^{1/2} \\
&= \overline{\nu} \, |u|_{(H^1(a,b))^2} |v|_{(H^1(a,b))^2}
\end{aligned}$$

and

$$\begin{aligned}
|(Mu, v)| &= \left| \alpha\omega \int_a^b u^c v^s dx - \alpha\omega \int_a^b u^s v^c dx \right| \\
&\leq \alpha\omega \left| \int_a^b u^c v^s dx \right| + \alpha\omega \left| \int_a^b u^s v^c dx \right| \\
&\leq \alpha\omega \, \|u^c\|_{L^2(a,b)} \|v^s\|_{L^2(a,b)} + \alpha\omega \, \|u^s\|_{L^2(a,b)} \|v^c\|_{L^2(a,b)} \\
&\leq \alpha\omega \left(\|u^c\|^2_{L^2(a,b)} + \|u^s\|^2_{L^2(a,b)} \right)^{1/2} \left(\|v^c\|^2_{L^2(a,b)} + \|v^s\|^2_{L^2(a,b)} \right)^{1/2} \\
&= \alpha\omega \, \|u\|_{(L^2(a,b))^2} \|v\|_{(L^2(a,b))^2}.
\end{aligned}$$

Now combining these two estimates and applying Cauchy's inequality again gives

$$\begin{aligned} a(u,v) &= (Au,v) + (Mu,v) \\ &\le \max\{\overline{\nu}, \alpha\omega\} \left(|u|_{(H^1(a,b))^2} \, |v|_{(H^1(a,b))^2} + \|u\|_{(L^2(a,b))^2} \, \|v\|_{(L^2(a,b))^2} \right) \\ &\le \max\{\overline{\nu}, \alpha\omega\} \left(|u|_{(H^1(a,b))^2} + \|u\|_{(L^2(a,b))^2} \right)^{1/2} \left(|v|_{(H^1(a,b))^2} + \|v\|_{(L^2(a,b))^2} \right)^{1/2} \\ &= \mu_2 \, \|u\|_{(H^1(a,b))^2} \, \|v\|_{(H^1(a,b))^2} , \end{aligned}$$

with $\mu_2 = \max\{\overline{\nu}, \alpha\omega\}$.

So due to Lax Milgram (Theorem 5.1) we have the existence and uniqueness in $\left(H_0^1(a,b)\right)^2$ with the generic constants

$$\mu_1 = \frac{\underline{\nu}}{C_F^2 + 1} \quad \text{and} \quad \mu_2 = \max\{\overline{\nu}, \alpha\omega\}.$$

Since S_h^1 is a subspace of $\left(H_0^1(a,b)\right)^2$, we have also the existence and uniqueness of the solution in the discrete case: Find $u_h \in S_h^1$, such that

$$a(u_h, v_h) = \langle F, v_h \rangle \qquad \forall v_h \in S_h^1.$$

Finite Element Systems

We use the usual nodal basis of S_h^1 as defined before to obtain the Galerkin system, that reads as follows: Find $\underline{u}_h = (\underline{u}_h^c, \underline{u}_h^s)^T \in \mathbb{R}^{2N_h}$, such that

$$A_h^M \underline{u}_h = \underline{f}_h. \tag{5.5}$$

The stiffness matrix A_h^M is given by

$$A_h^M = \begin{pmatrix} L_h & D_h \\ -D_h & L_h \end{pmatrix} = \underbrace{\begin{pmatrix} L_h & 0 \\ 0 & L_h \end{pmatrix}}_{=A_h} + \underbrace{\begin{pmatrix} 0 & D_h \\ -D_h & 0 \end{pmatrix}}_{=M_h},$$

with

$$(L_h)_{ij} = (L\varphi_i, \varphi_j) \quad \text{and} \quad (D_h)_{ij} = (D\varphi_i, \varphi_j).$$

The right-hand side $\underline{f}_h = (\underline{f}_h^c, \underline{f}_h^s)^T$ is obtained by evaluating

$$(f_h^c)_i = (f^c, \varphi_i)_{L^2(a,b)} \quad \text{and} \quad (f_h^s)_i = (f^s, \varphi_i)_{L^2(a,b)}.$$

In the case that $\nu(x) = \nu = const$, the stiffness matrix L_h and mass matrix D_h are given by

$$L_h = \frac{\nu(b-a)}{h} \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix} \quad \text{and} \quad D_h = \frac{\alpha\omega(b-a)h}{6} \begin{pmatrix} 4 & 1 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4 & 1 \\ 0 & \cdots & 0 & 1 & 4 \end{pmatrix}.$$

5.2.2 GMRES Preconditioner

In order to solve this non-symmetric problem (5.5), we want to use GMRES and the preconditioning technique proposed in Chapter 3. Now we want to use this theory and apply the Cai-Xu preconditioner

$$B^M = I_H^h \left(A_H^M\right)^{-1} Q_h^H + \beta B_h. \tag{5.6}$$

We want to treasure, that B_h is a good preconditioner for the symmetric part A_h, β is some balancing constant, A_H^M the coarse stiffness matrix, I_H^h the interpolation from the coarse space into the fine one and Q_h^H the obverse restriction.

Coarse Stiffness Matrix A_H^M

$$A_H^M = \begin{pmatrix} L_H & D_H \\ -D_H & L_H \end{pmatrix} = \underbrace{\begin{pmatrix} L_H & 0 \\ 0 & L_H \end{pmatrix}}_{=A_H} + \underbrace{\begin{pmatrix} 0 & D_H \\ -D_H & 0 \end{pmatrix}}_{=M_H}$$

with

$$(L_H)_{ij} = (L\Phi_i, \Phi_j) \quad \text{and} \quad (D_H)_{ij} = (D\Phi_i, \Phi_j).$$

Constructing the interpolation operator I_H^h: Prolongation

We want to construct $\tilde{I}_H^h : \mathbb{R}^{N_H} \to \mathbb{R}^{N_h}$. This gives $\tilde{I}_H^h \in \mathbb{R}^{N_h \times N_H}$. The corresponding matrix has the following structure (for example for $H = 2h$):

$$\tilde{I}_H^h = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

This gives the following interpolation formula: $\underline{u}_h^c = \tilde{I}_H^h \underline{u}_H^c$ respectively $\underline{u}_h^s = \tilde{I}_H^h \underline{u}_H^s$. So, finally, we arrive at the representation

$$I_H^h = \begin{pmatrix} \tilde{I}_H^h & 0 \\ 0 & \tilde{I}_H^h \end{pmatrix}.$$

Constructing the L^2 projector Q_h^H: Restriction

In [4] and therefore Chapter 3 the restriction of the fine subspace to the coarse subspace is done by an L^2 projection. i.e. $Q_H : S_H^1 \longrightarrow S_h^1$

$$(Q_h^H u_h, v_H) = (u_h, v_H) \quad \forall u_h \in S_h^1, v_H \in S_H^1$$

So in fact by using our nodal basis, we have to solve the equation: Find $\underline{u}_H \in \mathbb{R}^{2N_H}$ such that for a given $\underline{u}_h \in \mathbb{R}^{2N_h}$

$$M_H \underline{u}_H^* = M_h^H \underline{u}_h^*, \quad \text{where } * \in \{s, c\},$$

with the M_H mass matrix on the coarse space, given by $(M_H)_{ij} = (\Phi_i, \Phi_j)$ and $\left(M_h^H\right)_{ij} = (\phi_i, \Phi_j)$. So $\tilde{Q}_h^H : \mathbb{R}^{N_h} \longrightarrow \mathbb{R}^{N_H}$ is given by

$$\tilde{Q}_h^H = M_H^{-1} M_h^H$$

respectively $Q_h^H : \mathbb{R}^{2N_h} \longrightarrow \mathbb{R}^{2N_H}$ is given by

$$Q_h^H = \begin{pmatrix} \tilde{Q}_h^H & 0 \\ 0 & \tilde{Q}_h^H \end{pmatrix}.$$

Weighted Restriction I_h^H: Restriction

Another approach is to use weighted restriction I_h^H. The main advantage of this operator is, that we have $I_h^H = \left(I_H^h\right)^T$.

Rescaling

The first step is to choose the scalar product and norm. Since the constants C_1 and C_2 are needed to be independent of the meshparameter h, we have to introduce some necessary rescaling.
In fact L_h and D_h and therefore A_h and M_h depend on the mesh parameter h. The following estimates of the eigenvalues can be obtained straight forward.

$$\frac{h\underline{\nu}(b-a)}{6C_F^2} \leq \lambda_{\min}(L_h) \leq \lambda_{\max}(L_h) \leq \frac{4(b-a)}{h}$$

$$\frac{h\alpha\omega(b-a)}{6} \leq \lambda_{\min}(D_h) \leq \lambda_{\max}(D_h) \leq \frac{h\alpha\omega(b-a)}{2}$$

In order to apply the theory of Chapter 3, a proper scaling of (5.5) has to be performed. The following scaling argument will yield the desired result.

$$\frac{1}{h} A_h^M \underline{u}_h = \frac{1}{h} \underline{f}_h.$$

We choose the discrete scalar product and the discrete norm as

$$(\underline{u}_h, \underline{v}_h)_{A_h} := \frac{1}{h}(A_h \underline{u}_h, \underline{v}_h)_{(l^2)^2} = \frac{1}{h}((L_h \underline{u}_h^c, \underline{v}_h^c)_{l^2} + (L_h \underline{u}_h^s, \underline{v}_h^s)_{l^2})$$
$$\|\underline{u}_h\|_{A_h} := \sqrt{(\underline{u}_h, \underline{u}_h)_{A_h}}.$$

We also define the following scalar product and norm on the space of finite element functions S_h^1.

$$(u_h, v_h)_A := (A u_h, v_h) = (L u_h^c, v_h^c) + (L u_h^s, v_h^s)$$
$$\|u_h\|_A := \sqrt{(A u_h, u_h)} = \sqrt{(L u_h^c, u_h^c) + (L u_h^s, u_h^s)}$$

Therefore we obtain the norm equivalences

$$\|u_h\|^2_{L^2(a,b)} \cong h \|\underline{u}_h\|^2_{l^2} \quad \text{resp.} \quad \|u_h\|^2_A \cong \|\underline{u}_h\|^2_{A_h} . \tag{5.7}$$

Due to Theorem 3.2, Assumption 1, Assumption 2 and Assumption 3 have to be fulfilled. Hence we have the norm equivalence (5.7), it is sufficient to prove the required assumptions in S^1_h.

Corollary 5.3.

$$\|u_h\|_{(L^2(a,b))^2} \le C_1 \|u_h\|_A , \quad \forall u_h \in S^1_h,$$

with $C_1 = \frac{C_F}{\sqrt{\underline{\nu}}}$.

Proof. Using Friedrich's inequality (5.4), we obtain the estimate

$$\begin{aligned}\|u_h\|^2_A &= \underline{\nu} \int_a^b \frac{du^c_h}{dx}^2 dx + \underline{\nu} \int_a^b \frac{du^s_h}{dx}^2 dx \ge \frac{\underline{\nu}}{C_F^2} \left(\|u^c_h\|^2_{L^2(a,b)} + \|u^s_h\|^2_{L^2(a,b)} \right) \\ &\ge \frac{\underline{\nu}}{C_F^2} \|u_h\|^2_{(L^2(a,b))^2} .\end{aligned}$$

□

Corollary 5.4.

$$(Mu_h, v_h) \le C_2 \|u_h\|_{(L^2(a,b))^2} \|v_h\|_A , \quad \forall u_h, v_h \in S^1_h,$$

with $C_2 = C_1 \alpha\omega$.

Proof. Using Cauchy's inequality and Corollary 5.3, we arrive at

$$|(Mu_h, v_h)| \le \alpha\omega \|u_h\|_{(L^2(a,b))^2} \|v_h\|_{(L^2(a,b))^2} \le C_1 \alpha\omega \|u_h\|_{(L^2(a,b))^2} \|v_h\|_A .$$

□

Since in our case we have $M = -M^T$, we also have:

$$|(u, Mv)| = |(Mv, u)| \le C_2 \|v\|_{(L^2(a,b))^2} \|u\|_A , \quad \forall u \in \left(H^1_0(a,b)\right)^2 .$$

Corollary 5.5. *Let us assume the H^2-coercivity of the adjoint problem: Find $w \in \left(H^1_0(a,b)\right)^2$, such that*

$$a(v, w) = (f, v)_{(L_2(a,b))^2}, \quad \forall v \in \left(H^1_0(a,b)\right)^2 , \tag{5.8}$$

for a given $f \in (L_2(a,b))^2$, i.e. the solution w of the variational problem (5.8) fulfills

1. $w \in \left(H^1_0(a,b)\right)^2 \cap \left(H^2(a,b)\right)^2$ *and*
2. $\exists \mu_3 > 0 : \|w\|_{(H^2(a,b))^2} \le \mu_3 \|f\|_{(L_2(a,b))^2}$.

Then the parameter δ_H, defined by

$$\delta_H = \sup_{v_h \in S_h^1} \frac{\|(I - P_H)v_h\|_{(L^2(a,b))^2}}{\|v_h\|_A}$$

fulfills the inequality

$$\delta_H \leq C_3 H$$

with $C_3 = \left(1 + \frac{\mu_2}{\mu_1}\right) \frac{\sqrt{1+C_F^2}}{C_1 \mu_3}$, where the Ritz projection

$$\begin{aligned} P_H : S_h^1 &\to S_H^1 \\ P_H u_h &= u_H \end{aligned}$$

is defined by the variational problem: Find $u_H \in S_H^1$ such that

$$a(u_H, v_H) = a(u_h, v_H), \quad \forall v_H \in S_H^1 \tag{5.9}$$

for a given right-hand side $u_h \in S_h^1$.

Proof. Due to Lax-Milgram, there exists a unique $u_H = P_H u_h \in S_H^1$ of the variational problem (5.9). Next we choose $f = u_h - u_H \in \left(H_0^1(a,b)\right)^2$ in the adjoint problem (5.8). The H^2-coercivity of the adjoint problem (5.8) guarantees the existence of the solution $w \in \left(H_0^1(a,b)\right)^2 \cap \left(H^2(a,b)\right)^2$ and gives us the estimate

$$\|w\|_{(H^2(a,b))^2} \leq \frac{1}{\mu_3} \|u_h - u_H\|_{(L^2(a,b))^2} \tag{5.10}$$

Taking $v = u_h - u_H \in S_h^1 \subset \left(H_0^1(a,b)\right)^2$ in (5.8) and using the Galerkin-orthogonality gives us the estimate

$$\begin{aligned} \|u_h - u_H\|^2_{(L_2(a,b))^2} &= a(u_h - u_H, w) \\ &= a(u_h - u_H, w) - (a(u_h, w_H) - a(u_H, w_H)) \\ &= a(u_h - u_H, w - w_H) \\ &\leq \mu_2 \|u_h - u_H\|_{(H^1(a,b))^2} \|w_h - w_H\|_{(H^1(a,b))^2}, \quad \forall w_H \in S_H^1. \end{aligned}$$

Using triangle inequality, we arrive at

$$\|u_h - u_H\|^2_{(L_2(a,b))^2} \leq \mu_2 \left(\|u_h\|_{(H^1(a,b))^2} + \|u_H\|_{(H^1(a,b))^2}\right) \inf_{w_H \in S_H^1} \|w - w_H\|_{(H^1(a,b))^2}. \tag{5.11}$$

Next we give estimates for the two terms of (5.11). Using the H^1-approximation theorem and the H^2-coercivity of the adjoint problem (5.8), we get

$$\inf_{w_H \in S_H^1} \|w - w_H\|_{(H^1(a,b))^2} \leq \sqrt{1 + C_F^2} H \|w\|_{(H^2(a,b))^2} \leq \frac{\sqrt{1 + C_F^2}}{\mu_3} H \|u_h - u_H\|_{(L_2(a,b))^2}. \tag{5.12}$$

Boundedness and coercivity of the bilinear form $a(\cdot, \cdot)$ gives the chain of inequalities

$$\mu_1 \|u_H\|^2_{(H_1(a,b))^2} \leq a(u_H, u_H) = a(u_h, u_H) \leq \mu_2 \|u_h\|_{(H_1(a,b))^2} \|u_H\|_{(H_1(a,b))^2},$$

that yields the estimate

$$\|u_H\|_{(H_1(a,b))^2} \leq \frac{\mu_2}{\mu_1} \|u_h\|_{(H_1(a,b))^2}. \tag{5.13}$$

Inserting the estimates (5.12) and (5.13) in equation (5.11) gives the following result:

$$\|u_h - u_H\|_{(L_2(a,b))^2} \leq \left(1 + \frac{\mu_2}{\mu_1}\right) \frac{\sqrt{1 + C_F^2}\mu_2}{\mu_3} H \|u_h\|_{(H^1(a,b))^2}. \tag{5.14}$$

Now by the definition of δ_H, Corollary 5.3 and (5.14), we can conclude, that

$$\delta_H = \sup_{u_h \in S_h^1} \frac{\|u_h - P_H u_h\|_{(L_2(a,b))^2}}{\|u_h\|_A} \leq \frac{1}{C_1} \sup_{u_h \in S_h^1} \frac{\|u_h - P_H u_h\|_{(L_2(a,b))^2}}{\|u_h\|_{(H^1(a,b))^2}}$$
$$\leq \left(1 + \frac{\mu_2}{\mu_1}\right) \frac{\sqrt{1 + C_F^2}\mu_2}{C_1 \mu_3} H.$$

This finishes the proof. □

From the previous discussion, we can conclude the following theorem.

