



TNF

Technisch-Naturwissenschaftliche
Fakultät

Multiharmonic Finite Element Analysis of Parabolic Time-Periodic Simulation and Optimal Control Problems

DISSERTATION

zur Erlangung des akademischen Grades

Doktorin

im Doktoratsstudium der

Technischen Wissenschaften

Eingereicht von:

Dipl.-Ing. Monika Wolfmayr Bakk.techn.

Angefertigt am:

Institut für Numerische Mathematik

Beurteilung:

O. Univ. Prof. Dipl.-Ing. Dr. Ulrich Langer (Betreuung)

Prof. Dr. Alfio Borzi

Mitwirkung:

Priv. Doz. Dr. Johannes Kraus

Linz, Februar 2014

For Sebastian and my parents

Abstract

This thesis is devoted to the construction, analysis and implementation of efficient and robust numerical methods for linear parabolic time-periodic simulation and optimal control problems. The discretization of these problems is based on the multiharmonic finite element method whereas new algebraic multilevel preconditioned minimal residual methods are developed for solving the discrete problems, which have saddle point structure.

The mathematical and numerical analysis include existence and uniqueness results in a new variational framework and full a priori and a posteriori error estimates in space and time. Since we consider time-periodic problems, the multiharmonic finite element method is a very natural approach to discretize this type of parabolic problems. More precisely, we expand all – given and unknown – functions into Fourier series in time, truncate them, and then approximate the Fourier coefficients by the finite element method. This method reduces a large linear time-dependent problem to a sequence of smaller time-independent ones.

The multiharmonic finite element discretization of linear parabolic time-periodic simulation and optimal control problems leads to large systems of symmetric but indefinite linear algebraic equations, which fortunately decouple into smaller linear systems each of them defining the cosine and sine Fourier coefficients with respect to a single frequency. The resulting smaller systems have saddle point structure and can be solved by the preconditioned minimal residual method totally in parallel. Hence, we construct block-diagonal preconditioners resulting in fast converging minimal residual solvers with parameter-independent convergence rates. The diagonal blocks of these preconditioners are sums of stiffness and mass matrices, which can be seen as finite element discretization of reaction-diffusion type problems with heterogeneous reaction and diffusion coefficients.

Moreover, we present efficient preconditioners for reaction-diffusion type problems that are optimal in terms of the computational complexity and robust with respect to the reaction and diffusion coefficients. The considered preconditioners belong to the class of so-called algebraic multilevel iteration methods, which are based on multilevel block factorization and polynomial stabilization. One of the main achievements of this thesis is not only the construction of preconditioners via the algebraic multilevel iteration method but also the presentation of a rigorous proof of the robustness and optimal complexity of these preconditioners. This analysis benefits from the use of symbolic techniques.

Although the main focus of this thesis is the numerical analysis of linear parabolic time-periodic simulation and optimal control problems, we finally implement the algorithms in C++, perform many numerical experiments and discuss numerical results which impressively confirm our theoretical findings.

Zusammenfassung

Diese Dissertation beschäftigt sich mit der Entwicklung, Analyse sowie Implementierung von effizienten und robusten numerischen Verfahren zur Lösung von linearen parabolischen zeitperiodischen Problemen. Dies beinhaltet sowohl die Simulation als auch die Steuerung dieser Probleme. Wir verwenden die multiharmonische Finite Elemente Methode zur Diskretisierung der Probleme und entwickeln neue algebraische Multilevel-Vorkonditionierer, um wiederum die diskreten Probleme, die Sattelpunktstruktur haben, mittels des vorkonditionierten MINRES-Verfahrens (vom engl. minimal residual) zu lösen.

Die detaillierte mathematische und numerische Analyse umfasst sowohl Existenz- und Eindeutigkeitsresultate sowie a priori und a posteriori Fehlerabschätzungen in Raum und Zeit. Wir definieren hier neue Funktionenräume im Fourierraum und schaffen somit eine neue Umgebung für entsprechende Variationsformulierungen. Man kann die multiharmonische Finite Elemente Methode durchaus als einen natürlichen Zugang zur Diskretisierung von zeitperiodischen Probleme bezeichnen, weil hier alle – sowohl gegebene als auch unbekannte – Funktionen in Fourierreihen entwickelt werden und somit die Periodizität auf natürliche Weise erfüllt wird. Weiters werden die Fourierreihen abgebrochen und die Fourierkoeffizienten mittels der Finiten Elemente Methode approximiert. Somit reduziert die multiharmonische Finite Elemente Methode ein großes lineares zeitabhängiges Problem auf eine Reihe von kleineren zeitunabhängigen Problemen.

Die multiharmonische Finite Elemente Diskretisierung von linearen parabolischen zeitperiodischen Problemen führt zu großen Systemen linearer algebraischer Gleichungen mit symmetrischen, aber indefiniten Systemmatrizen, die glücklicherweise in kleinere lineare Systeme in den Fourierkoeffizienten zerfallen und alle parallel gelöst werden können. Diese kleineren Systeme haben ebenfalls eine Sattelpunktstruktur und können somit mittels des vorkonditionierten MINRES-Verfahrens gelöst werden. Wir behandeln die Konstruktion von block-diagonalen Vorkonditionierern, welche uns ein schnell konvergierendes MINRES-Verfahren mit einer von den Problemparametern unabhängigen Konvergenzrate liefern. Die Diagonalblöcke dieser Vorkonditionierer bestehen aus Summen von Steifigkeits- und Massenmatrizen, welche man als Finite Elemente Diskretisierung von Reaktions-Diffusions-Gleichungen mit inhomogenen Reaktions- und Diffusionskoeffizienten betrachten kann.

Weiters werden effiziente Vorkonditionierer für Reaktions-Diffusions-Probleme präsentiert, die optimal in der Berechnungskomplexität und robust in Bezug auf die Reaktions- und Diffusionskoeffizienten sind. Die betrachteten Vorkonditionierer gehören der Familie der sogenannten algebraischen Multilevel-Iterationsverfahren an, welche auf einer Multilevel-Blockfaktorisierung und Stabilisierung mittels Polynome basieren. Ein wichtiger theoretischer Beitrag dieser Dissertation ist, dass neben der Konstruktion dieser algebraischen Multilevel-Vorkonditionierer ein detaillierter Beweis für deren Optimalität in der Berechnungskomplexität und Robustheit in Bezug auf die Reaktions- und Diffusionskoeffizienten präsentiert wird, der unter anderem auch durch das Verwenden von symbolischen Methoden erreicht wird.

Obwohl das Hauptaugenmerk dieser Dissertation auf der numerischen Analyse von linearen parabolischen zeitperiodischen Problemen liegt, werden schlussendlich die theoretisch behandelten Algorithmen in C++ implementiert, verschiedene numerische Experimente durchgeführt und numerische Resultate diskutiert, welche eindrucksvoll die theoretischen Ergebnisse dieser Dissertation untermauern.

Acknowledgments

First of all, I would like to thank my supervisor Prof. Ulrich Langer for all his support and encouragement throughout the work on this thesis, especially, for introducing me to lots of different fields in numerics and optimization as well as for giving me the latitude to follow my own scientific interests within the scope of this project. I am very thankful for his time and scientific guidance, which he willingly gave me during the last couple of years, including lots of interesting and fruitful discussions. I also owe my thanks to Prof. Alfio Borzì for his interest in my work and for co-refereeing this thesis. Moreover, I especially thank my PhD co-supervisor Priv. Doz. Dr. Johannes Kraus for spending so much time on introducing me to algebraic multilevel iteration methods, for his help in the implementation of the theoretical algorithms as well as for his encouragement, guidance and all the fruitful and enlightening discussions.

I am very thankful to Prof. Walter Zulehner for introducing me to robust preconditioning techniques for saddle point systems and for his scientific influence and support ever since my graduate studies. Moreover, I thank Prof. Peter Paule for his support and for creating such a friendly and at the same time productive environment within the Doctoral Program “Computational Mathematics”. At this point, I express my thanks to Dr. Veronika Pillwein for introducing me to the symbolic computational algorithm Cylindrical Algebraic Decomposition (CAD) and to Dr. Clemens Pechstein for many nice discussions as well as to all my other colleagues at the Institute of Computational Mathematics, the Doctoral Program “Computational Mathematics” and the Johann Radon Institute for Computational and Applied Mathematics at the Johannes Kepler University for the friendly working environment and also for the administrative and technical support there.

Moreover, I thank Prof. Sergey Repin for showing so much interest in my work, for introducing me to a posteriori error estimation techniques as well as for the nice collaboration. I also express my thanks to Prof. Fredi Tröltzsch and Prof. Dietmar Hömberg for the nice scientific stay in spring 2013 in Berlin.

I gratefully acknowledge the financial support by the Austrian Science Foundation – Fond zur Förderung der wissenschaftlichen Forschung (FWF) – under the grants P19255 “Data-sparse Boundary and Finite Element Domain Decomposition Methods in Electromagnetics” and DK W1214 Doctoral Program “Computational Mathematics: Numerical Analysis and Symbolic Computation”, project DK04, and the financial support by the strategic program “Innovatives OÖ 2010 plus” by the Upper Austrian Government.

Finally, I would like to thank my parents for all their support and love throughout my whole life as well as my husband and best friend Sebastian for his encouragement, care, patience, love and understanding.

Monika Wolfmayr
Linz, February 2014

Contents

1	Introduction	1
2	Preliminaries	9
2.1	Function spaces	9
2.2	Fourier series	16
2.2.1	Fourier series expansion	16
2.2.2	Convergence of Fourier series	17
2.2.3	Fourier series expansions in the space-time cylinder	18
2.3	Variational problems	20
2.4	The finite element method	23
2.5	Parabolic partial differential equations	26
2.5.1	Parabolic initial-boundary value problems	26
2.5.2	Parabolic time-periodic boundary value problems	28
2.6	Optimal control problems	29
2.6.1	General linear-quadratic optimal control problems	29
2.6.2	Linear-quadratic parabolic optimal control problems	31
2.7	Robust block-diagonal preconditioning for the MINRES method	32
2.8	The AMLI method	36
2.8.1	The linear AMLI method	37
2.8.2	The nonlinear AMLI method	42
3	MhFE analysis of parabolic time-periodic BVPs	45
3.1	A parabolic time-periodic boundary value problem	45
3.2	Multiharmonic finite element discretization	56
3.3	Block-diagonal preconditioned MINRES solver	58
3.4	Discretization error analysis	62
3.4.1	Discretization error with respect to the truncation index	63
3.4.2	Discretization error with respect to the finite element discretization parameter	63
3.4.3	Complete discretization error	66
4	MhFE analysis of parabolic time-periodic OCPs	69
4.1	A parabolic time-periodic optimal control problem	69
4.2	Multiharmonic finite element discretization	71
4.3	Block-diagonal preconditioned MINRES solver	73
4.4	Discretization error analysis	80
4.4.1	Discretization error with respect to the truncation index	81
4.4.2	Discretization error with respect to the finite element discretization parameter	82
4.4.3	Complete discretization error	84

5	AMLI preconditioners	87
5.1	A reaction-diffusion type problem	88
5.1.1	The variational problem	88
5.1.2	The finite element discretization	88
5.2	Robust AMLI algorithms for conforming linear finite elements	90
5.2.1	The 2-refinement	92
5.2.2	The 3-refinement	94
5.2.3	The uniform m -refinement	97
5.3	Additive preconditioning of the pivot block	100
5.3.1	Additive preconditioning for the pivot block of the mass matrix	102
5.3.2	Additive preconditioner for the pivot block of the stiffness matrix	103
5.3.3	Additive preconditioning for the pivot block of the whole system	109
5.4	Stabilization polynomials of higher degree	110
5.5	The relative condition number of the preconditioned system	113
5.6	AMLI preconditioned MINRES solver	115
6	A posteriori error analysis	119
6.1	A posteriori error estimates for parabolic time-periodic BVPs	119
6.2	A posteriori error estimates for parabolic time-periodic OCPs	139
6.3	A posteriori estimates for cost functionals	163
7	Numerical results	167
7.1	Heterogeneous reaction-diffusion type problems	167
7.2	Saddle point systems	171
8	Conclusions and outlook	181
	Bibliography	185
	Eidesstattliche Erklärung	195
	Curriculum Vitae	197