Theorem 5.6. *If*

- *Corollary 5.3, Corollary 5.4 and Corollary 5.5 are valid,*
- *the coarse mesh size H is sufficiently small and*
- *B is a good preconditioner for A,*

then the Cai-Xu preconditioner is almost optimal and the convergence rates are given by Theorem 3.2.

Here the choice of the meshsize H is sufficient for the convergence of the method.
The next theorem provides information about the H^2-coercivity of the variational problem (4.5).

Theorem 5.7. *Let us assume, that the right-hand side f of the variational problem* (4.5) *is in L^2, i.e. $f \in \left(L^2(a,b)\right)^2$. Then, the solution $u \in \left(H_0^1(a,b)\right)^2$ of* (4.5) *fulfills the H^2-coercivity condition, i.e*

1. *$u \in \left(H_0^1(a,b)\right)^2 \cap \left(H^2(a,b)\right)^2$ and*
2. *$\exists \mu_3 > 0 : \|u\|_{(H^2(a,b))^2} \leq \mu_3 \|f\|_{(L^2(a,b))^2}$.*

Proof. Lax-Milgram guarantees the existence and uniqueness of a solution $u \in \left(H_0^1(a,b)\right)^2$ of the variational problem (4.5). The next step is to derive the weak derivative of order two of u. Therefore, we consider

$$a(u,v) = (f,v)_{(L^2(a,b))^2}, \quad \forall v \in \left(\dot{C}^\infty(a,b)\right)^2.$$

By the choice $v^s(x) = 0 \in \dot{C}^\infty(a,b)$, one obtains

$$\int_a^b \frac{du^c}{dx}\frac{dv^c}{dx}dx = \frac{1}{\nu}\int_a^b \left(f^c + \alpha\omega u^s\right) v^c dx, \quad \forall v^c \in \dot{C}^\infty(a,b).$$

By definition, the weak derivative of order two of u^c fulfills

$$-\int_a^b \frac{d^2u^c}{dx^2} v^c dx = \int_a^b \frac{du^c}{dx}\frac{dv^c}{dx}dx, \quad \forall v^c \in \dot{C}^\infty(a,b).$$

Since $(f^c + \alpha\omega u^s) \in L^2(a,b)$, we can conclude, that the weak derivative of order two of u^c is given by

$$\frac{d^2u^c}{dx^2} = -\frac{1}{\nu}\left(f^c + \alpha\omega u^s\right). \tag{5.15}$$

Similarly, by the choice $v^c(x) = 0$, we obtain the weak derivative of order two of u^s, namely

$$\frac{d^2u^s}{dx^2} = -\frac{1}{\nu}\left(f^s - \alpha\omega u^c\right). \tag{5.16}$$

Using (5.15) and triangle inequality, we arrive at

$$|u^c|_{H^2(a,b)} = \left\|\frac{1}{\nu}\left(f^c + \alpha\omega u^s\right)\right\|_{L^2(a,b)} \le \frac{1}{\underline{\nu}}\left(\|f^c\|_{L^2(a,b)} + \alpha\omega\,\|u^s\|_{L^2(a,b)}\right).$$

Analogous, using (5.16) and triangle inequality, we arrive at

$$|u^s|_{H^2(a,b)} \le \frac{1}{\underline{\nu}}\left(\|f^s\|_{L^2(a,b)} + \alpha\omega\,\|u^c\|_{L^2(a,b)}\right).$$

Combining the results and using Lax-Milgram, precisely the estimate

$$\|u\|_{(H^1(a,b))^2} \le \frac{1}{\mu_1}\,\|f\|_{(L^2(a,b))^2},$$

we can derive the estimation

$$\begin{aligned} |u|^2_{(H^2(a,b))^2} &\le \frac{1}{\underline{\nu}^2}\left(\|f\|^2_{(L^2(a,b))^2} + \alpha\omega\,\|u\|^2_{(L^2(a,b))^2}\right) \\ &\le \frac{1}{\underline{\nu}^2}\left(\frac{1}{\mu_1} + \alpha\omega\right)\|f\|^2_{(L^2(a,b))^2}. \end{aligned}$$

Hence, the second assertion follows. □

The Vectorspace $\mathbb{R}^{2N_h}$

For better understanding, we also want to formulate the three assumptions in the matrix-vector formulation. The proofs follow directly from the normequivalence given in (5.7).

Corollary 5.8.

$$\|\underline{u}_h\|_{(l^2)^2} \le C_1\,\|\underline{u}_h\|_{A_h} \quad \forall \underline{u}_h \in \mathbb{R}^{2N_h}.$$

for some constant C_1>0.

Corollary 5.9.

$$(M_h \underline{u}_h, \underline{v}_h) \leq C_2 \left\| \underline{u}_h \right\|_{(l^2)^2} \left\| \underline{v}_h \right\|_{A_h} \quad \forall \underline{u}_h, \underline{v}_h \in \mathbb{R}^{2N_h}.$$

for some constant $C_2 > 0$.

Corollary 5.10. *The parameter δ_H, defined by*

$$\delta_H = \sup_{\underline{v}_h \in \mathbb{R}^{2N_h}} \frac{\left\| \underline{v}_h - I_H^h P_H \underline{v}_h \right\|_{(l^2)^2}}{\left\| \underline{v}_h \right\|_{A_h}} \leq C_3 H.$$

fulfills the inequality

$$\delta_h \leq C_3 H$$

with some positive constant C_3, where the Ritz projection

$$\begin{aligned} P_H : & \mathbb{R}^{2N_h} && \to \mathbb{R}^{2N_H} \\ & P_H \underline{u}_h && = \underline{u}_H. \end{aligned}$$

is defined by the variational problem: Find $\underline{u}_H \in \mathbb{R}^{2N_H}$, such that

$$(A_h^M I_H^h \underline{u}_H, I_H^h \underline{v}_H) = (A_h^M \underline{u}_h, I_H^h \underline{v}_H) \quad \forall \underline{v}_H \in \mathbb{R}^{N_H^2}.$$

for a given right-hand side $\underline{u}_h \in \mathbb{R}^{2N_h}$.

5.2.3 Spectral Analysis

In fact we want to get an idea of the quantitative value of the convergence rates and the choice of the balancing parameter β. So we want to study the eigenvalues resp. singular values of the corresponding problem for the case that $\nu(x) = \nu = \text{const}$. The goal of this section is to obtain quantitative convergence results for the GMRES method, when the proposed Cai-Xu preconditioner is used.
Indeed we consider only two special cases, since the analysis of the full problem is too complex. Firstly, we use only the symmetric part as a preconditioner and completely neglect the skew symmetric one, i.e. we do not solve the system on the coarse subspace. This case can be interpreted as balancing the parameter $\beta = \infty$. Secondly, we use the full Cai-Xu preconditioner but we restrict the coarse subspace in the sense that $H = 2h$.
This section starts by computing eigenvalues of some needed matrices and concludes with presenting the corresponding convergence results.

Modelproblem

We start to study the eigenvalues resp. singular values of the corresponding problem for the case that $\nu(x) = \nu = \text{const}$.

$$A_h^M = \begin{pmatrix} L_h & D_h \\ -D_h & L_h \end{pmatrix} = \underbrace{\begin{pmatrix} L_h & 0 \\ 0 & L_h \end{pmatrix}}_{=A_h} + \underbrace{\begin{pmatrix} 0 & D_h \\ -D_h & 0 \end{pmatrix}}_{=M_h}$$

The stiffness matrix L_h and mass matrix D_h are given by

$$L_h = \frac{\nu(b-a)}{h}\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix} = \nu(b-a)A_1$$

$$D_h = \frac{\alpha\omega(b-a)h}{6}\begin{pmatrix} 4 & 1 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4 & 1 \\ 0 & \cdots & 0 & 1 & 4 \end{pmatrix} = \alpha\omega(b-a)M_1$$

where A_1 and M_1 are the stiffness matrix resp. mass matrix of the Laplace equation in 1D with homogeneous Dirichlet boundary.

Eigensystem of the Laplace equation in 1D

$$\begin{aligned} -\frac{\partial^2 u}{\partial x^2}(x) &= f(x), \qquad x \in (0,1), \\ u(0) = u(1) &= 0. \end{aligned}$$

Eigenvalues λ_k and eigenvectors v_k $(k = 1, \ldots, n-1)$ of the stiffness matrix $A_1 \in \mathbb{R}^{(n-1)\times(n-1)}$ are given by [16], namely

$$\lambda_k(A_1) = \frac{4}{h}\sin^2\left(\frac{k\pi}{2n}\right), \quad \text{where } h = \frac{1}{n}$$

and

$$\underline{v}_h^k = \sqrt{2h}\sin\left(kl\pi h\right)_{l=1}^{n-1}.$$

Note, that here we have

$$\lambda_{\min}(A_1) = \lambda_1(A_1) \leq \ldots \leq \lambda_{n-1}(A_1) = \lambda_{\max}(A_1).$$

Hence, the mass matrix M_1 has the same eigenfunctions we can compute the corresponding eigenvalues

$$\lambda_k(M_1) = \frac{h}{3}\left(3 - 2\sin^2\left(\frac{k\pi}{2n}\right)\right).$$

Note, that here we have

$$\lambda_{\max}(M_1) = \lambda_1(M_1) \geq \ldots \geq \lambda_{n-1}(M_1) = \lambda_{\min}(M_1).$$

Eigensystem of the matrix A_h^M

The eigenvalues of the matrices L_h and D_h are given by

$$\lambda_k^{(L)} = \lambda_k(L_h) = \frac{4\nu(b-a)}{h} \sin^2\left(\frac{k\pi}{2n}\right)$$

and

$$\lambda_k^{(D)} = \lambda_k(D_h) = \frac{\alpha\omega(b-a)h}{3}\left(3 - 2\sin^2\left(\frac{k\pi}{2n}\right)\right).$$

Now we have the following relations. Let $\lambda_k^{(A_M)} = \lambda_k(A_h^M)$. For $m = 1, \ldots, n-1$ we have

$$\lambda_{2m-1}^{(A_M)} = \lambda_m^{(L)} + i\lambda_m^{(D)} \quad \text{and} \quad \lambda_{2m}^{(A_M)} = \lambda_m^{(L)} - i\lambda_m^{(D)}$$

with the eigenvectors

$$\underline{v}_{(A_M)}^{2m-1} = \begin{pmatrix} -i\underline{v}^m \\ \underline{v}^m \end{pmatrix} \quad \text{and} \quad \underline{v}_{(A_M)}^{2m} = \begin{pmatrix} i\underline{v}^m \\ \underline{v}^m \end{pmatrix}.$$

Proof. For all odd m we have

$$\begin{aligned} A_h^M \underline{v}v_{(A_M)}^m &= \begin{pmatrix} L_h & D_h \\ -D_h & L_h \end{pmatrix} \begin{pmatrix} -i\underline{v}^m \\ \underline{v}^m \end{pmatrix} = \begin{pmatrix} -iL_h\underline{v}^m + D_h\underline{v}^m \\ iD_h\underline{v}^m + L_h\underline{v}^m \end{pmatrix} \\ &= \begin{pmatrix} -i\lambda_m^{(L)}\underline{v}^m + \lambda_m^{(D)}\underline{v}^m \\ i\lambda_m^{(D)}\underline{v}^m + \lambda_m^{(L)}\underline{v}^m \end{pmatrix} = \begin{pmatrix} -i\underline{v}^m(\lambda_m^{(L)} + i\lambda_m^{(D)}) \\ \underline{v}^m(i\lambda_m^{(D)} + \lambda_m^{(L)}) \end{pmatrix} \\ &= \lambda_m^{(A_M)}\underline{v}_{(A_M)}^m. \end{aligned}$$

For all even m we have

$$\begin{aligned} A_h^M \underline{v}_{(A_M)}^m &= \begin{pmatrix} L_h & D_h \\ -D_h & L_h \end{pmatrix} \begin{pmatrix} i\underline{v}^m \\ \underline{v}^m \end{pmatrix} = \begin{pmatrix} iL_h\underline{v}^m + D_h\underline{v}^m \\ -iD_h\underline{v}^m + L_h\underline{v}^m \end{pmatrix} \\ &= \begin{pmatrix} i\lambda_m^{(L)}\underline{v}^m + \lambda_m^{(D)}\underline{v}^m \\ -i\lambda_m^{(D)}\underline{v}^m + \lambda_m^{(L)}\underline{v}^m \end{pmatrix} = \begin{pmatrix} i\underline{v}^m(\lambda_m^{(L)} - i\lambda_m^{(D)}) \\ \underline{v}^m(-i\lambda_m^{(D)} + \lambda_m^{(L)}) \end{pmatrix} \\ &= \lambda_m^{(A_M)}\underline{v}_{(A_M)}^m. \end{aligned}$$

□

Eigensystem of the matrix $A_h^{-1}A_h^M$

Now we have the following relations. Let $\lambda_k^{(A^{-1}A_M)} = \lambda_k(A_h^{-1}A_h^M)$. For $m = 1, \ldots, n-1$ we have

$$\lambda_{2m-1}^{(A^{-1}A_M)} = 1 + i\frac{\lambda_m^{(L)}}{\lambda_m^{(D)}} \quad \text{and} \quad \lambda_{2m}^{(A^{-1}A_M)} = 1 - i\frac{\lambda_m^{(L)}}{\lambda_m^{(D)}}$$

with the eigenvectors

$$\underline{v}_{(A^{-1}A_M)}^{2m-1} = \underline{v}_{(A_M)}^{2m-1} = \begin{pmatrix} -i\underline{v}^m \\ \underline{v}^m \end{pmatrix} \quad \text{and} \quad \underline{v}_{(A^{-1}A_M)}^{2m} = \underline{v}_{(A_M)}^{2m} = \begin{pmatrix} i\underline{v}^m \\ \underline{v}^m \end{pmatrix}.$$

Proof. For all odd m we have

$$
\begin{aligned}
A_h^{-1} A_h^M \underline{v}_{(A_M)}^m &= \begin{pmatrix} L_h^{-1} & 0 \\ 0 & L_h^{-1} \end{pmatrix} \begin{pmatrix} L_h & D_h \\ -D_h & L_h \end{pmatrix} \begin{pmatrix} -i\underline{v}^m \\ \underline{v}^m \end{pmatrix} \\
&= \begin{pmatrix} I & L_h^{-1} D_h \\ -L_h^{-1} D_h & I \end{pmatrix} \begin{pmatrix} -i\underline{v}^m \\ \underline{v}^m \end{pmatrix} = \begin{pmatrix} -i\underline{v}^m + L_h^{-1} D_h \underline{v}^m \\ iL_h^{-1} D_h \underline{v}^m + I\underline{v}^m \end{pmatrix} \\
&= \begin{pmatrix} -i\underline{v}^m + \lambda_m^{(D)} L_h^{-1} \underline{v}^m \\ i\lambda_m^{(D)} L_h^{-1} \underline{v}^m + \underline{v}^m \end{pmatrix} = \begin{pmatrix} -i\underline{v}^m + \lambda_m^{(D)} \left(\lambda_m^{(L)}\right)^{-1} \underline{v}^m \\ i\lambda_m^{(D)} \left(\lambda_m^{(L)}\right)^{-1} \underline{v}^m + \underline{v}^m \end{pmatrix} \\
&= \lambda_m^{(A^{-1}A_M)} \underline{v}_{(A_M)}^m .
\end{aligned}
$$

For all even m we have

$$
\begin{aligned}
A_h^{-1} A_h^M \underline{v}_{(A_M)}^m &= \begin{pmatrix} L_h^{-1} & 0 \\ 0 & L_h^{-1} \end{pmatrix} \begin{pmatrix} L_h & D_h \\ -D_h & L_h \end{pmatrix} \begin{pmatrix} i\underline{v}^m \\ \underline{v}^m \end{pmatrix} \\
&= \begin{pmatrix} I & L_h^{-1} D_h \\ -L_h^{-1} D_h & I \end{pmatrix} \begin{pmatrix} i\underline{v}^m \\ \underline{v}^m \end{pmatrix} = \begin{pmatrix} i\underline{v}^m + L_h^{-1} D_h \underline{v}^m \\ -iL_h^{-1} D_h \underline{v}^m + I\underline{v}^m \end{pmatrix} \\
&= \begin{pmatrix} i\underline{v}^m + \lambda_m^{(D)} L_h^{-1} \underline{v}^m \\ -i\lambda_m^{(D)} L_h^{-1} \underline{v}^m + \underline{v}^m \end{pmatrix} = \begin{pmatrix} i\underline{v}^m + \lambda_m^{(D)} \left(\lambda_m^{(L)}\right)^{-1} \underline{v}^m \\ -i\lambda_m^{(D)} \left(\lambda_m^{(L)}\right)^{-1} \underline{v}^m + \underline{v}^m \end{pmatrix} \\
&= \lambda_m^{(A^{-1}A_M)} \underline{v}_{(A_M)}^m .
\end{aligned}
$$

□

Next we want to discuss the distribution of the imaginary part.

$$
\begin{aligned}
\frac{\lambda_m^{(L)}}{\lambda_m^{(D)}} &= \frac{12\nu \sin^2\left(\frac{k\pi}{2n}\right)}{h^2 \alpha\omega \left(3 - \sin^2\left(\frac{k\pi}{2n}\right)\right)} \\
&\leq \frac{12\nu \sin^2\left(\frac{(n-1)\pi}{2n}\right)}{h^2 \alpha\omega \left(3 - \sin^2\left(\frac{(n-1)\pi}{2n}\right)\right)} = O\left(h^{-2}\right)
\end{aligned}
$$

In Figure 5.2 the distribution of the imaginary part is shown. It's important to mention, that larger imaginary ratio slows down the convergence of the GMRES method.