Chapter 1

Introduction

The numerical solution of evolution problems is the main building block in many simulation and optimization codes for time-dependent processes in different applications like heat conduction in thermodynamics, diffusion-convection-reaction in chemistry and biology, and eddy current diffusion in electromagnetics. The simulation is usually based on models typically described via time-dependent partial differential equations (PDEs) or systems of PDEs. Together with appropriate initial and boundary conditions, we usually arrive at initial-boundary value problems (I-BVPs), the numerical solution of which is the basis for their simulation. There is a huge amount of publications on the mathematical analysis and numerical solution as well as on the numerical analysis of these numerical methods. The parabolic I-BVPs are discussed and analyzed, for instance, in the books by Ladyzhenskaya et al. [109], Wloka [177] or Zeidler [181, 182], whereas the numerics can be found, e.g., in the monographs by Lang [110], Thomée [167], or, in many other books on the numerical treatment of PDEs such as by Grossmann et al. [68] and Zulehner [188]. Optimal control problems with parabolic I-BVPs as equality constraints have extensively been investigated, e.g., in the books by Lions [117], Hinze et al. [81], Tröltzsch [169] or Borzi and Schulz [38]. Nevertheless, the analysis and numerics of evolution equations are hot research topics. In particular, the construction and analysis of highly parallel, fast and robust solution algorithms becomes more and more important not only in the simulation of evolution processes but also in their optimization where one usually needs multiple simulations including simulations of the adjoint problem that is backward in time.

Let $\Omega \subset \mathbb{R}^d$ be the computational domain with the spatial dimension $d \in \{1, 2, 3\}$ and the boundary $\Gamma = \partial\Omega$, and let $(0, T)$ be a prescribed time interval. A general form of a parabolic I-BVP is, for instance, given by

$$\begin{aligned} \sigma \partial_t u + Lu &= f && \text{in } \Omega \times (0, T), \\ u &= g_D && \text{on } \Gamma \times (0, T), \\ u &= u_0 && \text{on } \bar{\Omega} \times \{0\}, \end{aligned}$$

with the unknown function u , which describes the temperature evolution in heat conduction or the concentration of some substance in chemistry, and the following given data: the source term f , the (Dirichlet) boundary data g_D (other boundary conditions are possible), the coefficient σ and the initial data u_0 . In heat conduction, the coefficient $\sigma = c\rho$ is the product of the heat capacity c and the density ρ , whereas in electromagnetics, the coefficient σ refers to the (electric) conductivity which is positive in conducting regions like iron, but zero in non-conducting regions like air. We mention that in two space dimensions the eddy current problems turn to scalar parabolic problems as given above. In the scalar parabolic PDE given above, the elliptic operator L can have the quite general form

$$Lu(\mathbf{x}, t) = -\operatorname{div}(a(\mathbf{x}) \nabla u(\mathbf{x}, t)) + b(\mathbf{x})^T \nabla u(\mathbf{x}, t) + c(\mathbf{x}) u(\mathbf{x}, t) \quad (\mathbf{x}, t) \in \Omega \times (0, T),$$

with some coefficients a , b and c , which may depend on the function u (or on the gradient ∇u of the function u), leading to nonlinear problems with nonlinear operators, often denoted by $L(u)$.

A very general but often used linear parabolic model problem is the one with the operator

$$Lu(\mathbf{x}, t) = -\operatorname{div}(\nu(\mathbf{x}) \nabla u(\mathbf{x}, t)) \quad (\mathbf{x}, t) \in \Omega \times (0, T),$$

and homogeneous Dirichlet boundary conditions, i.e., $g_D = 0$ on $\Gamma \times (0, T)$. Other boundary conditions are Neumann or Robin boundary conditions, but, for simplicity, we focus on (homogeneous) Dirichlet boundary conditions in this work. Here, the diffusion coefficient is denoted by ν , since ν often refers to the reluctivity, e.g., in electromagnetics, where $\nu = 1/\mu$ with μ being the (magnetic) permeability. There exist many publications on parabolic I-BVPs including their numerical treatment such as the already mentioned books [109, 177, 181, 182, 110, 167, 68, 188], the book by Gustafsson et al. [72], the papers [53, 56, 55] and the references therein.

During the last couple of decades, PDE-constrained optimization has become more and more important in research and application, for which Lions has definitely paved the way with his work [117] in 1971. A typical distributed parabolic optimal control problem (OCP) is given by the minimization of a cost functional $\mathcal{J}(y, u)$ with respect to a control u and the corresponding state y , e.g.,

$$\mathcal{J}(y, u) = \frac{1}{2} \int_0^T \int_{\Omega} (y - y_d)^2 \, d\mathbf{x} \, dt + \frac{\lambda}{2} \int_0^T \int_{\Omega} u^2 \, d\mathbf{x} \, dt$$

subject to some PDE constraints such as

$$\begin{aligned} \sigma(\mathbf{x}) \partial_t y(\mathbf{x}, t) - \operatorname{div}(\nu(\mathbf{x}) \nabla y(\mathbf{x}, t)) &= u(\mathbf{x}, t) & (\mathbf{x}, t) \in \Omega \times (0, T), \\ y(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Gamma \times (0, T), \\ y(\mathbf{x}, 0) &= y_0(\mathbf{x}) & \mathbf{x} \in \overline{\Omega}. \end{aligned}$$

Here, y_d denotes the desired state. Besides the PDE constraints, there are often inequality or box constraints imposed on the control and the state, i.e.,

$$u_a \leq u \leq u_b \quad \text{and} \quad y_a \leq y \leq y_b \quad \text{a.e. in } \Omega \times (0, T),$$

respectively, where u_a , u_b , y_a , y_b are some given data leading to PDE-constrained optimization with control or state constraints. Some recent published books considering PDE-constrained optimization are, e.g., the books by Hinze et al. [81], Tröltzsch [169] and Borzi and Schulz [38], which have already been mentioned in the context of I-BVPs, and, there are also many other publications on OCPs for time-dependent I-BVPs such as [36, 125, 129, 71, 5, 65, 140, 170].

State of the art

Parabolic *time-periodic* boundary value and optimal control problems

Time-periodic conditions occur in many practical applications, as, e.g., in chemistry or electromagnetics. In electromagnetics, source terms or target functions are often time-periodic or even time-harmonic and, hence, also the solution of the problems has the corresponding property, see, e.g., [69, 70], where the authors consider the optimal control of so-called magnetohydrodynamic (MHD) equations for viscous, incompressible, electrically conducting fluids. In the time-periodic case, the initial condition is typically replaced by the periodicity condition

$$u(\mathbf{x}, 0) = u(\mathbf{x}, T) \quad \mathbf{x} \in \overline{\Omega},$$

and we call T the time period. Moreover, the solution of T -anti-periodic problems is closely related to $2T$ -periodic ones, which is, for instance, studied in [135] by Okochi.

Time-periodic PDEs have been discussed, e.g., in the works of Steuerwalt [163], Hackbusch [74], Vejvoda [175], Vandewalle and Piessens [171, 172], Lieberman [116] and Pao [137]. As Steuerwalt comments in [163], the existence of a time-periodic solution implies the solvability of the corresponding initial-value problem and vice versa. The unique solvability of parabolic I-BVPs as well as of the parabolic problems in the time-periodic case are also discussed by Zeidler [181, 182]. Nevertheless, the variety of publications discussing I-BVPs is much wider than of those discussing time-periodic BVPs. This disparity becomes even bigger in the context of PDE-constrained optimization, although the interest in analyzing and solving time-periodic OCPs is increasing, see, e.g., [1, 88, 146, 89, 112] and [94, 91, 90, 95, 180, 96]. The latter group of papers is devoted to the optimal control of time-harmonic and time-periodic eddy current problems. In this context, we want to mention the theses [24, 92, 93, 87] as well.

Fourier series and multiharmonic finite element approximation

By now, the approximation via Fourier series is commonly used for both, the space and the time discretization of PDEs. Since we consider time-periodic problems, it is a very natural approach to expand the given data as well as the unknown functions into Fourier series in time, i.e., to decompose them into sums of a (infinite) set of cosine and sine functions. Fourier series approximation is also used for space discretization. For instance, Heinrich [77] and Bernardi et al. [32] apply Fourier series approximation on problems with axisymmetric domains reducing three-dimensional problems to problems of several two-dimensional equations. This approach has been used and analyzed even earlier, e.g., in Canuto et al. [44], where the authors have also presented discretization error estimates. Fourier series are trigonometric series and are named after *Jean-Baptiste Joseph Fourier* (1768-1830), a French mathematician and physicist. Although other prominent mathematicians as Leonhard Euler, Jean le Rond d'Alembert and Daniel Bernoulli studied trigonometric series as well, Fourier introduced them in order to solve the heat equation in a metal plate in [58], and, hence, pioneered the solution to the heat equation in the general case and also to many other mathematical and physical problems. The Fourier series expansion of a time-periodic function $f(\mathbf{x}, t)$ with time period T and with frequency $\omega = 2\pi/T$ is given by

$$f(\mathbf{x}, t) = f_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [f_k^c(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \sin(k\omega t)],$$

where $k \in \mathbb{N}$ are referred as the modes, and f_0^c as well as f_k^c and f_k^s denote the corresponding amplitudes and are called Fourier coefficients. They depend on the spatial variable \mathbf{x} . For the numerical treatment, we apply the so-called multiharmonic approximation of the Fourier series, which means that we truncate the Fourier series at a finite index $N \in \mathbb{N}$. The multiharmonic approximation of a Fourier series is a trigonometric polynomial. If the boundary value and optimal control problems are linear, then, due to the orthogonality of the cosine and sine functions, the equations decouple into those depending only on the Fourier coefficients with respect to each single mode. The approximation of the Fourier coefficients, which depend on the spatial variable \mathbf{x} , can be performed by means of the finite element method (FEM). Altogether, the method combining the multiharmonic and the finite element discretization is called *multiharmonic finite element method (MhFEM)* or harmonic-balanced finite element method. It was successfully used for simulating electromagnetic devices which can be described by the eddy current approximation to Maxwell's equations, see [179, 138, 52, 73, 25, 26, 27, 49]. Later it has been applied to time-periodic parabolic optimal control problems [88, 105, 89, 112] and to time-periodic eddy current optimal control problems [91, 90, 95].

Space-time methods for parabolic problems

The MhFEM belongs to the family of space-time discretization techniques. The main idea of space-time discretization methods is to treat the time variable simply as an additional spatial variable, i.e., the so-called space-time cylinder $\Omega \times (0, T)$ with $\Omega \subset \mathbb{R}^d$ has dimension $d + 1$. Space and time can

contemporaneously be discretized in advance yielding very efficient solvers which can be parallelized much easier than time-integration methods. During the last couple of years, the analysis and practical application of space-time discretization techniques are of much interest and are now more than ever a hot research topic due to the appearance of powerful parallel computers with several thousands or even hundreds of thousands processors.

The rather traditional methods for solving parabolic problems are the so-called time-stepping or time-marching methods, where the solution is computed iteratively on successive time subintervals, see, e.g., [167]. In general, time-stepping methods are difficult to perform in parallel. Hence, methods were developed to overcome this difficulty. In [75], Hackbusch introduced a scheme for a simultaneous execution of the elliptic multigrid method on successive time steps, where the time-direction is treated as an axis in the space-time grid. Moreover, Womble [178] considered the so-called parallel time-stepping method and another space-time method came up, which was only used for space parallelism before, i.e., the multigrid waveform relaxation method, see, e.g., Vandewalle and Piessens [171]. In [83], Horton and Vandewalle presented a space-time multigrid method for solving the whole space-time parabolic problem in parallel. Other works on space-time methods are, e.g., Babuška and Janik [21, 22], who consider the h-p version of the finite element method for parabolic equations, and Costabel [50] studying boundary integral operators for the heat equation. Moreover, there have been presented a lot of new techniques and results in space-time discretization in the last couple of years as, for instance, the so-called parareal method proposed by Lions et al. [118] and further analyzed in, e.g., [63]. For more details and further investigations regarding space-time methods, we refer the reader to, e.g., [36, 160, 158, 131] or the recent publications [6, 7, 121, 130].

Robust and optimal solvers for parabolic problems

There are a couple of parameters involved in parabolic time-periodic simulation and optimal control problems, e.g., the discretization parameters in space and time as well as parameters which correspond to the conductivity and the reluctivity in electromagnetics. Moreover, the regularization or cost parameter comes into play in the optimal control problems. Besides the discretization error analysis, the construction of fast solvers, which are robust with respect to these "bad" parameters, is an important issue and a hot research topic during the last couple of years, see, e.g., [82, 157, 187, 124, 105, 141], the paper [37], which reviews research on multigrid methods for optimization problems, or the paper [1], where the authors present a nested multigrid method for solving time-periodic parabolic optimal control problems based on the works by Hackbusch [74, 76].