Computation of the constants

The previous spectral analysis also delivers estimations for the constants C_1 and C_2 computed in Corollary 5.3 and Corollary 5.4. Estimate for $C1$:

$$
\|\underline{v}_h\|_{(l_2)^2} \leq \frac{1}{\sqrt{\lambda_{\min(L_h)}}} \|\underline{v}_h\|_{A_h} .
$$

Estimate for $C2$:

$$
|(M\underline{u}_h, \underline{v}_h)| \leq \frac{\left|\lambda_{\max(D_h)}\right|}{\sqrt{\lambda_{\min(L_h)}}} \|\underline{u}_h\|_{(l_2)^2} \|\underline{v}_h\|_{A_h} .
$$

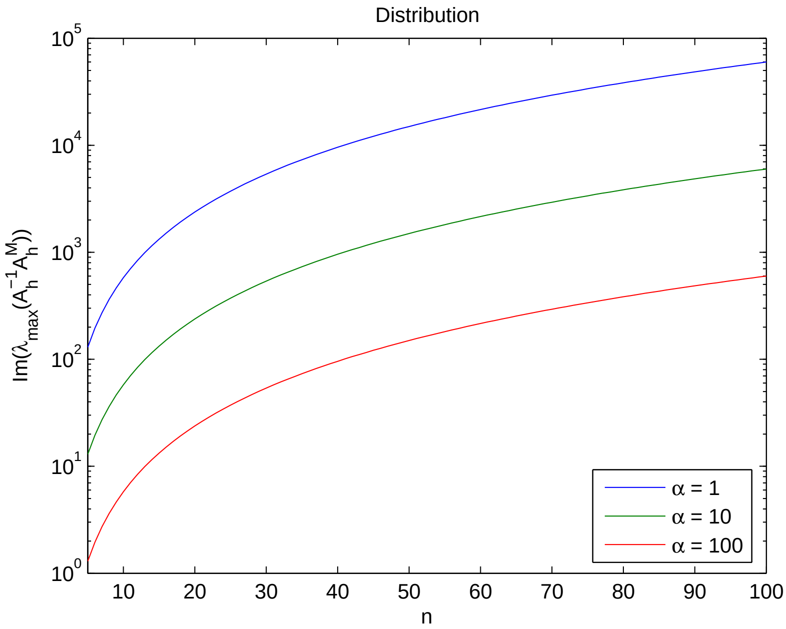

Figure 5.2: Distribution $\left|Im(\lambda_{max}(A_h^{-1}A_h^M)\right|$

Sharp computation of the convergence rate for GMRES (1)

Here we consider the case $\beta = \infty$. This means, that we just use the symmetric part as a preconditioner and neglect the non-symmetric. We want to compute the convergence rate of the resulting GMRES method.

Lemma 5.11 (Upper Bound)**.**

$$\left\|A_h^{-1}A_h^M \underline{u}_h\right\|_{A_h} \leq \mu \left\|\underline{u}_h\right\|_{A_h}$$

with $\mu = \sqrt{1 + \left(\frac{\lambda_1^{(D)}}{\lambda_1^{(L)}}\right)^2}$.

Proof. By simple manipulations and the substitution $\underline{u}u_h = A_h^{-1/2}\underline{w}_h$ we obtain

$$\begin{aligned}\left\|A_h^{-1}A_h^M \underline{u}_h\right\|_{A_h}^2 &= (A_h A_h^{-1}A_h^M \underline{u}_h, A_h^{-1}A_h^M \underline{u}_h) \\ &= (A_h^{-1}A_h^M \underline{u}_h, A_h^M \underline{u}_h) \\ &= ((A_h - M_h)A_h^{-1}(A_h + M_h)\underline{u}_h, \underline{u}_h) \\ &= ((A_h - M_h A_h^{-1} M_h)\underline{u}_h, \underline{u}_h) \\ &= (\underbrace{(I - A_h^{-1/2} M_h A_h^{-1} M_h A_h^{-1/2})}_{=:G_u} \underline{w}_h, \underline{w}_h).\end{aligned}$$

Now we want to compute the corresponding eigenvalues. The resulting operator G_u is SPD. Since all involved matrices $A_h^{-1/2}$, A_h^{-1} and M_h have the same eigenvectors, we can compute the corresponding eigenvalues, namely

$$\lambda_{2m}^{(G_u)} = \lambda_{2m-1}^{(G_u)} = 1 + \left(\frac{\lambda_m^{(D)}}{\lambda_m^{(L)}} \right)^2 .$$

The corresponding eigenvectors are given by

$$\underline{v}_{(G_u)}^{2m} = \underline{v}_{(G_u)}^{2m+1} = \begin{pmatrix} \underline{v}^m \\ \underline{v}^m \end{pmatrix} .$$

In fact, we have

$$\left\| A_h^{-1} A_h^M \underline{u}_h \right\|_{A_h} \leq \sqrt{1 + \left(\frac{\lambda_1^{(D)}}{\lambda_1^{(L)}} \right)^2} \, \|\underline{u}_h\|_{(l^2)^2} .$$

□

Lemma 5.12 (Lower bound)**.**

$$\left(A_h^{-1} A_h^M \underline{u}_h, \underline{u}_h \right)_{A_h} \geq \gamma \left(\underline{u}_h, \underline{u}_h \right)_{A_h}$$

with $\gamma = 1$*.*

Proof. By simple manipulations and the substitution $\underline{u}_h = A_h^{-1/2} \underline{w}_h$ we obtain

$$\begin{aligned} \left(A_h^{-1} A_h^M \underline{u}_h, \underline{u}_h \right)_{A_h} &= \left(A_h^M \underline{u}_h, \underline{u}_h \right) \\ &= (\underbrace{A_h^{-1/2} A_h^M A_h^{-1/2}}_{=:G_l} \underline{w}_h, \underline{w}_h) \end{aligned}$$

We have already computed the eigenvalues of G_l. For $k = 1, \ldots, 2n-2$

$$\lambda_k^{G_l} = \lambda_k^{(A_h^{-1} A_h^M)}$$

So due to Section 3.1 we can conclude, that the minimal eigenvalue of the symmetric part is 1 and therefore $\gamma = 1$. □

Theorem 5.13 (Case $\beta = \infty$)**.** *With the choice* $B_M = A_h^{-1}$ *we obtain the convergence rate*

$$\|\underline{r}_m\|_{A_h} \leq q^{m/2} \, \|\underline{r}_0\|_{A_h}$$

with

$$q = 1 - \frac{1}{1 + \frac{h^4 \alpha^2 \omega^2}{144 \nu^2} \left(\frac{3}{\sin^2\left(\frac{\pi h}{2}\right)} - 2 \right)^2} .$$

Additionally we have an asymptotic upper bound

$$q_{\lim} = \lim_{h \longrightarrow 0} q = \frac{\alpha^2 \omega^2}{\pi^4 \nu^2 + \alpha^2 \omega^2} \leq 1.$$

The corresponding upper bounds $q_{\lim}$ in dependence of the parameter α can be seen in Figure 5.3. The choice of the remaining parameters is $\omega = 1$ and $\nu = 1$.

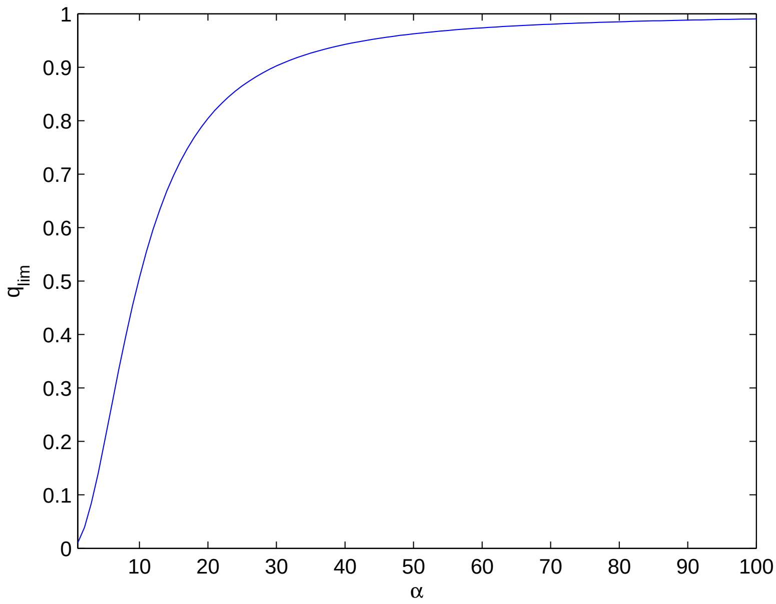

Figure 5.3: Convergence rate $q_{\lim}$ for $\beta = \infty$ in dependence of α

Sharp computation of the convergence rate for GMRES (2)

We choose $B_h = A_h^{-1}$ and therefore want to compute the eigensystem of

$$\begin{aligned} B^M A_h^M &= \left(I_H^h \left(A_H^M\right)^{-1} Q_H + \beta A_h^{-1} \right) A_h^M \\ &= I_H^h \left(A_H^M\right)^{-1} Q_H A_h^M + \beta A_h^{-1} A_h^M \end{aligned}$$

Since there are two different grids involved, the computation of the singular values seems to be very difficult in general. In [22], a spectral analysis for multigrid method was developed for the special case $H = 2h$. We follow those computings to obtain the eigenvalues of the operator of interest for this special case.

In fact we will consider not the L^2-projection Q_H, but the weighted restriction I_h^H. This simplifies the analysis since we can use the property $I_h^H = \left(I_H^h\right)^T$. So the Cai-Xu preconditioner reads as

$$B^M = I_H^h \left(A_H^M\right)^{-1} I_h^H + \beta A_h^{-1}. \tag{5.17}$$

Lemma 5.14 (Lower Bound)**.** *For B^M given by* (5.17)*, we have*

$$(B^M A_h^M \underline{u}_h, \underline{u}_h)_{A_h} \geq \gamma (\underline{u}_h, \underline{u}_h)_{A_h}$$

with $\gamma = \beta$.

Proof. By simple manipulations and $\underline{u}_h = A_h^{-1/2} \underline{w}_h$.

$$\begin{aligned} (B^M A_h^M \underline{u}_h, \underline{u}_h)_A &= ((A_h I_H^h \left(A_H^M\right)^{-1} I_h^H A_h^M + \beta A_h^M) \underline{u}_h, \underline{u}_h) \\ &= (A_h^{1/2} I_H^h \left(A_H^M\right)^{-1} I_h^H A_h^M A_h^{-1/2} \underline{w}_h, \underline{w}_h) + \beta (A_h^{-1/2} A_h^M A_h^{-1/2} \underline{w}_h, \underline{w}_h). \end{aligned}$$

Let $C = I_H^h \left(A_H^M\right)^{-1} I_h^H A_h^M$. So we are interested in $\lambda(C)$.

$$\left(A_H^M\right)^{-1} = \begin{pmatrix} \left(L_H + D_H L_H^{-1} D_H\right)^{-1} & L_H^{-1} D_H \left(L_H + D_H L_H^{-1} D_H\right)^{-1} \\ -L_H^{-1} D_H \left(L_H + D_H L_H^{-1} D_H\right)^{-1} & \left(L_H + D_H L_H^{-1} D_H\right)^{-1} \end{pmatrix} =: \begin{pmatrix} S_H & T_H \\ -T_H & S_H \end{pmatrix}$$

We have

$$\lambda_k^{(S_H)} = \frac{\lambda_k^{(L)}}{\left(\lambda_k^{(L)}\right)^2 + \left(\lambda_k^{(D)}\right)^2} \quad \text{and} \quad \lambda_k^{(T_H)} = \frac{\lambda_k^{(D)}}{\left(\lambda_k^{(L)}\right)^2 + \left(\lambda_k^{(D)}\right)^2}.$$

Now we arrive at the matrix C, precisely

$$C = \begin{pmatrix} \tilde{I}_H^h S_H \tilde{I}_h^H L_h - \tilde{I}_H^h T_H \tilde{I}_h^H D_h & \tilde{I}_H^h S_H \tilde{I}_h^H D_h + \tilde{I}_H^h T_H \tilde{I}_h^H L_h \\ -(\tilde{I}_H^h S_H \tilde{I}_h^H D_h + \tilde{I}_H^h T_H \tilde{I}_h^H L_h) & \tilde{I}_H^h S_H \tilde{I}_h^H L_h - \tilde{I}_H^h T_H \tilde{I}_h^H D_h \end{pmatrix} := \begin{pmatrix} C_1 & C_2 \\ -C_2 & C_1 \end{pmatrix}.$$

Since multiplication with unitary matrices doesn't change the spectral radius, we consider the following expression amongst others

$$Q_h^T C_1 Q_h = Q_h^T \tilde{I}_H^h Q_H Q_H^T S_H Q_H Q_H^T \tilde{I}_h^H Q_h Q_h^T L_h Q_H,$$

where the unitary matrices $Q_h^T = Q_h^{-1}$ and $Q_H^T = Q_H^{-1}$ are given by

$$Q_h = \left(\underline{v}_h^1, \underline{v}_h^{n-1}, \underline{v}_h^2, \underline{v}_h^{n-2}, \ldots, \underline{v}_h^{n/2-1}, \underline{v}_h^{n/2+1}, \underline{v}_h^{n/2}\right) \quad \text{and} \quad Q_H = \left(\underline{v}_H^1, \underline{v}_H^2 1, \ldots, \underline{v}_H^{n/2-1}\right).$$

We have the representation

$$Q_h^T L_h Q_h = \begin{pmatrix} L_h^{(1)} & & & \\ & L_h^{(2)} & & \\ & & \ddots & \\ & & & L_h^{(n/2)} \end{pmatrix},$$

with the blocks

$$L_h^{(k)} = \frac{4\nu(b-a)}{h} \begin{pmatrix} \sin^2\left(\frac{1}{2}k\pi h\right) & 0 \\ 0 & \cos^2\left(\frac{1}{2}k\pi h\right) \end{pmatrix} \quad k = 1, \ldots, \frac{n}{2} - 1$$

$$L_h^{(n/2)} = \frac{2h}{3}.$$

Similarly, we have the representation

$$Q_h^T D_h Q_h = \begin{pmatrix} D_h^{(1)} & & & \\ & D_h^{(2)} & & \\ & & \ddots & \\ & & & D_h^{(n/2)} \end{pmatrix},$$

with the blocks

$$D_h^{(k)} = \frac{\alpha\omega(b-a)h}{3} \begin{pmatrix} 3 - 2\sin^2\left(\frac{1}{2}k\pi h\right) & 0 \\ 0 & 3 - 2\cos^2\left(\frac{1}{2}k\pi h\right) \end{pmatrix} \quad k = 1, \ldots, \frac{n}{2} - 1$$

$$D_h^{(n/2)} = \frac{2}{h}.$$

We also obtain

$$Q_H^T S_H Q_H = \operatorname{diag}\left(\lambda_k^{(S_H)}\right)_{k=1}^{n/2-1}$$
$$Q_H^T T_H Q_H = \operatorname{diag}\left(\lambda_k^{(T_H)}\right)_{k=1}^{n/2-1}.$$

Hence, have the following relation for $k = 1, \ldots, \frac{n}{2} - 1$:

$$\tilde{I}_h^H \underline{v}_h^k = 2\cos^2\left(\frac{1}{2}k\pi h\right)\underline{v}_H^k$$
$$\tilde{I}_h^H \underline{v}_h^{n-k} = -2\sin^2\left(\frac{1}{2}k\pi h\right)\underline{v}_H^k$$

Therefore, we get

$$Q_H^T \tilde{I}_h^H Q_h = \begin{pmatrix} r_{(1)} & & & & 0 \\ & r_{(2)} & & & 0 \\ & & \ddots & & \vdots \\ & & & r_{(n/2-1)} & 0 \end{pmatrix},$$

where

$$r_{(k)} = \sqrt{2}\left(\cos^2\left(\frac{1}{2}k\pi h\right), -\sin^2\left(\frac{1}{2}k\pi h\right)\right).$$

Since $I_H^h = \left(I_h^H\right)^T$, we can derive

$$Q_h^T \tilde{I}_H^h Q_H = \begin{pmatrix} r_{(1)}^T & & & \\ & r_{(2)}^T & & \\ & & \ddots & \\ & & & r_{(n/2-1)}^T \\ 0 & 0 & \ldots & 0 \end{pmatrix}.$$

Using the abbreviation $c_k = \cos\left(\frac{1}{2}k\pi h\right)$ and $s_k = \sin\left(\frac{1}{2}k\pi h\right)$, we can conclude, that

$$Q_h^T C_1 Q_h = \begin{pmatrix} C_1^{(1)} & & & \\ & C_1^{(2)} & & \\ & & \ddots & \\ & & & C_1^{(n/2)} \end{pmatrix},$$

where the blocks $C_1^{(k)}$ are given by

$$\begin{aligned} C_1^{(k)} &= \frac{8\nu(b-a)}{h}\lambda_k^{(S_H)} c_k^2 s_k^2 \begin{pmatrix} c_k^2 & -c_k^2 \\ -s_k^2 & s_k^2 \end{pmatrix} \\ &\quad - \frac{2\alpha\omega(b-a)h}{3}\lambda_k^{(T_H)} \begin{pmatrix} c_k^4(3-2s_k^4) & -c_k^2 s_k^2(3-2c_k^2) \\ -c_k^2 s_k^2(3-2s_k^2) & s_k^4(3-2c_k^4) \end{pmatrix} \quad k = 1, \ldots, \frac{n}{2}-1 \\ C_1^{(n/2)} &= 0. \end{aligned}$$