In the context of parabolic time-periodic problems, the preconditioned minimal residual (MINRES) method proves to provide efficient solvers, since the MhFEM leads to saddle point systems of the form

$$\underbrace{\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}}_{=: \mathcal{A} \in \mathbb{R}^{n \times n}} u = f,$$

where \mathcal{A} is a regular, symmetric, but indefinite system matrix. Hence, the goal is to construct preconditioners for the preconditioned MINRES method yielding robust convergence rates and optimal complexity, see, e.g., [187, 88, 89, 112]. The practical implementation of the preconditioners can be done by (algebraic) multigrid, multilevel or domain decomposition methods, see, e.g., [168, 173, 101, 142]. The (linear) algebraic multilevel iteration (AMLI) method stands out among these methods, since it provides a basic concept for proving robustness and optimality of the method, see the fundamental papers [14, 15] and, e.g., [13, 17, 98, 33, 126, 101]. In this context, we want to mention that in the last couple of years the use of tools from symbolic computation for proving statements arising in numerics have become much more popular, such as the standard tool Cylindrical Algebraic Decomposition (CAD), see, e.g., [47, 48, 86], of which several implementations are available, see, e.g., [164, 43, 159]. For instance, applications of this tool in the context of multigrid and multilevel solvers for optimal control problems can be found in [165, 145, 103].

Functional a posteriori error estimation

Nowadays, a posteriori error estimation together with mesh-adaptive methods is well-established in the numerical analysis of PDEs, see, e.g., [153, 123] and the references therein. Among many other techniques, the class of functional a posteriori error estimates provides a useful variational approach to a posteriori error estimation, which is based on the works by Repin, see, e.g., [150, 151, 153]. Other a posteriori error estimation techniques can be found in, e.g., the early works by Babuška and Rheinboldt [23], Bank and Weiser [28], Zienkiewicz and Zhu [183] as well as in the books by Verfürth [176] and Ainsworth and Oden [4]. In the context of functional a posteriori error estimates for optimal control problems, we refer the reader to, e.g., the papers [59, 60].

In view of the a posteriori error estimates presented later in this work, we want to emphasize the paper [152], where a method of deriving a posteriori error estimates for parabolic I-BVPs is suggested, and provides a basic concept for deriving functional a posteriori error estimates for parabolic time-periodic problems as well. More precisely, this a posteriori error estimation technique is a very useful approach to the MhFE approximation of parabolic time-periodic problems.

On this work

This thesis is devoted to parabolic time-periodic simulation and optimal control problems. More precisely, we present a complete MhFE analysis of parabolic time-periodic boundary value and the corresponding optimal control problems and use the MhFEM as discretization technique for the parabolic time-periodic problems. The resulting finite element systems are solved by the preconditioned MINRES method, where we construct robust preconditioners of optimal complexity by the AMLI method.

Main achievements

Existence and uniqueness results and a new variational framework. We prove the unique solvability of a parabolic time-periodic boundary value problem in a special variational setting after introducing function spaces and formulating variational problems in the spirit of Ladyzhenskaya et al. [109]. Here, we introduce the space $H^{1, \frac{1}{2}}$ defined for L^2 -functions on the $d+1$ -dimensional space-time cylinder, whose spatial gradients are in $[L^2]^d$ and the L^2 -norms of their half-time derivatives in the Fourier space are finite. In this setting, we are able to establish inf-sup and sup-sup conditions from which we deduce existence and uniqueness for the parabolic time-periodic problems by applying the theorem of Babuška and Aziz.

Robust preconditioners for the MINRES method. The multiharmonic finite element discretization of both problems, the parabolic time-periodic boundary value and optimal control problem, leads to systems of linear algebraic equations which decouple into smaller systems. We construct block-diagonal preconditioners for these systems which yield robust and fast convergence rates for the MINRES method following the work by Zulehner in [187].

Robust algebraic multilevel preconditioners of optimal complexity. The diagonal blocks of the MINRES preconditioners are “weighted” sums of stiffness and mass matrices, which can be seen as finite element discretization of reaction-diffusion type problems with heterogeneous reaction and diffusion coefficients. One of the main contributions of this work is not only to construct efficient practical preconditioners via the AMLI method but to present a rigorous proof of the robustness and optimal complexity of multilevel preconditioners for reaction-diffusion type problems in two space dimensions. Of course, we present some numerical results which confirm our theoretical findings.

A priori and a posteriori error analysis. We present full a priori and a posteriori error estimates for both, the parabolic time-periodic boundary value and optimal control problems, including the a

priori and a posteriori error analysis for the space $H^{1,\frac{1}{2}}$. The a posteriori error analysis is based on the method presented in Repin [152] but contains proper changes regarding the space $H^{1,\frac{1}{2}}$ and the special features of the MhFEM.

Outline

The thesis is organized as follows:

- Chapter 2 provides fundamental results needed in the subsequent chapters including definitions and theorems in the field of function spaces, Fourier series, variational problems, the finite element method as well as parabolic partial differential equations and optimal control problems. Moreover, we present a strategy for constructing robust preconditioners for the MINRES method in order to solve linear saddle point systems which arise in the context of parabolic time-periodic boundary value and optimal control problems. Finally, some basic results on the AMLI method are provided.
- In Chapter 3, we discuss parabolic time-periodic boundary value problems and present the multiharmonic finite element analysis including existence and uniqueness as well as a priori error estimates for these problems. Moreover, we present the construction of a block-diagonal preconditioner for the MINRES method which yields robust and fast convergence.
- Chapter 4 is designed and organized like Chapter 3, but considers parabolic time-periodic optimal control problems, which creates further challenges. Here, we present again the multiharmonic finite element analysis, a preconditioned MINRES solver and a priori error estimates for this type of optimal control problems.
- Chapter 5 provides a very detailed discussion and proofs on the robustness and optimality of multilevel preconditioners for reaction-diffusion type problems. This topic is strongly motivated by the previous two chapters.
- In Chapter 6, we present a functional a posteriori error analysis for parabolic time-periodic boundary value problems and the corresponding optimal control problems.
- Chapter 7 presents numerical experiments using the linear AMLI method which we have constructed and analyzed in Chapter 5 as well as numerical results on solving the boundary value and optimal control problems of Chapters 3 and 4 by various AMLI preconditioned MINRES solvers.
- In Chapter 8, we draw some conclusions and give an outlook on some future work in connection with the multiharmonic finite element approach.

Parts of this thesis have already been published by the author (and co-authors) in peer-reviewed international journals or proceedings of international conferences:

- Parts of Chapter 3 and Chapter 4 have been addressed in
 - [89] M. Kollmann, M. Kolmbauer, U. Langer, M. Wolfmayr, and W. Zulehner. A finite element solver for a multiharmonic parabolic optimal control problem. *Comput. Math. Appl.*, 65(3):469-486, 2013.
 - [111] U. Langer and M. Wolfmayr. Multiharmonic finite element solvers for time-periodic parabolic optimal control problems. *Proc. Appl. Math. Mech. (PAMM)*, 12(1):687-688, 2012.
 - [112] U. Langer and M. Wolfmayr. Multiharmonic finite element analysis of a time-periodic parabolic optimal control problem. *J. Numer. Math.*, 21(4):265-300, 2013.

- Parts of Chapter 5 have been addressed in [103] J. Kraus and M. Wolfmayr. On the robustness and optimality of algebraic multilevel methods for reaction-diffusion type problems. *Comput. Vis. Sci.*, 2014. (to appear).
- Parts of Chapter 7 have been presented in [89, 103, 112].

Chapter 2

Preliminaries

The purpose of this chapter is to provide definitions as well as fundamental results which are essential for the further investigations presented in this thesis. More precisely, our preliminary results address the following fields of research: function spaces, Fourier series, variational problems, the finite element method, parabolic partial differential equations and optimal control problems. Furthermore, we present a strategy for constructing robust preconditioners for the minimal residual method in order to solve linear saddle point systems. Finally, we also provide some basic results on the algebraic multilevel iteration method for solving symmetric and positive definite systems of algebraic equations. In this work, it will be step by step obvious that all these very different parts in mathematics are important for the multiharmonic finite element analysis of time-periodic parabolic simulation and optimal control problems.

2.1 Function spaces

In this section, we present definitions and basic results on Sobolev spaces and Bochner spaces which are used for the treatment of parabolic partial differential equations and optimal control problems. For more details, we mainly refer the reader to the books by Adams and Fournier [3], Evans [57], Tröltzsch [169], Steinbach [162] and Zeidler [181, 182].

Banach and Hilbert spaces

Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, where $d \in \{1, 2, 3\}$, and let $\bar{\Omega}$ be the closure of Ω . We denote by a *Banach space* $\{X, \|\cdot\|_X\}$ a vector space X equipped with a norm $\|\cdot\|_X$ such that the space X is complete with respect to this norm, i.e., every Cauchy sequence in X converges. An example for a classical Banach space is $C(\bar{\Omega})$, the space of all bounded and uniformly continuous functions $u \in C(\bar{\Omega})$ on $\bar{\Omega}$, equipped with the norm

$$\|u\|_{C(\bar{\Omega})} = \sup_{\mathbf{x} \in \bar{\Omega}} |u(\mathbf{x})|.$$

If the norm $\|\cdot\|_X$ is induced by an inner product $(\cdot, \cdot)_X : X \times X \rightarrow \mathbb{R}$, i.e.,

$$\|u\|_X = (u, u)_X^{1/2} \quad \forall u \in X,$$

then $\{X, (\cdot, \cdot)_X\}$ is called a *Hilbert space*. The *dual* X^* of a vector space X over the set of real numbers \mathbb{R} is defined as the set of all bounded, linear functionals f mapping from X to \mathbb{R} , equipped with the norm

$$\|f\|_{X^*} = \sup_{u \in X \setminus \{0\}} \frac{f(u)}{\|u\|_X}.$$

Here, we also introduce the *duality product* $\langle \cdot, \cdot \rangle_{X^*, X} : X^* \times X \rightarrow \mathbb{R}$ and

$$f(u) := \langle f, u \rangle_{X^*, X}.$$

We denote by $L^p(\Omega)$ the Banach space of all Lebesgue-measurable functions u defined on Ω for which

$$\|u\|_{L^p(\Omega)} = \left(\int_{\Omega} |u(\mathbf{x})|^p d\mathbf{x} \right)^{1/p} < \infty,$$

where $1 \leq p < \infty$. In the following, we mean by the denotation measurable always Lebesgue-measurable. It can be shown that $\|\cdot\|_{L^p(\Omega)}$ is a norm on $L^p(\Omega)$ by verifying the *triangle inequality*

$$\|u + v\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)} \quad \forall u, v \in L^p(\Omega),$$

which is known as the *Minkowski inequality*. The other two norm axioms are trivial. For $p = \infty$, we define the norm

$$\|u\|_{L^\infty(\Omega)} = \operatorname{ess\,sup}_{\mathbf{x} \in \Omega} |u(\mathbf{x})| = \inf\{C \geq 0 : |u(\mathbf{x})| \leq C \text{ a.e. on } \Omega\}.$$

We are especially interested in the space $L^2(\Omega)$ which is the set of all real-valued square-integrable functions on Ω . It can be shown that the space $\{L^2(\Omega), (\cdot, \cdot)_{L^2(\Omega)}\}$ is a Hilbert space introducing the inner product

$$(u, v)_{L^2(\Omega)} = \int_{\Omega} u(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}$$

and equipping with the corresponding norm $\|\cdot\|_{L^2(\Omega)} = (\cdot, \cdot)_{L^2(\Omega)}^{1/2}$. In this context, we want to mention another important inequality, namely the *Cauchy-Schwarz inequality*, also called *Cauchy-Bunyakowski-Schwarz (CBS) inequality*, i.e.,

$$|(u, v)| \leq \|u\| \|v\| \quad \forall u, v \in L^2(\Omega), \quad (2.1)$$

where, here, $(\cdot, \cdot) = (\cdot, \cdot)_{L^2(\Omega)}$ and $\|\cdot\| = \|\cdot\|_{L^2(\Omega)}$.

Strong and weak convergence

Let X be a real Banach space equipped with the norm $\|\cdot\|_X$. A sequence $\{u_n\}_{n=1}^\infty \subset X$ *converges strongly* to some $u \in X$, if

$$\lim_{n \rightarrow \infty} \|u_n - u\|_X = 0.$$

A sequence $\{u_n\}_{n=1}^\infty \subset X$ *converges weakly* to some $u \in X$, if

$$\lim_{n \rightarrow \infty} f(u_n) = f(u) \quad \forall f \in X^*.$$

Weak convergence in a Hilbert space $\{X, (\cdot, \cdot)_X\}$ is equivalent to

$$\lim_{n \rightarrow \infty} (u_n, v)_X = (u, v)_X \quad \forall v \in X.$$

We denote by the symbols \rightarrow and \rightharpoonup strong and weak convergence, respectively. Hence, we write, e.g., for weak convergence, $u_n \rightharpoonup u$ as $n \rightarrow \infty$. Strong convergence implies weak convergence. More precisely, if a sequence $\{u_n\}_{n=1}^\infty \subset X$ converges strongly to $u \in X$, then it converges weakly to u as well.