Therefore we have for $k = 1, \dots, \frac{n}{2} - 1$

$$\lambda_k^{(C_1)} = \frac{2\nu(b-a)}{h}\lambda_k^{(S_H)} \sin^2(k\pi h) - \frac{2\alpha\omega(b-a)h}{3}\lambda_k^{(T_H)}(2+\cos(2k\pi h)), \quad k = 1, \dots, \frac{n}{2} - 1$$
$$\lambda_k^{(C_1)} = 0, \quad k = \frac{n}{2}, \dots, n-1.$$

Similarly we can compute $\lambda_k^{(C_2)}$.

$$\lambda_k^{(C_2)} = \frac{2\nu(b-a)}{h}\lambda_k^{(T_H)} \sin^2(k\pi h) + \frac{2\alpha\omega(b-a)h}{3}\lambda_k^{(S_H)}(2+\cos(2k\pi h)), \quad k = 1, \dots, \frac{n}{2} - 1$$
$$\lambda_k^{(C_2)} = 0, \quad k = \frac{n}{2}, \dots, n-1.$$

Both $\lambda_k^{(C_1)}$ and $\lambda_k^{(C_2)}$ are equipped with the eigenvectors $\underline{v}^k$. For better understanding we mention, that the following order is given.

$$\lambda_{\min}^{(C_1)} = \lambda_1^{(C_1)} \le \dots \le \lambda_{\frac{n}{2}-1}^{(C_1)} = \lambda_{\max}^{(C_1)}$$
$$\lambda_{\max}^{(C_2)} = \lambda_1^{(C_2)} \ge \dots \ge \lambda_{\frac{n}{2}-1}^{(C_2)} = \lambda_{\max}^{(C_2)}$$

Consequently, for $k = 1, \dots, n-1$ we have

$$\lambda_{2k-1}^{(C)} = \lambda_k^{(C_1)} + i\lambda_k^{(C_2)} \quad \text{and} \quad \lambda_{2k}^{(C)} = \lambda_k^{(C_1)} - i\lambda_k^{(C_2)}$$

with the eigenvectors

$$\underline{v}_{(C)}^{2k-1} = \begin{pmatrix} -i\underline{v}^k \\ \underline{v}^k \end{pmatrix} \quad \text{and} \quad \underline{v}_{(C)}^{2k} = \begin{pmatrix} i\underline{v}^k \\ \underline{v}^k \end{pmatrix}.$$

Finally, we can conclude that $\lambda_{\min}^{(C^{\text{sym}})} = 0$ and therefore $\lambda_{\min}^{((B_M A_h^M)^{\text{sym}})} = \beta$. □

Lemma 5.15 (Upper Bound)**.** *For B_M given by (5.17), we have*

$$\left\|B_M A_h^M u_h\right\|_{A_h} \le \mu \left\|u_h\right\|_{A_h}$$

with $\mu^2 = 1 + \beta\mu^{(1)} + \beta^2\mu^{(2)}$, where the generic constants $\mu^{(1)}$ and $\mu^{(2)}$ are given by the proof.

Proof. We use the computations performed in the proof of the previous lemma to gain the upper bound. Decomposing the squared norm gives

$$\left\|B_M A_h^M u_h\right\|_{A_h}^2 = \left\|(I_H^h \left(A_H^M\right)^{-1} I_h^H A_h^M + \beta A_h^{-1} A_h^M) u_h\right\|_{A_h}^2 = \left\|I_H^h \left(A_H^M\right)^{-1} I_h^H A_h^M u_h\right\|_{A_h}^2$$
$$+ \beta(\left(A_h^M\right)^T I_H^h \left(\left(A_H^M\right)^{-1} + \left(A_H^M\right)^{-T}\right) I_h^H A_h^M u_h, u_h) + \beta^2 \left\|A_h^{-1} A_h^M u_h\right\|_{A_h}^2.$$

Now we estimate each summand individually. Since we have already computed all eigenvalues appearing in these three expressions, we just have to collect the effort of previous computations. For the first term we obtain

$$\begin{aligned}\left\|I_H^h \left(A_H^M\right)^{-1} I_h^H A_h^M u_h\right\|_{A_h}^2 &= (A_h^{-1/2} \left(A_h^M\right)^T I_H^h \left(A_H^M\right)^{-T} I_h^H A_h I_H^h \left(A_H^M\right)^{-1} I_h^H A_h^M A_h^{-1/2} w_h, w_h) \\ &= (A_h^{-1/2} C^T A_h C A_h^{-1/2} w_h, w_h) \\ &\le \bar{\lambda}_{\frac{n}{2}-1}^{(C)} \lambda_{\frac{n}{2}-1}^{(C)} (w_h, w_h) \\ &= \mu^{(0)} (w_h, w_h).\end{aligned}$$

The second term gives

$$\begin{aligned}(A_h^{-1/2}\left(A_h^M\right)^T I_H^h\left(\left(A_H^M\right)^{-1}+\left(A_H^M\right)^{-T}\right)I_h^H A_h^M A_h^{-1/2}w_h,w_h)\\ \leq \frac{\lambda_1^{(A_h^M)}\bar{\lambda}_1^{(A_h^M)}}{\lambda_1^{(L_h)}}2\lambda_1^{(S_H)}\frac{1}{4}(3+\cos(2\pi h))(w_h,w_h)\\ =\mu^{(1)}(w_h,w_h).\end{aligned}$$

The third term can be estimated by

$$\left\|A_h^{-1}A_h^M u_h\right\|_{A_h}^2 \leq \lambda_1^{(G_u)}(w_h,w_h)=\mu^{(2)}(w_h,w_h).$$

So the combination of the estimates gives the desired result. □

Theorem 5.16 (Case $H = 2h$). *With the choice* $B^M = I_H^h\left(A_H^M\right)^{-1}I_h^H+\beta A_h^{-1}$ *we obtain the convergence rate*

$$\|r_m\|_{A_h}\leq q^{m/2}\|r_0\|_{A_h}$$

with

$$q=1-\frac{\beta^2}{\mu^{(0)}+\beta\mu^{(1)}+\beta^2\mu^{(2)}}$$

where the generic constants $\mu^{(0)}$,$\mu^{(1)}$ *and* $\mu^{(2)}$ *are given by the proof. Additionally we have an asymptotic upper bound*

$$q_{\lim}=\lim_{h\longrightarrow 0}q=\frac{\pi^4\nu^2+\beta^2\alpha^2\omega^2}{\pi^4(1+\beta^2)\nu^2+\beta^2\alpha^2\omega^2}\leq 1$$

The corresponding upper bounds $q_{\lim}$ in dependence of the parameter α can be seen in Figure 5.4. The corresponding upper bounds $q_{\lim}$ in dependence of the parameter β can be seen in Figure 5.5. The choice of the remaining parameters is $\omega = 1$ and $\nu = 1$. In fact we have also the following rigorous bound for β tending to infinity:

$$\lim_{\beta\longrightarrow\infty}\frac{\pi^4\nu^2+\beta^2\alpha^2\omega^2}{\pi^4(1+\beta^2)\nu^2+\beta^2\alpha^2\omega^2}=\frac{\alpha^2\omega^2}{\pi^4\nu^2+\alpha^2\omega^2}.$$

Indeed $q_{\lim}$ is strict decrising in β. Therefore we attain the best convergence results for $\beta = \infty$. However we have to keep in mind, that we have to balance between the non-symmetric and symmetric part, and therefore this choice has to be investigated.

Conclusion

We have seen, that for $\alpha \longrightarrow \infty$ the convergence rate q tends to 1. By varying α we can easily simulate various choices of the higher frequencies $k\omega$, that are point of interest in the nonlinear case. So we can conclude, that although the Cai-Xu preconditioner promises quite good results, it is not robust in $k\omega$.

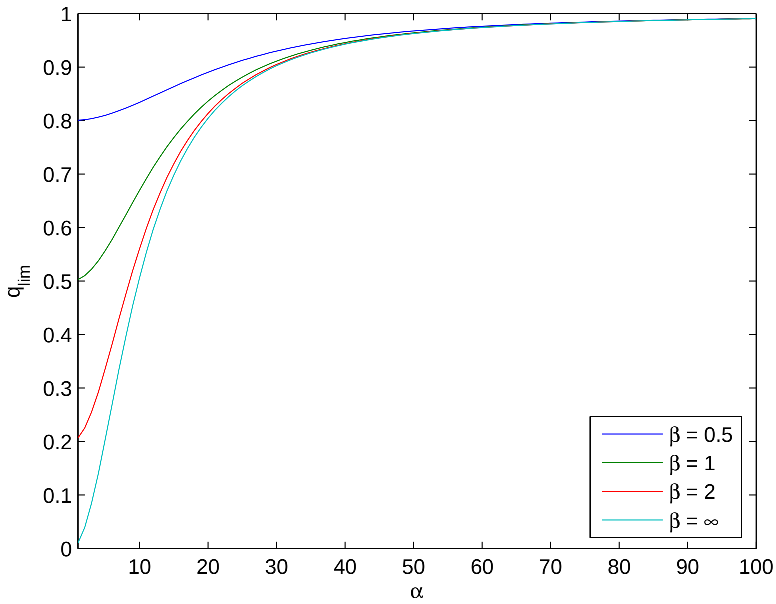

Figure 5.4: Convergence rate $q_{\lim}$ in dependence of α

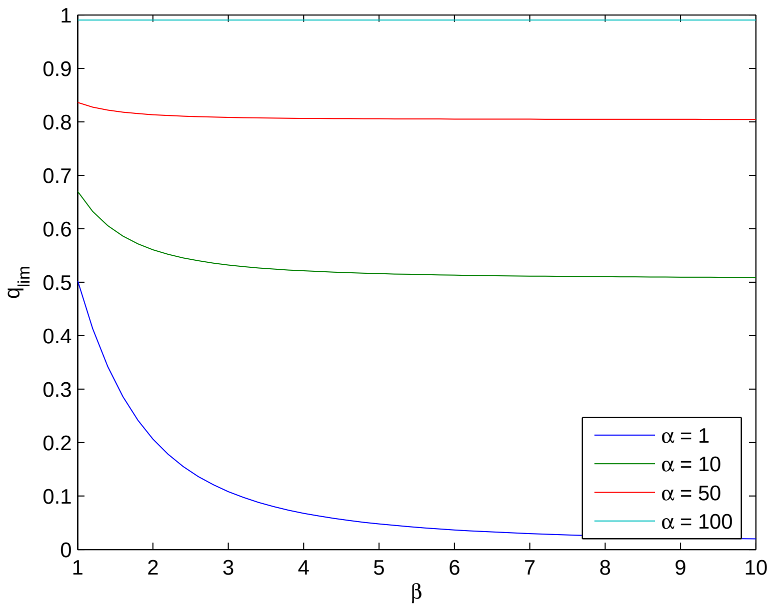

Figure 5.5: Convergence rate $q_{\lim}$ in dependence of β

5.3 The Nonlinear Problem

In this section we explore the nonlinear problem. The resulting system of nonlinear equations has to be solved by an appropriate iterative procedure, whereas our choice is the well known

Newton's method. So the proceedings turn to be in that way: We start by stating the already derivated variational formulation. In order to apply Newton's method the main focus is on building up the Frechet derivative of the nonlinear operator $L^{(N)}$ and building up the resulting finite element system.

5.3.1 Variational Formulation

As our starting point we use the variational formulation (4.14) achieved in Chapter 4.

$$\int_a^b \alpha\omega D_N u(x) \cdot v(x) + \psi[u](x) \cdot \nabla v(x) dx = \int_a^b f(x) \cdot v(x) dx$$

Now we have the nonlinear left-hand side

$$a(u,v) = \int_a^b \alpha\omega D_N u(x) \cdot v(x) + \psi[u](x) \cdot \nabla v(x) dx$$

and the linear right-hand side

$$(f,v) = \int_a^b f(x) \cdot v(x) dx.$$

This gives the nonlinear problem. Find $u \in \left(H_0^1(a,b)\right)^{2N+1}$, such that

$$a(u,v) = (f,v) \quad \forall v \in \left(H_0^1(a,b)\right)^{2N+1}. \tag{5.18}$$

For all $u, v \in \left(H_0^1(a,b)\right)^{2N+1}$, we define the operator $L^{(N)}$ by

$$\left(L^{(N)}[u], v\right) = \int_a^b \psi[u](x) \cdot \nabla v(x) dx.$$

So (5.18) now reads as:
Find $u \in \left(H_0^1(a,b)\right)^{2N+1}$, such that

$$\left(L^{(N)}[u], v\right) + \omega\alpha \left(D_N u, v\right) = (f,v) \quad \forall v \in \left(H_0^1(a,b)\right)^{2N+1}. \tag{5.19}$$

5.3.2 Newton's Method

Dealing with nonlinear problems, Newton's method is one of the most widely used iterative procedure for obtaining a solution. The prevalent use of this iterative technique is due to its fast convergence. The Newton iteration converges locally q-superlinear and under certain conditions even q-quadratic.

Framework of Newton's Method

The general idea of Newton's method reads as follows. We have two Banach spaces X and Y and a nonlinear mapping $F : X \longrightarrow Y$, that is Frechet-differentiable in an open subset of X. In fact we are interested in solving the nonlinear problem

$$F(x) = y$$

for some $y \in Y$. Therefore we apply the following iterative linearization procedure

$$x_{k+1} = x_k + \left(F'(x_k)\right)^{-1} \left(y - F(x_k)\right). \tag{5.20}$$

Theorem 5.17. *Let X and Y be Banach spaces, let $D \subset X$ be non-empty and open. Suppose the map $F : D \longrightarrow Q$ is Frechet-differentiable in D, and let $x^* \in D$ be a solution of $F(x) = y$ with regular derivative $F'(x^*)$. Then,*

- *if F' is continuous in x^*, the Newton iteration* (5.20) *is locally superlinear convergent.*
- *if there exists $\gamma > 0$ such that*

$$\left\|F'(x) - F'(x^*)\right\| \leq \gamma \left\|x - x^*\right\| \forall x \in U(x^*) \subset D,$$

the Newton iteration (5.20) *is locally quadratically convergent.*

Computing the Frechet Derivative

In order to apply Newton's method for solving the nonlinear multiharmonic problem (5.19), we have to calculate the Frechet derivative of the operators $L^{(N)}$ and D_N. Since D_N is a linear operator, it's derivative is just D_N again. For the nonlinear operator $L^{(N)}$, we first try to calculate the Gateaux derivative. In fact we want to compute the Gateaux differential of $L^{(N)}$ in direction w, denoted by $D_w(L^{(N)})$. So let $w \in \left(H_0^1(a,b)\right)^{2N+1}$ be an arbitrary direction. With the notation $p = \nabla u$ we can conclude

$$\begin{aligned}\left(D_w(L^{(N)})[u], v\right) &= \lim_{\delta \to 0} \frac{1}{\delta} \int_a^b \left(\psi[u + \delta w](x) - \psi[u](x)\right) \cdot \nabla v(x) dx \\ &= \int_a^b \lim_{\delta \to 0} \frac{1}{\delta} \left(\psi[u + \delta w](x) - \psi[u](x)\right) \cdot \nabla v(x) dx \\ &= \int_a^b D_w(\psi)[u] \cdot \nabla v dx = \int_a^b \left(\frac{\partial \psi}{\partial p}[u] \nabla w\right) \cdot \nabla v dx.\end{aligned} \tag{5.21}$$

Next we want to investigate the Jacobian $\frac{\partial \psi}{\partial p}[u]$. Indeed $\frac{\partial \psi}{\partial p}[u]$ is a $(2N+1) \times (2N+1)$ matrix containing all partial derivatives. In order to state the Jacobian, we have to clarify some notations.

$$p = \begin{pmatrix} p_0^c \\ p_1^c \\ p_1^s \\ \vdots \\ p_N^c \\ p_N^s \end{pmatrix} = \begin{pmatrix} \frac{du_0^c}{dx} \\ \frac{du_1^c}{dx} \\ \frac{du_1^s}{dx} \\ \vdots \\ \frac{du_N^c}{dx} \\ \frac{du_N^s}{dx} \end{pmatrix} = \nabla u$$

So the Jacobian reads as

$$\frac{\partial \psi}{\partial p}[u] = \begin{pmatrix} \frac{\partial \psi_0^c}{\partial p_0^c}[u] & \frac{\partial \psi_0^c}{\partial p_1^c}[u] & \frac{\partial \psi_0^c}{\partial p_1^s}[u] & \cdots & \frac{\partial \psi_0^c}{\partial p_N^c}[u] & \frac{\partial \psi_0^c}{\partial p_N^s}[u] \\ \frac{\partial \psi_1^c}{\partial p_0^c}[u] & \frac{\partial \psi_1^c}{\partial p_1^c}[u] & \frac{\partial \psi_1^c}{\partial p_1^s}[u] & \cdots & \frac{\partial \psi_1^c}{\partial p_N^c}[u] & \frac{\partial \psi_1^c}{\partial p_N^s}[u] \\ \frac{\partial \psi_1^s}{\partial p_0^c}[u] & \frac{\partial \psi_1^s}{\partial p_1^c}[u] & \frac{\partial \psi_1^s}{\partial p_1^s}[u] & \cdots & \frac{\partial \psi_1^s}{\partial p_N^s}[u] & \frac{\partial \psi_1^s}{\partial p_N^s}[u] \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \frac{\partial \psi_N^c}{\partial p_0^c}[u] & \frac{\partial \psi_N^c}{\partial p_1^c}[u] & \frac{\partial \psi_N^c}{\partial p_1^s}[u] & \cdots & \frac{\partial \psi_N^c}{\partial p_N^c}[u] & \frac{\partial \psi_N^c}{\partial p_N^s}[u] \\ \frac{\partial \psi_N^s}{\partial p_0^c}[u] & \frac{\partial \psi_N^s}{\partial p_1^c}[u] & \frac{\partial \psi_N^s}{\partial p_1^s}[u] & \cdots & \frac{\partial \psi_N^s}{\partial p_N^c}[u] & \frac{\partial \psi_N^s}{\partial p_N^s}[u] \end{pmatrix}.$$

Next we want to compute the individual entries of $\frac{\partial \psi}{\partial p}[u]$ and therefore we introduce the following notation for the first derivative of $u(t)$.

$$p(t) = \sum_{k=0}^{N} p_k^c \cos(k\omega t) + p_k^s \sin(k\omega t).$$

We will compute one specific element of the Jacobian by using simple analysis. By using chain rule we obtain

$$\begin{aligned}
\frac{\partial \psi_k^c}{\partial p_l^c}(p) &= \frac{\partial}{\partial p_l^c}\left(\frac{2}{T}\int_0^T \nu\left(|p(t)|\right) p(t) \cos(k\omega t) dt\right) \\
&= \frac{2}{T}\int_0^T \frac{\partial}{\partial p(t)}\left(\nu\left(|p(t)|\right) p(t)\right) \frac{\partial p(t)}{\partial p_l^c} \cos(k\omega t) dt \\
&= \frac{2}{T}\int_0^T \frac{\partial}{\partial p(t)}\left(\nu\left(|p(t)|\right) p(t)\right) \cos(l\omega t)\cos(k\omega t) dt.
\end{aligned}$$

Applying product rule gives the following result:

$$\frac{\partial}{\partial p(t)}\left(\nu\left(|p(t)|\right) p(t)\right) = \nu\left(|p(t)|\right) + \nu'\left(|p(t)|\right)|p(t)|$$

Analogous the remaining elements of the Jacobian can be computed. Summarising we have the following representations of the Jacobian elements

$$\begin{aligned}
\frac{\partial \psi_k^c}{\partial p_l^c}(p) &= \frac{2}{T}\int_0^T \big(\nu\left(|p(t)|\right) + \nu'\left(|p(t)|\right)|p(t)|\big) \cos(l\omega t)\cos(k\omega t) dt \\
\frac{\partial \psi_k^c}{\partial p_l^s}(p) &= \frac{2}{T}\int_0^T \big(\nu\left(|p(t)|\right) + \nu'\left(|p(t)|\right)|p(t)|\big) \sin(l\omega t)\cos(k\omega t) dt \\
\frac{\partial \psi_k^s}{\partial p_l^c}(p) &= \frac{2}{T}\int_0^T \big(\nu\left(|p(t)|\right) + \nu'\left(|p(t)|\right)|p(t)|\big) \cos(l\omega t)\sin(k\omega t) dt \\
\frac{\partial \psi_k^s}{\partial p_l^s}(p) &= \frac{2}{T}\int_0^T \big(\nu\left(|p(t)|\right) + \nu'\left(|p(t)|\right)|p(t)|\big) \sin(l\omega t)\sin(k\omega t) dt.
\end{aligned}$$

Obviously the Gateaux differential $D_w(L^{(N)})[u]$ given by (5.21) is linear and continuous in w, so $L^{(N)}$ is Gateaux differentiable in u with the Gateaux derivative $L^{(N)\prime}[u]$ defined by

$$L^{(N)\prime}[u]w := D_w(L^{(N)}[u])$$

The Gauteaux derivative $L^{(N)\prime}[u]$ exists for all u and is continuous, so $L^{(N)\prime}$ is the Frechet derivative.
Now the Newton iteration applied to the nonlinear multiharmonic problem (5.19) reads as follows.

1. Choose some initial guess u^0.

2. For $k = 0, 1, \ldots$ compute the defect

$$(r^k, v) = \left(f - \left(L^{(N)}[u^k] + \omega\alpha D_N u^k\right), v\right).$$

3. Solve the linear variational problem

$$\left(\left(L^{(N)\prime}[u^k] + \omega\alpha D_N\right) w^k, v\right) = \left(r^k, v\right).$$

4. Update the solution

$$u^{k+1} = u^k + w^k.$$

Properties of the Frechet derivative $L^{(N)\prime}[u]$

Similar to the linear case we have the following relations. The operator $L^{(N)\prime}[u]$ is symmetric and positive definite and D_N is skew symmetric. In order to show the positive definiteness of $L^{(N)\prime}[u]$ we have to verify the following spectral equivalence.

$$\int_a^b \frac{\partial \psi}{\partial p}[u]\nabla w \cdot \nabla w dx \cong \int_a^b B[u]\nabla w \cdot \nabla w dx, \quad \forall w \in \left(H_0^1(a,b)\right)^{2N+1} \tag{5.22}$$

with

$$B[u] = \operatorname{diag}\left(\frac{\partial \psi_0^c}{\partial p_0^c}[u], \frac{\partial \psi_1^c}{\partial p_1^c}[u], \frac{\partial \psi_1^s}{\partial p_1^s}[u], \ldots, \frac{\partial \psi_N^c}{\partial p_N^c}[u], \frac{\partial \psi_N^s}{\partial p_N^s}[u]\right).$$

Since $B[u]$ is positive definite, this property is also passed to $L^{(N)\prime}[u]$.

Corollary 5.18. *Let us assume that*

1. $\nu(\cdot) \in C^1(\mathbb{R}^+ \longrightarrow \mathbb{R})$ *and* $0 < \underline{\nu} \le \nu(s) \le \overline{\nu}$ *for all* $s \in \mathbb{R}_0^+$,
2. $s \mapsto \nu(s)s$ *is Lipschitz continuous with Lipschitz constant* L, *i.e*

$$|\nu(s)s - \nu(t)t| \le L\,|s-t|\,,$$

3. $s \mapsto \nu(s)s$ *is strongly monotone with monotonicity constant* m, *i.e*

$$(\nu(s)s - \nu(t)t)(s-t) \ge m\,|s-t|^2\,.$$

Then, we have the norm equivalence

$$\int_a^b \frac{\partial \psi}{\partial p}[u]\nabla w \cdot \nabla w dx \cong |w|^2_{(H^1(a,b))^{2N+1}}, \quad \forall w \in \left(H_0^1(a,b)\right)^{2N+1}.$$

Proof. The strong monotonicity of the operator $L^{(N)}$ is due to the strong monotonicity of the mapping $s \mapsto \nu(s)s$, see [21, 13], namely

$$\left\langle L^{(N)}[u] - L^{(N)}[v], u - v\right\rangle \ge m\,\|u-v\|^2_{(H^1(a,b))^{2N+1}}\,. \tag{5.23}$$

The Lipschitz-continuity of the operator $L^{(N)}$ is due to the Lipschitz-continuity of the mapping $s \mapsto \nu(s)s$, see [21, 13], namely

$$\sup_{\|w\|=1}\left|\left\langle L^{(N)}[u] - L^{(N)}[v], w\right\rangle\right| = \left\|L^{(N)}[u] - L^{(N)}[v]\right\| \le 3L\,\|u-v\|_{(H^1(a,b))^{2N+1}}\,. \tag{5.24}$$

Since we know the Frechet derivative $L^{(N)\prime}$ of $L^{(N)}$, the following representation is valid

$$L^{(N)}[u] - L^{(N)}[v] = \int_0^1 \left\langle L^{(N)\prime}[u + t(v-u)], v-u \right\rangle dt,$$

indeed, by the mean value theorem of integration, we arrive at

$$\begin{aligned}\int_a^b (\psi[u] - \psi[v])\nabla w dx &= \int_a^b \int_0^1 \left(\frac{\partial \psi}{\partial p}[u + t(v-u)]\nabla(v-u) \right) dt \cdot \nabla w dx \\ &= \int_a^b \left(\frac{\partial \psi}{\partial p}[u + \xi(v-u)]\nabla(v-u) \right) \cdot \nabla w dx,\end{aligned}$$

for all $u, v, w \in \left(H_0^1(a,b)\right)^{2N+1}$ and for some $\xi \in (0,1)$. With Lipschitz-continuity (5.24), strong monotonicity (5.23) and the choice $w = v - u$, we arrive at

$$m\,|w|^2_{(H^1(a,b))^{2N+1}} \le \int_a^b \frac{\partial \psi}{\partial p}[u + \xi w]\nabla w \cdot \nabla w dx \le 3L\,|w|^2_{(H^1(a,b))^{2N+1}}.$$

Hence, the norm equivalence is proven. □

Corollary 5.19. *Let us assume that*

1. $\nu(\cdot) \in C^1(\mathbb{R}^+ \longrightarrow \mathbb{R})$ *and* $0 < \underline{\nu} \le \nu(s) \le \overline{\nu}$ *for all* $s \in \mathbb{R}_0^+$*,*
2. $s \mapsto \nu(s)s$ *is Lipschitz continuous with Lipschitz constant* L*, i.e*

$$|\nu(s)s - \nu(t)t| \le L\,|s-t|,$$

3. $s \mapsto \nu(s)s$ *is strongly monotone with monotonicity constant* m*, i.e*

$$(\nu(s)s - \nu(t)t)(s-t) \ge m\,|s-t|^2.$$

Then, we have the norm equivalence

$$\int_a^b B[u]\nabla w \cdot \nabla w dx \cong |w|^2_{(H^1(a,b))^{2N+1}}, \quad \forall w \in \left(H_0^1(a,b)\right)^{2N+1}.$$

Proof. By the identity $\nu(s) + \nu'(s)s = (\nu(s)s)'$, the strong monotonicity and the Lipschitz-continuity, we arrive

$$m \le \nu(s) + \nu'(s)s \le L, \quad \forall s > 0.$$

Using $0 < \underline{\nu} \le \nu(s) \le \overline{\nu}$, we arrive at

$$\min\{m, \underline{\nu}\} \le \nu(s) + \nu'(s)s \le \max\{L, \overline{\nu}\}, \quad \forall s \ge 0.$$

Hence, using the orthogonality arguments (4.12), each diagonal-entry of $B[u]$, can be estimated by

$$\begin{aligned}\min\{m, \underline{\nu}\} &\le \frac{\partial \psi_i^c}{\partial p_i^c}[u] \le \max\{L, \overline{\nu}\}, \quad \forall i = 0, \dots, N, \\ \min\{m, \underline{\nu}\} &\le \frac{\partial \psi_i^s}{\partial p_i^s}[u] \le \max\{L, \overline{\nu}\}, \quad \forall i = 1, \dots, N.\end{aligned}$$

Therefore, the norm equivalence is proven. □

Lemma 5.20. *Let us assume*

1. $\nu(\cdot) \in C^1(\mathbb{R}^+ \longrightarrow \mathbb{R})$ *and* $0 < \underline{\nu} \leq \nu(s) \leq \overline{\nu}$ *for all* $s \in \mathbb{R}_0^+$,
2. $s \mapsto \nu(s)s$ *is Lipschitz continuous with Lipschitz constant* L*, i.e*
$$|\nu(s)s - \nu(t)t| \leq L\,|s-t|\,,$$
3. $s \mapsto \nu(s)s$ *is strongly monotone with monotonicity constant* m*, i.e*
$$(\nu(s)s - \nu(t)t)(s-t) \geq m\,|s-t|^2\,.$$

Then, we have the spectral equivalence

$$\int_a^b \frac{\partial \psi}{\partial p}[u]\nabla w \cdot \nabla w dx \cong \int_a^b B[u]\nabla w \cdot \nabla w dx, \quad \forall w \in \left(H_0^1(a,b)\right)^{2N+1}.$$

Proof. The statement follows from Corollary 5.18 and Corollary 5.19. □

5.3.3 Discrete Version

As in the linear case, we approximate $H_0^1(a,b)$ by piecewise linear functions and therefore $\left(H_0^1(a,b)\right)^{2N+1}$ by $S_h^N = \left(\text{span } \{\varphi_i\}_{i=1}^{N_h}\right)^{2N+1}$. Next we want to build up the matrices and the resulting system of linear equations, that has to be solved in each step of Newton's iteration. In order to achieve clear arrangements, we have to clarify some notations.

Notations

By $u_h(t)$ we denote the multiharmonic function that is determined by

$$u_h(t) = \sum_{k=0}^{N} u_k^c(x)\cos(k\omega t) + u_k^s(x)\sin(k\omega t)$$

with the the finite element functions

$$u_k^c(x) = \sum_{i=1}^{N_h} u_{k,i}^c \varphi_i(x) \quad \text{and} \quad u_k^s(x) = \sum_{i=1}^{N_h} u_{k,i}^s \varphi_i(x).$$

So therefore the notation u_h is self-explanatory, i.e

$$u_h = (u_0^c, u_1^c, u_1^s, \ldots, u_N^c, u_N^s) \in \left(H_0^1(a,b)\right)^{2N+1}.$$

In fact we will also need the first space-derivative of $u_h(t)$, namely $p_h(t)$. Similar to the original one, it is given by the multiharmonic function

$$p_h(t) = \frac{\partial u}{\partial x} = \sum_{k=0}^{N} p_k^c(x)\cos(k\omega t) + p_k^s(x)\sin(k\omega t)$$

with the finite element functions

$$p_k^c(x) = \sum_{i=1}^{N_h} u_{k,i}^c \varphi_i'(x) \quad \text{and} \quad p_k^s(x) = \sum_{i=1}^{N_h} u_{k,i}^s \varphi_i'(x).$$

Further on we use the vector valued notation

$$\begin{aligned} \underline{u}_h &= (\underline{u}_0^c, \underline{u}_1^c, \underline{u}_1^s, \ldots, \underline{u}_N^c, \underline{u}_N^s)^T \in \mathbb{R}^{(2N+1)N_h} \\ \underline{w}_h &= (\underline{w}_0^c, \underline{w}_1^s, \ldots, \underline{w}_N^c, \underline{w}_N^s)^T \in \mathbb{R}^{(2N+1)N_h} \\ \underline{v}_h &= (\underline{v}_0^c, \underline{v}_1^s, \ldots, \underline{v}_N^c, \underline{v}_N^s)^T \in \mathbb{R}^{(2N+1)N_h} \end{aligned}$$

where for $j = 1, \ldots, N$ each individual entry is given by

$$\underline{u}_j^c = \left(u_{j,1}^c, u_{j,2}^c, \ldots, u_{j,N_h}^c\right) \in \mathbb{R}^{N_h}$$

Finite Element Equation

At each step k of Newton's iteration, we have to perform two major steps. (For simplicity I omit the iteration index k.)

- Evaluating $(r_h, v_h) = \left(f_h - \left(L^{(N)}[u_h] + \omega\alpha D_N u_h\right), v_h\right)$. One of the main tasks is to build up the operator $L^{(N)}$.
- Solving $\left(\left(L^{(N)\prime}[u_h] + \omega\alpha D_N\right) w_h, v_h\right) = (r_h, v_h)$. Similar here the expression $L^{(N)\prime}[u_h]$ has to be evaluated.