Differentiability classes and weak derivatives

A function u is of class C^k with $k \in \mathbb{N}$, if all derivatives of u up to the order k exist and are continuous. We denote by $C^\infty(\Omega)$ the space of all infinitely differentiable functions, also called *smooth* functions, and by $C_0^\infty(\Omega)$ the space of all smooth functions that have compact support in Ω . Moreover, we call a function u *analytic*, if u is smooth and if it equals its Taylor series expansion around any point in its domain.

Let $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ be a multi-index and x^α be the monomial $x_1^{\alpha_1} \dots x_d^{\alpha_d}$ with the degree $|\alpha| = \sum_{i=1}^d \alpha_i$. We define the differential operator

$$D^\alpha u = \frac{\partial^{|\alpha|}}{\partial x^\alpha} u = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}} u$$

for all functions $u \in C^{|\alpha|}(\Omega)$. If $|\alpha| = 1$ with $\alpha_i = 1$ for $1 \leq i \leq d$, then we write

$$\partial_{x_i} u = \frac{\partial}{\partial x_i} u.$$

The spaces $C^k(\overline{\Omega})$ are Banach spaces equipped with the norm

$$\|u\|_{C^k(\overline{\Omega})} = \max_{0 \leq |\alpha| \leq k} \sup_{\mathbf{x} \in \Omega} |D^\alpha u(\mathbf{x})|.$$

Moreover, we denote by $w = D^\alpha u \in L^2(\Omega)$ the α -th *weak derivative* of $u \in L^2(\Omega)$, if

$$(w, v)_{L^2(\Omega)} = (-1)^{|\alpha|} (u, D^\alpha v)_{L^2(\Omega)} \quad \forall v \in C_0^\infty(\Omega).$$

Sobolev spaces

Let $1 \leq p < \infty$ and $k \in \mathbb{N}$. We define the Banach space $W_p^k(\Omega)$ by

$$W_p^k(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega) \text{ with } |\alpha| \leq k\},$$

which is called *Sobolev space* and is equipped with the norm

$$\|u\|_{W_p^k(\Omega)} = \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{1/p}.$$

Analogously, we introduce the space $W_\infty^k(\Omega)$ for $p = \infty$ with the norm

$$\|u\|_{W_\infty^k(\Omega)} = \max_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)}.$$

Again, we are especially interested in the case $p = 2$. We set

$$H^k(\Omega) := W_2^k(\Omega),$$

which is a Hilbert space by introducing the inner product

$$(u, v)_{H^k(\Omega)} = \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}.$$

In this work, we are mostly concerned with the Hilbert space $H^1(\Omega)$ equipped with the norm

$$\|u\|_{H^1(\Omega)} = \left(\|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 \right)^{1/2}.$$

Here, the symbol $\nabla = \nabla_{\mathbf{x}}$ denotes the *weak (spatial) gradient*. Moreover, we have that

$$\|\nabla u\|_{L^2(\Omega)}^2 = \int_{\Omega} |\nabla u(\mathbf{x})|^2 d\mathbf{x} = \int_{\Omega} \left((\partial_{x_1} u(\mathbf{x}))^2 + \cdots + (\partial_{x_d} u(\mathbf{x}))^2 \right) d\mathbf{x},$$

where we introduce the seminorm in the space $H^1(\Omega)$ by

$$|u|_{H^1(\Omega)} = \|\nabla u\|_{L^2(\Omega)}.$$

Altogether, we can define the Hilbert space $H^1(\Omega)$ as follows

$$H^1(\Omega) = \{u \in L^2(\Omega) : \nabla u \in [L^2(\Omega)]^d\},$$

where the weak gradient meets the identity

$$\int_{\Omega} \nabla u \cdot \mathbf{v} d\mathbf{x} = - \int_{\Omega} u \operatorname{div} \mathbf{v} d\mathbf{x} \quad \forall \mathbf{v} \in [C_0^\infty(\Omega)]^d.$$

Let $\Gamma = \partial\Omega$ be the boundary of the domain Ω . Later on, we will include homogeneous Dirichlet boundary conditions in our partial differential equations. For that, let $\gamma_{\Gamma} u \in C(\Gamma)$ denote the restriction of a continuous function $u \in C(\bar{\Omega})$ to the boundary Γ , i.e.,

$$(\gamma_{\Gamma} u)(\mathbf{x}) = u(\mathbf{x}) \quad \forall \mathbf{x} \in \Gamma.$$

The operator γ_{Γ} is called (Dirichlet) *trace operator*, for which it can be shown the extension to a linear and bounded operator

$$\gamma_{\Gamma} : H^s(\Omega) \rightarrow H^{s-1/2}(\Gamma)$$

for $1/2 < s < 3/2$. So, the notation

$$u = 0 \text{ on } \Gamma$$

stands for

$$\gamma_{\Gamma} u = 0.$$

Hence, we are now in the position to introduce the space $H_0^1(\Omega)$ by

$$H_0^1(\Omega) = \{u \in H^1(\Omega) : u = 0 \text{ on } \Gamma\}.$$

The space $H_0^1(\Omega)$ can also be defined as the closure of the space $C_0^\infty(\Omega)$ in $H^1(\Omega)$, has the same norm as $H^1(\Omega)$ and is a closed subspace of this Hilbert space.

Moreover, let us define the Hilbert space $H(\operatorname{div}, \Omega)$ as follows

$$H(\operatorname{div}, \Omega) = \{\mathbf{u} \in [L^2(\Omega)]^d : \operatorname{div} \mathbf{u} \in L^2(\Omega)\},$$

where the *weak (spatial) divergence* is defined by the identity

$$\int_{\Omega} \operatorname{div} \mathbf{u} v d\mathbf{x} = - \int_{\Omega} \mathbf{u} \cdot \nabla v d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega).$$

For ease of notation, we will use the symbols $(\cdot, \cdot)_{L^2(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$ as well as the symbols $(\cdot, \cdot)_{H^1(\Omega)}$ and $\|\cdot\|_{H^1(\Omega)}$ for indicating both the scalar and the vector-valued case. We denote the L^2 -inner product by

$$(\mathbf{u}, \mathbf{v})_{L^2(\Omega)} = \sum_{i=1}^n (u_i, v_i)_{L^2(\Omega)},$$

where $\mathbf{u} = (u_1, \dots, u_n)^T$ and $\mathbf{v} = (v_1, \dots, v_n)^T$ are vector functions. The associated norm is given by

$$\|\mathbf{u}\|_{L^2(\Omega)}^2 = (\mathbf{u}, \mathbf{u})_{L^2(\Omega)}.$$

Since we are considering time-dependent problems in this work, we need to define Sobolev spaces which include the time domain as well.