In the following proceedings we want to compute the following expressions.

$$\begin{aligned} (M_h)_{ij} &= \alpha\omega\,(D_N \phi_i, \phi_j) \\ (A_h)_j\,[u_h] &= \left(L^{(N)}[u_h], \phi_j\right) \\ (A_h')_{ij}\,[u_h] &= \left(L^{(N)\prime}[u_h]\phi_i, \phi_j\right) \end{aligned}$$

where $\{\phi_i\}_{i=1}^{(2N+1)N_h}$ is the standard nodal basis for S_h^N.
In both cases the evaluation of the the term $\omega\alpha D_N$ by applying the usual nodal basis isn't tricky at all. We refer the resulting matrix as the mass matrix M_h that is given by

$$M_h = \begin{pmatrix} 0 & & & & & & \\ & 0 & D_h & & & & \\ & -D_h & 0 & & & & \\ & & & 0 & 2D_h & & \\ & & & -2D_h & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 & ND_h \\ & & & & & & -ND_h & 0 \end{pmatrix}$$

with the already - in the linear case - appeared modified mass matrix of the Laplace equation in one dimension, namely

$$D_h = \frac{\alpha\omega(b-a)h}{6}\begin{pmatrix} 4 & 1 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4 & 1 \\ 0 & \cdots & 0 & 1 & 4 \end{pmatrix}.$$

Assembling the stiffness matrix $A_h[u_h]$

Now we want to develop the working process of assembling the stiffness matrix $A_h[u_h]$.

$$\left(L^{(N)}[u_h], v_h\right) = \int_a^b \psi_0^c[u_h]\frac{dv_1^c}{dx} + \sum_{j=1}^{N}\left(\psi_j^c[u_h]\frac{dv_j^c}{dx} + \psi_j^s[u_h]\frac{dv_j^s}{dx}\right)dx$$

We fix j and regard for example the cosine part of the sum.

$$\begin{aligned}
\int_a^b \psi_j^c[u_h]\frac{dv_j^c}{dx}dx &= \int_a^b \frac{2}{T}\int_0^T \nu\left(|p_h(t)|\right)p_h(t)\cos(j\omega t)dt\frac{dv_j^c}{dx}dx \\
&= \frac{2}{T}\int_0^T\int_a^b \nu\left(|p_h(t)|\right)\sum_{i=0}^{N}\left(p_i^c(x)\cos(i\omega t) + p_i^s(x)\sin(i\omega t)\right)\frac{dv_j^c}{dx}dx\cos(j\omega t)dt \\
&= \sum_{i=0}^{N}\int_a^b \frac{2}{T}\int_0^T \nu\left(|p_h(t)|\right)\cos(i\omega t)\cos(j\omega t)dt\frac{du_i^c}{dx}\frac{dv_j^c}{dx}dx \\
&+ \sum_{i=1}^{N}\int_a^b \underbrace{\frac{2}{T}\int_0^T \nu\left(|p_h(t)|\right)\sin(i\omega t)\cos(j\omega t)dt}_{=:\theta^{(s,c)}[u_h]}\frac{du_i^s}{dx}\frac{dv_j^c}{dx}dx.
\end{aligned}$$

We want to assemble the stiffness matrix element-wise and therefore investigate a fixed submatrix. First we will regard only one part of the expression above and therefore fix i and j.

$$\begin{aligned}
\int_a^b \theta^{(s,c)}[u_h]\frac{du_i^s}{dx}\frac{dv_j^c}{dx}dx &= \sum_{k=2}^{N_h-1}\int_{T_k}\theta^{(s,c)}[u_h]\frac{du_i^s}{dx}\frac{dv_j^c}{dx}dx \\
&= {}^{(cs)}_{(ji)}\mathrm{G}_h^{(1)}\, u_{i,1}^s v_{j,1}^c + \sum_{k=2}^{N_h-1}\begin{pmatrix} u_{i,k-1}^s \\ u_{i,k}^s \end{pmatrix}^T {}^{(cs)}_{(ji)}\mathrm{G}_h^{(k)}\begin{pmatrix} v_{j,k-1}^c \\ v_{j,k}^c \end{pmatrix} \\
&+ {}^{(cs)}_{(ji)}\mathrm{G}_h^{(N_h)}\, u_{i,N_h}^s v_{j,N_h}^c
\end{aligned}$$

with

$${}^{(sc)}_{(ji)}\mathrm{G}_h^{(k)} = \begin{pmatrix} \int_{T_k}\theta^{(s,c)}[u_h]\varphi_{k-1}'(x)^2dx & \int_{T_k}\theta^{(s,c)}[u_h]\varphi_{k-1}'(x)\varphi_k'(x)dx \\ \int_{T_k}\theta^{(s,c)}[u_h]\varphi_k'(x)\varphi_{k-1}'(x)dx & \int_{T_k}\theta^{(s,c)}[u_h]\varphi_k'(x)^2dx \end{pmatrix}$$

Next we consider

$$\theta^{(s,c)}[u_h|_{T_k}] = \frac{2}{T}\int_0^T \nu\left(|p_h(t)|_{T_k}|\right)\sin(i\omega t)\cos(j\omega t)dt$$

Regarding this expression on a fixed element T_k we first consider

$$p_h(t)|_{T_k} = \frac{1}{h} \sum_{l=0}^{N} \left(u^c_{l,k} - u^c_{l,k-1}\right) \cos(l\omega t) + \left(u^s_{l,k} - u^s_{l,k-1}\right) \sin(l\omega t)$$

Since this expression is constant on an element of the subdivision, also $\theta^{(s,c)}[u_h|_{T_k}]$ is constant. So we arrive at

$$\begin{aligned} {}^{(sc)}_{(ji)}\mathrm{G}_h^{(k)} &= \theta^{(s,c)}[u_h|_{T_k}] \begin{pmatrix} \int_{T_k} \varphi'_{k-1}(x)^2 dx & \int_{T_k} \varphi'_{k-1}(x)\varphi'_k(x) dx \\ \int_{T_k} \varphi'_k(x)\varphi'_{k-1}(x) dx & \int_{T_k} \varphi'_k(x)^2 dx \end{pmatrix} \\ &= \frac{1}{h}\theta^{(s,c)}[u_h|_{T_k}] \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \end{aligned}$$

Now we can assemble these local stiffness matrices. Here I will omit all indices of ${}^{(cc)}_{(ji)}\mathrm{G}_h^{(k)}$ except the right upper (k).

$${}^{(sc)}_{(ji)}\mathrm{G}_h = \begin{pmatrix} G^{(1)}_{22} + G^{(2)}_{11} & G^{(2)}_{12} & 0 & \cdots & \cdots & 0 \\ G^{(2)}_{21} & G^{(2)}_{22} + G^{(3)}_{11} & G^{(3)}_{12} & \ddots & & \vdots \\ 0 & G^{(3)}_{12} & G^{(3)}_{22} + G^{(4)}_{11} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & G^{(N_h-1)}_{12} \\ 0 & \cdots & \cdots & 0 & G^{(N_h-1)}_{12} & G^{(N_h-1)}_{22} + G^{(N_h)}_{11} \end{pmatrix}$$

and similary ${}^{(ss)}_{(ji)}\mathrm{G}_h$ and ${}^{(cs)}_{(ji)}\mathrm{G}_h = {}^{(cc)}_{(ji)}\mathrm{G}_h$. Finally we obtain the full stiffness matrix

$$\bar{A}_h[\underline{u}_h] = \begin{pmatrix} {}^{(cc)}_{(00)}\mathrm{G}_h & {}^{(cc)}_{(01)}\mathrm{G}_h & {}^{(sc)}_{(01)}\mathrm{G}_h & {}^{(cc)}_{(02)}\mathrm{G}_h & {}^{(sc)}_{(02)}\mathrm{G}_h & \cdots & \cdots & {}^{(cc)}_{(0N)}\mathrm{G}_h & {}^{(sc)}_{(0N)}\mathrm{G}_h \\ {}^{(cc)}_{(10)}\mathrm{G}_h & {}^{(cc)}_{(11)}\mathrm{G}_h & {}^{(sc)}_{(11)}\mathrm{G}_h & {}^{(cc)}_{(12)}\mathrm{G}_h & {}^{(sc)}_{(12)}\mathrm{G}_h & \cdots & \cdots & {}^{(cc)}_{(1N)}\mathrm{G}_h & {}^{(sc)}_{(1N)}\mathrm{G}_h \\ {}^{(cs)}_{(10)}\mathrm{G}_h & {}^{(cs)}_{(11)}\mathrm{G}_h & {}^{(ss)}_{(11)}\mathrm{G}_h & {}^{(cs)}_{(12)}\mathrm{G}_h & {}^{(ss)}_{(12)}\mathrm{G}_h & \cdots & \cdots & {}^{(cs)}_{(1N)}\mathrm{G}_h & {}^{(ss)}_{(1N)}\mathrm{G}_h \\ {}^{(cc)}_{(20)}\mathrm{G}_h & {}^{(cc)}_{(21)}\mathrm{G}_h & {}^{(sc)}_{(21)}\mathrm{G}_h & {}^{(cc)}_{(22)}\mathrm{G}_h & {}^{(sc)}_{(22)}\mathrm{G}_h & & & \vdots & \vdots \\ {}^{(cs)}_{(20)}\mathrm{G}_h & {}^{(cs)}_{(21)}\mathrm{G}_h & {}^{(ss)}_{(21)}\mathrm{G}_h & {}^{(cs)}_{(22)}\mathrm{G}_h & {}^{(ss)}_{(22)}\mathrm{G}_h & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & & \ddots & \ddots & \vdots & \vdots \\ {}^{(cc)}_{(N0)}\mathrm{G}_h & {}^{(cc)}_{(N1)}\mathrm{G}_h & {}^{(sc)}_{(N1)}\mathrm{G}_h & \cdots & \cdots & \cdots & \cdots & {}^{(cc)}_{(NN)}\mathrm{G}_h & {}^{(sc)}_{(NN)}\mathrm{G}_h \\ {}^{(cs)}_{(N0)}\mathrm{G}_h & {}^{(cs)}_{(N1)}\mathrm{G}_h & {}^{(ss)}_{(N1)}\mathrm{G}_h & \cdots & \cdots & \cdots & \cdots & {}^{(cs)}_{(NN)}\mathrm{G}_h & {}^{(ss)}_{(NN)}\mathrm{G}_h \end{pmatrix}$$

So in order to compute the defect, we have to evaluate

$$A_h[\underline{u}_h] = \bar{A}_h[\underline{u}_h]\underline{u}_h.$$

Assembling the Derivative $A_h'[u_h]$

Now we want to apply analogous proceedings to assemble the stiffness matrix of the Frechet derivative, which we denote by $A_h'[u_h]$. For a simplified notation we define the non appearing sinus parts to be zero, i.e.

$$\psi_0^s = 0 \quad \text{and} \quad \frac{\partial}{\partial p_0^s} = 0.$$

This allows us to sum over all indices from 0 to N. Consequently evaluating the matrix and the vectors of finite element function, we obtain the following expression.

$$\begin{aligned}\int_a^b \frac{\partial \psi}{\partial p}[u_h]\nabla w_h \cdot \nabla v_h dx = \sum_{i,j=0}^{N} \int_a^b \frac{\partial \psi_j^c}{\partial p_i^c}[u_h]\frac{dw_i^c}{dx}\frac{dv_j^c}{dx}dx + \int_a^b \frac{\partial \psi_j^c}{\partial p_i^s}[u_h]\frac{dw_i^s}{dx}\frac{dv_j^c}{dx}dx \\ + \int_a^b \frac{\partial \psi_j^s}{\partial p_i^c}[u_h]\frac{dw_i^c}{dx}\frac{dv_j^s}{dx}dx + \int_a^b \frac{\partial \psi_j^s}{\partial p_i^s}[u_h]\frac{dw_i^s}{dx}\frac{dv_j^s}{dx}dx\end{aligned}$$

Even at this point, we can already see, that in the end we will obtain a fully populated matrix A_h', which in fact will be filled by tridiagonal matrices.
We want to assemble the stiffness matrix element-wise and therefore investigate a fixed submatrix. First we will regard only one part of the expression above and therefore fix i and j.

$$\begin{aligned}\int_a^b \frac{\partial \psi_j^c}{\partial p_i^c}[u_h]\frac{dw_i^c}{dx}\frac{dv_j^c}{dx}dx &= \sum_{k=2}^{N_h-1} \int_{T_k} \frac{\partial \psi_j^c}{\partial p_i^c}[u_h]\frac{dw_i^c}{dx}\frac{dv_j^c}{dx}dx \\ &= {}^{(cc)}_{(ji)}\mathrm{K}_h^{(1)}\, w_{i,1}^c v_{j,1}^c + \sum_{k=2}^{N_h-1} \begin{pmatrix} w_{i,k-1}^c \\ w_{i,k}^c \end{pmatrix}^T {}^{(cc)}_{(ji)}\mathrm{K}_h^{(k)} \begin{pmatrix} v_{j,k-1}^c \\ v_{j,k}^c \end{pmatrix} \\ &+ {}^{(cc)}_{(ji)}\mathrm{K}_h^{(N_h)}\, w_{i,N_h}^c v_{j,N_h}^c\end{aligned}$$

with

$${}^{(cc)}_{(ji)}\mathrm{K}_h^{(k)} = \begin{pmatrix} \int_{T_k} \frac{\partial \psi_j^c}{\partial p_i^c}[u_h]\varphi_{k-1}'(x)^2 dx & \int_{T_k} \frac{\partial \psi_j^c}{\partial p_i^c}[u_h]\varphi_{k-1}'(x)\varphi_k'(x)dx \\ \int_{T_k} \frac{\partial \psi_j^c}{\partial p_i^c}[u_h]\varphi_k'(x)\varphi_{k-1}'(x)dx & \int_{T_k} \frac{\partial \psi_j^c}{\partial p_i^c}[u_h]\varphi_k'(x)^2 dx \end{pmatrix}$$

Next we consider

$$\frac{\partial \psi_j^c}{\partial p_i^c}[u_h|_{T_k}] = \frac{2}{T}\int_0^T \left(\nu\left(|p_h(t)|_{T_k}|\right) + \nu'\left(|p_h(t)|_{T_k}|\right)|p_h(t)|_{T_k}|\right)\cos(i\omega t)\cos(j\omega t)dt.$$

Regarding this expression on a fixed element T_k we first consider

$$p_h(t)|_{T_k} = \frac{1}{h}\sum_{l=0}^{N}\left(u_{l,k}^c - u_{l,k-1}^c\right)\cos(l\omega t) + \left(u_{l,k}^s - u_{l,k-1}^s\right)\sin(l\omega t).$$

Since this expression is constant on an element of the subdivision, also $\frac{\partial \psi_j^c}{\partial p_i^c}[u_h]$ is constant. So we arrive at

$$\begin{aligned}{}^{(cc)}_{(ji)}\mathrm{K}_h^{(k)} &= \frac{\partial \psi_j^c}{\partial p_i^c}[u_h|_{T_k}] \begin{pmatrix} \int_{T_k} \varphi_{k-1}'(x)^2 dx & \int_{T_k} \varphi_{k-1}'(x)\varphi_k'(x)dx \\ \int_{T_k} \varphi_k'(x)\varphi_{k-1}'(x)dx & \int_{T_k} \varphi_k'(x)^2 dx \end{pmatrix} \\ &= \frac{1}{h}\frac{\partial \psi_j^c}{\partial p_i^c}[u_h|_{T_k}] \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.\end{aligned}$$

Now we can assemble these local stiffness matrices. Here I will omit all indices of ${}^{(cc)}_{(ji)}\mathrm{K}_h^{(k)}$ except the right upper (k).

$$
{}^{(cc)}_{(ji)}\mathrm{K}_h = \begin{pmatrix}
K_{22}^{(1)} + K_{11}^{(2)} & K_{12}^{(2)} & 0 & \cdots & \cdots & 0 \\
K_{21}^{(2)} & K_{22}^{(2)} + K_{11}^{(3)} & K_{12}^{(3)} & \ddots & & \vdots \\
0 & K_{12}^{(3)} & K_{22}^{(3)} + K_{11}^{(4)} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
\vdots & & \ddots & \ddots & \ddots & K_{12}^{(N_h-1)} \\
0 & \cdots & \cdots & 0 & K_{12}^{(N_h-1)} & K_{22}^{(N_h-1)} + K_{11}^{(N_h)}
\end{pmatrix}
$$

and similarly ${}^{(ss)}_{(ji)}\mathrm{K}_h$ and ${}^{(cs)}_{(ji)}\mathrm{K}_h = {}^{(sc)}_{(ji)}\mathrm{K}_h$ Finally we obtain the full stiffness matrix

$$
A'_h[\underline{u}_h] = \begin{pmatrix}
{}^{(cc)}_{(00)}\mathrm{K}_h & {}^{(cc)}_{(01)}\mathrm{K}_h & {}^{(sc)}_{(01)}\mathrm{K}_h & {}^{(cc)}_{(02)}\mathrm{K}_h & {}^{(sc)}_{(02)}\mathrm{K}_h & \cdots & \cdots & {}^{(cc)}_{(0N)}\mathrm{K}_h & {}^{(sc)}_{(0N)}\mathrm{K}_h \\
{}^{(cc)}_{(10)}\mathrm{K}_h & {}^{(cc)}_{(11)}\mathrm{K}_h & {}^{(sc)}_{(11)}\mathrm{K}_h & {}^{(cc)}_{(12)}\mathrm{K}_h & {}^{(sc)}_{(12)}\mathrm{K}_h & \cdots & \cdots & {}^{(cc)}_{(1N)}\mathrm{K}_h & {}^{(sc)}_{(1N)}\mathrm{K}_h \\
{}^{(cs)}_{(10)}\mathrm{K}_h & {}^{(cs)}_{(11)}\mathrm{K}_h & {}^{(ss)}_{(11)}\mathrm{K}_h & {}^{(cs)}_{(12)}\mathrm{K}_h & {}^{(ss)}_{(12)}\mathrm{K}_h & \cdots & \cdots & {}^{(cs)}_{(1N)}\mathrm{K}_h & {}^{(ss)}_{(1N)}\mathrm{K}_h \\
{}^{(cc)}_{(20)}\mathrm{K}_h & {}^{(cc)}_{(21)}\mathrm{K}_h & {}^{(sc)}_{(21)}\mathrm{K}_h & {}^{(cc)}_{(22)}\mathrm{K}_h & {}^{(sc)}_{(22)}\mathrm{K}_h & & & \vdots & \vdots \\
{}^{(cs)}_{(20)}\mathrm{K}_h & {}^{(cs)}_{(21)}\mathrm{K}_h & {}^{(ss)}_{(21)}\mathrm{K}_h & {}^{(cs)}_{(22)}\mathrm{K}_h & {}^{(ss)}_{(22)}\mathrm{K}_h & \ddots & & \vdots & \vdots \\
\vdots & \vdots & \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \vdots & & & \ddots & \ddots & \vdots & \vdots \\
{}^{(cc)}_{(N0)}\mathrm{K}_h & {}^{(cc)}_{(N1)}\mathrm{K}_h & {}^{(sc)}_{(N1)}\mathrm{K}_h & \cdots & \cdots & \cdots & \cdots & {}^{(cc)}_{(NN)}\mathrm{K}_h & {}^{(sc)}_{(NN)}\mathrm{K}_h \\
{}^{(cs)}_{(N0)}\mathrm{K}_h & {}^{(cs)}_{(N1)}\mathrm{K}_h & {}^{(ss)}_{(N1)}\mathrm{K}_h & \cdots & \cdots & \cdots & \cdots & {}^{(cs)}_{(NN)}\mathrm{K}_h & {}^{(ss)}_{(NN)}\mathrm{K}_h
\end{pmatrix}
$$

We have the following symmetry properties

$$
{}^{(sc)}_{(ij)}\mathrm{K}_h = {}^{(cs)}_{(ij)}\mathrm{K}_h \quad \text{and} \quad {}^{(**)}_{(ij)}\mathrm{K}_h = {}^{(**)}_{(ji)}\mathrm{K}_h
$$

where $(**) \in \{(cc), (cs), (sc), (ss)\}$. So we can conclude that $A'_h[\underline{u}_h] = A'_h[\underline{u}_h]^T$ and therefore this stiffness matrix is symmetric. So analogous to the linear case with $N = 1$ we end up with a system of equations, consisting of a symmetric and positive definite part $A_h[\underline{u}_h]$ and a skew symmetric part M_h. So we want to solve the equation:
Find $\underline{w}_h = (\underline{w}_0^c, \underline{w}_1^c, \underline{w}_1^s, \ldots, \underline{w}_N^c, \underline{w}_N^s)^T \in \mathbb{R}^{(2N+1)N_h}$, such that

$$
A_h^{M\prime}[\underline{u}_h]\underline{w}_h = \underline{f}_h
$$

where

$$
A_h^{M\prime}[\underline{u}_h] = A'_h[\underline{u}_h] + M_h.
$$

Since we have a fully coupled system with $(2N+1)N_h$ unknowns, efficient preconditioning is important for the achievement of a solution in acceptable time.

5.3.4 GMRES Preconditioner

We propose the following block diagonal preconditioner $B_h[\underline{u}_h]$ for the symmetric part $A'_h[\underline{u}_h]$.

$$B_h[\underline{u}_h] = \operatorname{diag}\left({}^{(cc)}_{(00)}\mathrm{K}_h,\ {}^{(cc)}_{(11)}\mathrm{K}_h,\ {}^{(ss)}_{(11)}\mathrm{K}_h,\ {}^{(cc)}_{(22)}\mathrm{K}_h,\ {}^{(ss)}_{(22)}\mathrm{K}_h, \dots,\ {}^{(cc)}_{(NN)}\mathrm{K}_h,\ {}^{(ss)}_{(NN)}\mathrm{K}_h \right)$$

It can be shown, that $B_h[\underline{u}_h]$ is equivalent to $A'_h[\underline{u}_h]$. Due to the norm equivalence in (5.22) we can conclude that

$$\int_a^b L^{(N)\prime}[u_h]\nabla w_h \cdot \nabla w_h dx \cong \int_a^b B[u_h]\nabla w_h \cdot \nabla w_h dx \quad \forall w_h \in S_h^1$$

and consequently the following spectral equivalence is satisfied; i.e. there exist constants $\underline{b} > 0$ and $\overline{b} > 0$, such that

$$\underline{b}\left(A'_h[\underline{u}_h]\underline{w}_h, \underline{w}_h\right) \le \left(B_h[\underline{u}_h]\underline{w}_h, \underline{w}_h\right) \le \overline{b}\left(A'_h[\underline{u}_h]\underline{w}_h, \underline{w}_h\right).$$

So summarising $B_h[\underline{u}_h]$ is a good preconditioner for the Jacobian $A'_h[\underline{u}_h]$ and therefore in each step of Newton's iteration, we have to solve the following system of linear equations $B_h^{A'_h[\underline{u}_h]+M_h}\underline{u}_h = \underline{f}_h$ with

$$B_h^{A'_h[\underline{u}_h]+M_h} = \begin{pmatrix} {}^{(cc)}_{(00)}\mathrm{K}_h &&&&&&&& \\ & {}^{(cc)}_{(11)}\mathrm{K}_h & D_h &&&&&& \\ & -D_h & {}^{(ss)}_{(11)}\mathrm{K}_h &&&&&& \\ &&& {}^{(cc)}_{(22)}\mathrm{K}_h & 2D_h &&&& \\ &&& -2D_h & {}^{(ss)}_{(22)}\mathrm{K}_h &&&& \\ &&&&& \ddots &&& \\ &&&&&& \ddots && \\ &&&&&&& {}^{(cc)}_{(NN)}\mathrm{K}_h & ND_h \\ &&&&&&& -ND_h & {}^{(ss)}_{(NN)}\mathrm{K}_h \end{pmatrix}.$$

Appart from the first block ${}^{(cc)}_{(00)}\mathrm{K}_h$, this matrix consists of blocks of the form equal to the linear case. So for $i = 1, \dots, N$ the uncoupled matrices

$$\left(B_h^{A'_h[\underline{u}_h]+M_h}\right)^{(k)} = \begin{pmatrix} {}^{(cc)}_{(kk)}\mathrm{K}_h & kD_h \\ -kD_h & {}^{(ss)}_{(kk)}\mathrm{K}_h \end{pmatrix} \tag{5.25}$$

can be solved by applying the same Cai-Xu preconditioning technique equal to the linear case.

In the next chapter, an alternative approach of constructing the block diagonal preconditioner $B_h[\underline{u}_h]$ is proposed. It is shown, that the uncoupled diagonal structure can already be attained by choosing a proper integration rule for the evaluation of the Jacobian.

Chapter 6

Numerical Integration

In every step of Newton's iteration, we have to set up the matrices $A_h^M[\underline{u}_h]$ and $A_h^{M\prime}[\underline{u}_h]$. Evaluating these expressions enforces the performance of integration in time. Typically this integration is approximated by some numerical integration, e.g. some Gauss method.
In fact we have two miscellaneous situations: The matrix $A_h^M[\underline{u}_h]$ is involved in computing the defect, and therefore has only to be assembled. The Jacobian $A_h^{M\prime}[\underline{u}_h]$ is part of the left-hand side of a system of linear equations, therefore an easily solvable structure is obsolete. It turns out, that in the first case a more accurate quadrature rule is necessary, while in the second case a rough approximation, e.g. the Gaussian midpoint rule, can be used.

6.1 Computing the Stiffness Matrix

In each step of the Newton iteration we have to evaluate these integrals

$$\theta^{(c,c)}[\underline{u}_h] = \frac{2}{T}\int_0^T \Theta[\underline{u}_h](t)\cos(l\omega t)\cos(k\omega t)dt$$
$$\theta^{(s,c)}[\underline{u}_h] = \frac{2}{T}\int_0^T \Theta[\underline{u}_h](t)\sin(l\omega t)\cos(k\omega t)dt$$
$$\theta^{(c,s)}[\underline{u}_h] = \frac{2}{T}\int_0^T \Theta[\underline{u}_h](t)\cos(l\omega t)\sin(k\omega t)dt$$
$$\theta^{(s,s)}[\underline{u}_h] = \frac{2}{T}\int_0^T \Theta[\underline{u}_h](t)\sin(l\omega t)\sin(k\omega t)dt$$

with

$$\Theta[\underline{u}_h](t) = \nu\left(|p(t)|\right).$$

Approximating these expressions yields a systematic error in evaluation of the residual. Therefore an adequate quadrature rule is compulsory. Our strategy is to use some adaptive numerical integration in order to reach a high accuracy. We use an adaptive Simpson quadrature rule.

6.2 Computing the Derivative

Preconditioning this system can be done via choosing a proper integration rule in time integration. The fully coupled system is due to the fact, that the nonlinearity depends on

the function u. By choosing integration rules, some variables can be uncoupled.
In each step of the Newton iteration we have to evaluate these integrals in order to compute the Jacobian.

$$\begin{aligned}
\frac{\partial \psi_k^c}{\partial p_l^c}[\underline{u}_h] &= \frac{2}{T}\int_0^T \Psi[\underline{u}_h](t)\cos(l\omega t)\cos(k\omega t)dt \\
\frac{\partial \psi_k^c}{\partial p_l^s}[\underline{u}_h] &= \frac{2}{T}\int_0^T \Psi[\underline{u}_h](t)\sin(l\omega t)\cos(k\omega t)dt \\
\frac{\partial \psi_k^s}{\partial p_l^c}[\underline{u}_h] &= \frac{2}{T}\int_0^T \Psi[\underline{u}_h](t)\cos(l\omega t)\sin(k\omega t)dt \\
\frac{\partial \psi_k^s}{\partial p_l^s}[\underline{u}_h] &= \frac{2}{T}\int_0^T \Psi[\underline{u}_h](t)\sin(l\omega t)\sin(k\omega t)dt
\end{aligned}$$

with

$$\Psi[\underline{u}_h](t) = \nu\left(|p(t)|\right) + \nu'\left(|p(t)|\right)|p(t)|\,.$$

We want to evaluate these integrals numerically, bearing in mind, that an easy solvable structure is attained. Therefore we apply Gaussian midpoint rule to the first part of the integral. The main advantage of this choice is the resulting block structure of the obtained matrix.

$$\begin{aligned}
\frac{\partial \psi_k^c}{\partial p_l^c}[\underline{u}_h] &\approx \Psi\left(\frac{T}{2}\right)\frac{2}{T}\int_0^T \cos(l\omega t)\cos(k\omega t)dt = \Psi\left(\frac{T}{2}\right)\delta_{lk} \\
\frac{\partial \psi_k^c}{\partial p_l^s}[\underline{u}_h] &\approx \Psi\left(\frac{T}{2}\right)\frac{2}{T}\int_0^T \sin(l\omega t)\cos(k\omega t)dt = 0 \\
\frac{\partial \psi_k^s}{\partial p_l^c}[\underline{u}_h] &\approx \Psi\left(\frac{T}{2}\right)\frac{2}{T}\int_0^T \cos(l\omega t)\sin(k\omega t)dt = 0 \\
\frac{\partial \psi_k^s}{\partial p_l^s}[\underline{u}_h] &\approx \Psi\left(\frac{T}{2}\right)\frac{2}{T}\int_0^T \sin(l\omega t)\sin(k\omega t)dt = \Psi\left(\frac{T}{2}\right)\delta_{lk}
\end{aligned}$$

In order to obtain the grade of quality of the chosen rule we give an error estimate. By Taylor expansion

$$\Psi(t) = \Psi\left(\frac{T}{2}\right) + \Psi'\left(\frac{T}{2}\right)\left(\frac{T}{2}-t\right) + O\left(\left(\frac{T}{2}-t\right)^2\right),$$

we obtain

$$\begin{aligned}
\frac{2}{T}\Psi'\left(\frac{T}{2}\right)\int_0^T\left(\frac{T}{2}-t\right)\cos(l\omega t)\cos(k\omega t)dt = 0 \\
\frac{2}{T}\Psi'\left(\frac{T}{2}\right)\int_0^T\left(\frac{T}{2}-t\right)\sin(l\omega t)\sin(k\omega t)dt = 0
\end{aligned}$$

and

$$\frac{2}{T}\Psi'\left(\frac{T}{2}\right)\int_0^T\left(\frac{T}{2}-t\right)\sin(l\omega t)\cos(k\omega t)dt = T\Psi'\left(\frac{T}{2}\right)\begin{cases}\frac{l}{\pi(l^2-k^2)} & k\neq l \\ \frac{1}{4\pi k} & k = l\end{cases}.$$

In fact this approximation yields an inexact Newton method, since we use an approximation of the Jacobian instead of the exact one. So for the whole iteration procedure, we basically can expect q-linear convergence.

Chapter 7

Numerical Results

In this chapter we present the numerical results of the developed solver. As in the theoretical proceedings we start with the calculations of the linear problem. Especially we consider the two special cases of Section 5.2 to verify the theoretical convergence rates.
We continue with some results of the nonlinear problem and apply an inexact damped Newton iteration. Precisely high and low order approximations are discussed and various numbers of harmonics in the multiharmonic ansatz are compared.

7.1 Linear Problem

Since we have

$$\frac{\|r_i\|_{A_h}}{\|r_0\|_{A_h}} \leq \frac{\|r_i\|_{l_2}}{\|r_0\|_{l_2}} \sqrt{\kappa(A_h)}$$

we expect the same behaviour, if we measure the residual in the usual l_2 norm. Here $\kappa(A_h)$ denotes the condition number of the matrix A_h.
We set up the following test conditions:

- Full GMRES without restarting.
- Initial guess all zero.
- Stopping criteria $\frac{\|r_i\|_{l_2}}{\|r_0\|_{l_2}} \leq 10^{-5}$.
- Cai-Xu preconditioner with the exact inverse of A_h as a preconditioner for the symmetric part, i.e

  $$B^M = I_H^h \left(A_H^M\right)^{-1} I_h^H + \beta A_h^{-1}$$

  In practice we don't use A_h^{-1}, but any other optimal preconditioner for the symmetric part A_h.
- Exakt solver in preconditioning.

For simplicity we choose $\nu(x) = 1$ and the frequency $\omega = 1$. Furthermore the parameter α is used for simulating high frequencies. In order to compute the excitation, we prescribe the solution $u(x,t)$ in such a way, that the initial condition $u(x,0)$ is fulfilled. (Note, that this isn't really necessary, since we are looking for a steady state solution.) The nonsymmetry of the resulting finite element equations is manageable by the choice of α. So the problem reads as

$$\begin{aligned} \alpha\frac{\partial u}{\partial t}(x,t) - \left(\frac{\partial^2 u}{\partial x^2}\right)(x,t) &= f(x,t) && (x,t) \in (0,1)\times(0,2\pi) \\ u(x,0) &= 0 && x \in (0,1) \\ u(0,t) = u(1,t) &= 0 && t \in (0,2\pi). \end{aligned} \tag{7.1}$$

We choose $f(x,t)$ in such a way, that

$$u(x,t) = x(1-x)e^x \sin(t)$$

and therefore we obtain the harmonic excitation

$$f(x,t) = \alpha(x-x^2)e^x \cos(t) + (3x+x^2)e^x \sin(t).$$

In Table 7.1, Table 7.2 and Table 7.3 the number of GMRES iteration for various parameter settings are prepared.
The Cai-Xu preconditioner is quite robust for the choice of the balancing parameter β.

$\alpha = 1$											
β		1	2	3	4	5	6	7	8	9	10
$H = 1/5$	$h = 1/60$	5	4	4	4	4	4	5	5	5	5
$H = 1/5$	$h = 1/120$	4	4	4	5	5	5	5	5	5	5

Table 7.1: GMRES iterations

$\alpha = 100$											
β		1	2	3	4	5	6	7	8	9	10
$H = 1/5$	$h = 1/60$	11	12	12	12	12	12	13	13	13	13
$H = 1/10$	$h = 1/60$	11	12	12	12	12	12	12	13	13	13
$H = 1/5$	$h = 1/120$	12	12	12	13	13	13	13	13	13	13
$1/\beta$		1	2	3	4	5	6	7	8	9	10
$H = 1/5$	$h = 1/60$	11	11	11	11	11	11	11	11	11	11
$H = 1/10$	$h = 1/60$	11	11	11	11	11	11	11	11	11	11
$H = 1/5$	$h = 1/120$	12	11	11	11	11	11	11	11	11	11

Table 7.2: GMRES iterations

Special Case $\beta = \infty$

From our spectral analysis we can conclude, that we will asymptotically obtain the fewest number of GMRES iterations for the case $\beta = \infty$. The resulting GMRES iterations are in Table 7.4. For $\beta > 1$ this behaviour is verified by Table 7.1, Table 7.2 and Table 7.3. For $\beta < 1$ the lower bound isn't sharp anymore. This is due to the fact, that for β tends to 0, the solution of the non symmetric part is appreciated by the coarse space solver.

$\alpha = 1000$											
β		1	2	3	4	5	6	7	8	9	10
$H = 1/5$	$h = 1/60$	19	20	20	21	21	22	22	22	22	22
$H = 1/10$	$h = 1/60$	18	20	20	21	21	21	22	22	22	22
$H = 1/5$	$h = 1/120$	20	21	22	22	22	22	23	23	23	23
$1/\beta$		1	2	3	4	5	6	7	8	9	10
$H = 1/5$	$h = 1/60$	19	18	17	16	16	16	15	15	15	14
$H = 1/10$	$h = 1/60$	18	18	17	16	16	16	15	15	15	14
$H = 1/5$	$h = 1/120$	20	19	18	18	17	17	16	16	16	16

Table 7.3: GMRES iterations

$\beta = \infty$						
α	1	5	10	50	100	1000
$h = 1/60$	4	6	7	10	12	19
$h = 1/120$	4	6	7	10	12	20

Table 7.4: GMRES iterations

7.2 Nonlinear Problem

Since in practice we cannot guarantee that the initial guess is sufficiently close enough to the solution for the iteration to converge, we choose a damped Newton method to obtain more global convergence properties. This modification differs in the updating procedure given in (5.20), i.e.

$$x_{k+1} = x_k + \tau_k \left(F'(x_k)\right)^{-1} (y - F(x_k))$$

with $\tau_k \in (0, 1]$. For $\tau_k = 1$ one obviously obtains Newton's method. The damping parameter τ_k is determined by line search.

We set up the following test conditions: We apply a damped Newton iteration, where in each step the damping parameter is calculated by line search. As an initial guess for the Newton iteration, we use all zero.

In each step of Newton's iteration we apply analogous proceedings as in the linear part for computing the inverse of the block diagonal Jacobian. More precisely, for $k = 1, \ldots, N$ each skewsymmetric 2×2 block given by (5.25)

$$\left(B_h^{A'_h[\underline{u}_h]+M_h}\right)^{(k)} = \begin{pmatrix} {}^{(cc)}_{(kk)}\mathrm{K}_h & kD_h \\ -kD_h & {}^{(ss)}_{(kk)}\mathrm{K}_h \end{pmatrix}$$

is solved by the Cai-Xu GMRES preconditioning technique with the following settings.

- Full GMRES without restarting.
- Initial guess all zero.
- Stopping criteria $\frac{\|r_i\|_{l_2}}{\|r_0\|_{l_2}} \leq 10^{-5}$.

- Cai-Xu preconditioner with the exact inverse as a preconditioner for the symmetric part.
- We set the balancing paramter β equal to one. Investigating the linear case we observed, that the choice of the balancing parameter β isn't really important to the convergence of the preconditioned GMRES method. Therefore this choice seems to be valid.
- The fine mesh is choosen to be of size $h = 1/60$ and the coarse mesh of size $H = 1/10$.

For $k = 0$ the symmetric and positive definite tridiagonal matrix ${}^{(cc)}_{(00)}\mathrm{K}_h$ is solved by an appropriate direct solver.
For test issues we choose the nonlinear potential to be

$$\nu(s) = 1 + s^2.$$

So in fact we are interested of a steady state solution of the problem

$$\begin{aligned} \alpha\frac{\partial u}{\partial t}(x,t) - \frac{\partial}{\partial x}\left(\nu\left(\left|\frac{\partial u}{\partial x}\right|\right)\frac{\partial u}{\partial x}\right)(x,t) &= f(x,t) \qquad (x,t) \in (0,1)\times(0,T) \\ u(0,t) = u(1,t) &= 0 \qquad\qquad t \in (0,T). \end{aligned} \tag{7.2}$$

Our first approach is to use a pure consine harmonic excitation

$$f(x,t) = cos(\omega t).$$

As mentioned in Chapter 4, we only expect odd harmonics to be involved in the solution.
Now we apply inexact Newton's method with the following presettings.

- As an initial guess we choose all zero.
- The Newton iteration is stopped, if the residual is less than 10^{-5}.
- In case that the initial guess isn't sufficiently close to the exact solution, we apply a line search algorithm by halving τ_k in each step and comparing the resulting residuals.

Table 7.5, Table 7.6 and Table 7.7 list the corresponding Newton iterations, refered as outer iteration, and GMRES iteration, refered as inner iteration, for various choices of ω. Note, that in each step of Newton's iteration, we have to perform $2N$ times a GMRES iteration, where each of them requires the amount of an inner iteration, as printed in the table.
In Table 7.5 we additionally add the total number of unknowns.

$\omega = 1$						
N	1	2	3	4	5	6
outer iteration	7	7	7	7	7	7
inner iteration	5	5	6	6	6	6
unknowns	177	295	413	531	649	767

Table 7.5: Newton and GMRES iterations

$\omega = 10$						
N	1	2	3	4	5	6
outer iteration	5	5	6	6	6	6
inner iteration per step	7	7	8	8	9	9

Table 7.6: Newton and GMRES iterations

$\omega = 50$						
N	1	2	3	4	5	6
outer iteration	3	3	3	3	3	3
inner iteration per step	10	10	10	10	10	10

Table 7.7: Newton and GMRES iterations

Convergence of the Newton iteration

Now we want to investigate the convergence behaviour of Newton's iteration. Since we apply an inexact version, we can at least expect q-linear convergence. Figure 7.1 shows the decreasing of the relative residual measured in l^2 norm, namely

$$\frac{\left\|\underline{f} - A_h^M[\underline{u}^k]\right\|_{l^2}}{\left\|\underline{f}\right\|_{l^2}}$$

for various numbers of involved harmonics. The plots verifies the theoretical convergence.

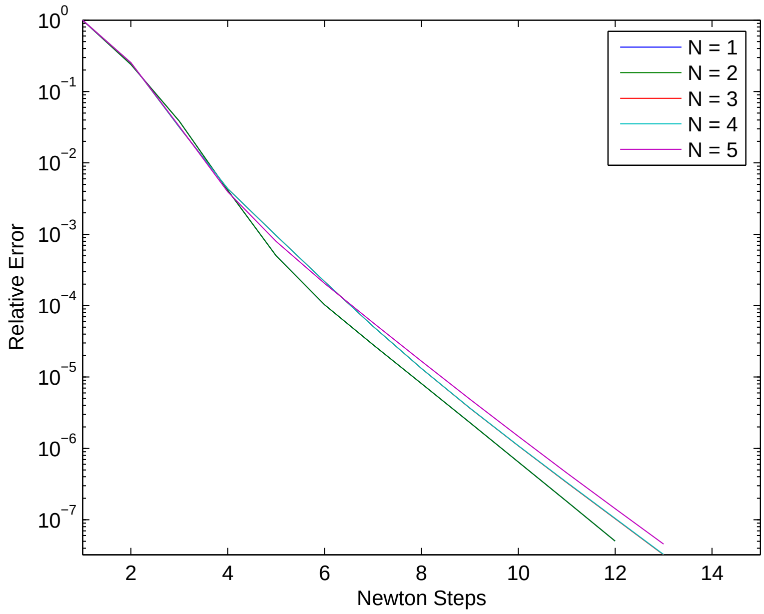

Figure 7.1: The relative residual

Investigation of involved harmonics

We use pure cosine excitation $f(x) = \cos(t)$ again and consider the nonzero Fourier coefficients of the steady state solution. In Figure 7.2 the essential Fourier coefficients are visualised. We can conclude, that it is sufficient to take three harmonics into account. Here we can also observe, that due to the pure cosine excitation, only odd harmonics are involved.

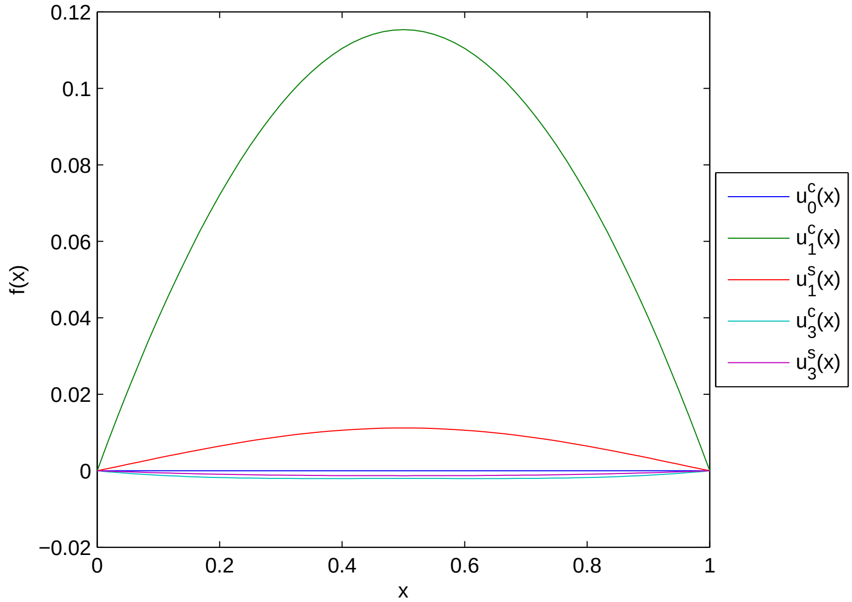

Figure 7.2: Relevant harmonics

In comparision we perform the same proceeding for a combination of sine and cosine excitation $f(x) = \cos(t) + \sin(t)$. The resulting Fourier coefficients can be seen in Figure 7.3.

Comparing the solution obtained from different numbers of involved harmonics

Figure 7.4 shows various solutions of our model equation with pure cosine excitation. Therefore different numbers of involved harmonics are used and the resulting solutions are shown for a fixed space-coordinate $x = 0.5$.

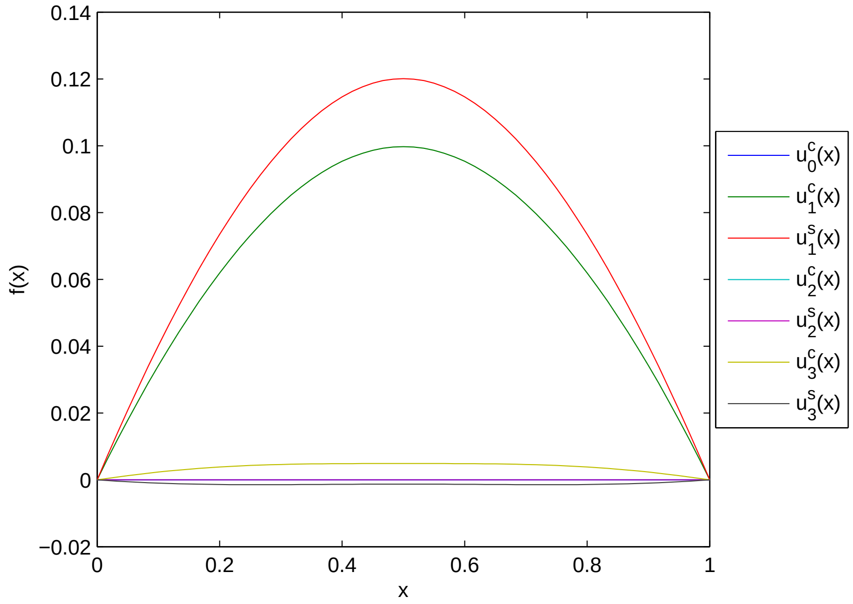

Figure 7.3: Relevant harmonics

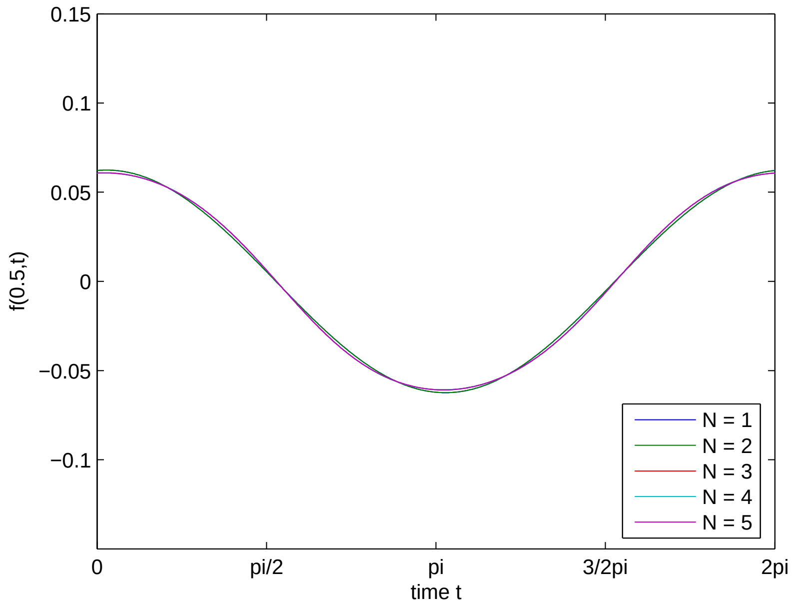

Figure 7.4: Different numbers of harmonics at $x = 0.5$.

Chapter 8

Conclusion and Outlook

8.1 Conclusion

In this thesis, the mathematical treatment of one-dimensional scalar potential, parabolic partial differential equation with time-harmonic excitation by the usage of a multiharmonic approach is embraced.
As preliminary considerations an abstract analysis of the Cai-Xu GMRES preconditioning technique has been cited from Cai/Xu [4].
After considering the relevance of a periodic steady state solution, we have applied a harmonic ansatz to the linear problem and a multiharmonic ansatz in terms of a truncated Fourier series to the nonlinear one.
In the linear case, we have discretised the resulting system of elliptic partial differential equations by means of the finite element method. In order to solve the system of linear equations, we have used the Cai-Xu preconditioner for GMRES and performed spectral analysis in order to verify convergence rates of the corresponding GMRES method.
Considering the nonlinear problem, the system of partial differential equations has been linearised by a Newton-type method and discretised by the finite element method as before. Performing an uncoupling of the variables in the Jacobian system by choosing a special type of Gaussian integration has yielded the same substructure to justify the application of the Cai-Xu preconditioner as in the linear case.
Finally we have discussed some numerical results conforming our theoretical investigations.

8.2 Outlook

This thesis provides a starting point of preparing theoretical results for a multiharmonic solver for general parabolic partial differential equations in 1D. So as a matter of course, there is still a load of open questions. Just mentioning a couple of them, the work can be continued in the following directions:

1. *Robust preconditioner in respect of higher frequencies ωk.*
   The Cai-Xu preconditioner is not robust in respect of high frequencies $k\omega$. For large frequencies, i.e. $k\omega$ tends to infinity, the convercence rate tends to one and therefore stagnates. In fact we want to precondition the resulting system of linear equations in

such a way, that the convergence rate is asymptotically bounded by a positive number less than one.

2. *Number of involved harmonics N.*
Due to the nonlinearity of the potential ν, in the multiharmonic solution u also higher frequencies of the basic excitation of the right-hand side are involved. Due to the fact, that each additionally involved harmonic dramatically increasis the number of unknowns, a low number is favoured. Numerical experiments exculpate the hope, that the truncation error behaves like $O\left(e^{-N}\right)$. The next is to verify this behaviour by developing some Fourier analysis.

3. *Choice of quadrature rule for integration in time.*
In order to achieve a diagonal block structure and therefore to uncouple the system of linear equations, the evaluation of the Jacobian $A_h^{M\prime}$ is done by a very rough approximation of the integration in time. In order to improve the error bound, one can think of more accurate quadrature rule, that in addition conserve an easy solveable structure of the Jacobian as well.

4. *Comparison to time-stepping methods.*
The obtained results can be compared to time-stepping methods, e.g Runge Kutta, in order to obtain comparable results concerning the complexity of the individual methods.

5. *Analysis and Simulation of 3D Problems.*
This thesis provides basic analysis of the scalar potential parabolic problems in 1D. However, the superordinate target is to construct an efficient multiharmonic solver for more practical problems in 3D. Therefore a continuative step is to provide theory and analyis for equations of the kind

$$
\begin{aligned}
\alpha\frac{\partial u}{\partial t} - \nabla\cdot\left(\nu\left(|\nabla u|\right)\nabla u\right) = f \qquad & \text{in } \Omega\times(0,T)\\
u(x,0) = u_0(x) \qquad & \text{in } \Omega\\
u(x,t) = 0 \qquad & \text{on } \partial\Omega\times(0,T)
\end{aligned}
$$

where $\Omega\subset\mathbb{R}^3$ is a bounded Lipschitz domain. Analogous to the 1D problem, we assume the right-hand side $f(x,t)$ to be given by a harmonic excitation.

6. *Generalization to the eddy current problem.*

We want to close by mentioning, that the work on this topic is planed to be continued in terms of exploring these declared and further open tasks.

Bibliography

[1] F. Bachinger. Multigrid solvers for 3D multiharmonic nonlinear magnetic field computations. Master's thesis, Institute for Computational Mathematics, Johannes Kepler University, Linz, Oktober 2003.

[2] F. Bachinger, U. Langer, and J. Schöberl. Numerical analysis of nonlinear multiharmonic eddy current problems. *Numer. Math.*, 100:593–616, 2005.

[3] D. Braess. *Finite Elemente - Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer Verlag, Berlin Heidelberg, 4 edition, 2007.

[4] X.-C. Cai and J. Xu. A preconditioned GMRES method for nonsymmetric or indefinite problems. *Mathematics of Computation*, 59(200):311–319, Oct. 1992.

[5] D. M. Copeland and U. Langer. Domain decomposition solvers for nonlinear multiharmonic finite element equations. (private communication).

[6] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis*, 20(2):345 – 357, 1983.

[7] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 2002.

[8] H. D. Gersem, H. V. Sande, and K. Hameyer. Strong coupled multi-harmonic finite element simulation package. *COMPEL*, 20:535 – 546, 2001.

[9] J. Gyselinck, P. Dularand, C. Geuzainand, and W. Legros. Harmonic-balance finite-element modeling of electromagnetic devices: a novel approach. *IEEE Transactions on Magnetics*, 38:521 – 524, 2002.

[10] M. Jung and U. Langer. *Methode der finiten Elemente für Ingenieure: Eine Einführung in die numerischen Grundlagen der Computersimulation*. Teubner Verlag, Stuttgart, June 2001.

[11] O. Ladyschenskaja. *Aufgaben der Mathematischen Physik*. Nauka, Moskau, 1973.

[12] G. Paoli, O. Biro, and G. Buchgraber. Complex representation in nonlinear time harmonic eddy current problems. *IEEE Transactions on Magnetics*, 34:2625–2628, 1998.

[13] C. Pechstein. Multigrid-newton-methods for nonlinear magnetostatic problems. Master's thesis, Johannes Kepler University, Linz, 2004.

[14] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856 – 869, 1986.

[15] O. Steinbach. *Numerische Näherungsverfahren für elliptische Randwertprobleme - Finite Elemente und Randelemente*. Teubner Verlag, Stuttgart - Leipzig - Wiesbaden, 2004.

[16] O. Steinbach. *Lösungsverfahren für lineare Gleichungssysteme - Algorithmen und Anwendungen*. Teubner Verlag, Wiesbaden, 2005.

[17] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems - Finite and Boundary Elements*. Springer Verlag, 2008.

[18] A. Toselli and O. Widlund. *Domain Decomposition Methods - Algorithm and Theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer Verlag, Berlin, 2005.

[19] S. Yamada and K. Bessho. Harmonic field calculation by the combination of finite elementanalysis and harmonic balance method. *IEEE Transactions on Magnetics*, 24:2588–2590, 1988.

[20] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/A: Linear Monotone Operators*. Springer, New York, 1990.

[21] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/B: Nonlinear Monotone Operators*. Springer, New York, 1990.

[22] Y. Zhou. Fourier analysis and local Fourier analysis for multigrid methods. Master's thesis, Institute for Computational Mathematics, Johannes Kepler University, Linz, 2009.

[23] W. Zulehner. *Numerische Mathematik: Eine Einführung anhand von Differentialgleichungsproblemen; Band 1: Stationäre Probleme*. Mathematik Kompakt. Birkhäuser Verlag, 2008.

Eidesstattliche Erklärung

Ich, Michael Kolmbauer, erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Kremsmünster, date

Signatur