Sobolev spaces defined in the space-time cylinder

Let $Q_T := \Omega \times (0, T)$ denote the space-time cylinder, where $(0, T)$ is a given time interval, and let $\Sigma_T = \Gamma \times (0, T)$ be its mantle boundary. To begin with, we introduce two Sobolev spaces of functions living on the space-time cylinder Q_T in the spirit of Ladyzhenskaya et al., which are used, e.g., in the monograph [109] for studying initial-boundary value problems, i.e., the Sobolev spaces $H^{1,0}(Q_T)$ and $H^{1,1}(Q_T)$, see also [108]. The space $H^{1,0}(Q_T)$ is defined as

$$H^{1,0}(Q_T) = \{u \in L^2(Q_T) : \nabla u \in [L^2(Q_T)]^d\} \quad (2.2)$$

equipped with the norm

$$\|u\|_{H^{1,0}(Q_T)} = \left(\int_0^T \int_{\Omega} (u(\mathbf{x}, t)^2 + |\nabla u(\mathbf{x}, t)|^2) \, d\mathbf{x} \, dt \right)^{1/2}, \quad (2.3)$$

and the space $H^{1,1}(Q_T)$ as

$$H^{1,1}(Q_T) = \{u \in L^2(Q_T) : \nabla u \in [L^2(Q_T)]^d, \partial_t u \in L^2(Q_T)\} \quad (2.4)$$

with the corresponding norm

$$\|u\|_{H^{1,1}(Q_T)} = \left(\int_0^T \int_{\Omega} (u(\mathbf{x}, t)^2 + |\nabla u(\mathbf{x}, t)|^2 + |\partial_t u(\mathbf{x}, t)|^2) \, d\mathbf{x} \, dt \right)^{1/2}, \quad (2.5)$$

where $\nabla = \nabla_{\mathbf{x}}$ and ∂_t denote the weak spatial gradient and the weak time derivative, respectively. Analogously, we define the Sobolev space

$$H^{0,1}(Q_T) = \{u \in L^2(Q_T) : \partial_t u \in L^2(Q_T)\}.$$

Furthermore, we can include the condition $u = 0$ on Σ_T by defining the Sobolev spaces

$$H_0^{1,0}(Q_T) = \{u \in H^{1,0}(Q_T) : u = 0 \text{ on } \Sigma_T\} \quad \text{and} \quad H_0^{1,1}(Q_T) = \{u \in H^{1,1}(Q_T) : u = 0 \text{ on } \Sigma_T\}.$$

In the context of time-periodic problems, we will also consider the time-periodicity condition $u(\mathbf{x}, 0) = u(\mathbf{x}, T)$. Hence, we define the Sobolev spaces

$$\begin{aligned} H_{per}^{0,1}(Q_T) &= \{u \in H^{0,1}(Q_T) : u(\mathbf{x}, 0) = u(\mathbf{x}, T) \text{ for almost all } \mathbf{x} \in \Omega\}, \\ H_{per}^{1,1}(Q_T) &= \{u \in H^{1,1}(Q_T) : u(\mathbf{x}, 0) = u(\mathbf{x}, T) \text{ for almost all } \mathbf{x} \in \Omega\}, \\ H_{0,per}^{1,1}(Q_T) &= \{u \in H_0^{1,1}(Q_T) : u(\mathbf{x}, 0) = u(\mathbf{x}, T) \text{ for almost all } \mathbf{x} \in \Omega\}. \end{aligned}$$

Bochner spaces

The concept of L^p -spaces can be generalized to functions which have values in a Banach space. Let $\{X, \|\cdot\|_X\}$ be a Banach space. We denote by the *Bochner space* $L^p(a, b; X)$, $1 \leq p < \infty$, the linear space of all equivalence classes of Lebesgue-measurable functions $u : (a, b) \rightarrow X$ such that the corresponding norm is finite, i.e.,

$$\|u\|_{L^p(a,b;X)} = \left(\int_a^b \|u(t)\|_X^p \, dt \right)^{1/p} < \infty.$$

The space $L^p(a, b; X)$ is a Banach space with respect to this norm. Analogously, we define the Bochner space for $p = \infty$ equipped with the norm

$$\|u\|_{L^\infty(a,b;X)} = \operatorname{ess\,sup}_{t \in (a,b)} \|u(t)\|_X < \infty.$$

Furthermore, we define the space $C([a, b]; X)$ consisting of all functions $u : [a, b] \rightarrow X$ which are continuous at every $t \in [a, b]$. The space $C([a, b]; X)$ is equipped with the norm

$$\|u\|_{C([a,b];X)} = \max_{t \in [a,b]} \|u(t)\|_X.$$

We can extend the concept of a Bochner space by defining a more general space, i.e.,

$$Y(a, b; X) = \{u : (a, b) \rightarrow X : \|u\|_{Y(a,b;X)} < \infty\},$$

where $\{Y, \|\cdot\|_Y\}$ and $\{X, \|\cdot\|_X\}$ are Banach spaces.

For the treatment of parabolic problems, we are interested in the Bochner spaces $L^p(a, b; X)$ with $(a, b) = (0, T)$ being the time interval, $p = 2$ and either $X = H^1(\Omega)$ or $X = H_0^1(\Omega)$ depending on the boundary conditions. For example, the space $L^2(0, T; H^1(\Omega))$ is defined as

$$L^2(0, T; H^1(\Omega)) = \{u : (0, T) \rightarrow H^1(\Omega) : \|u\|_{L^2(0,T;H^1(\Omega))} < \infty\},$$

where

$$\begin{aligned} \|u\|_{L^2(0,T;H^1(\Omega))} &= \left(\int_0^T \|u(t)\|_{H^1(\Omega)}^2 dt \right)^{1/2} \\ &= \left(\int_0^T \int_{\Omega} (u(\mathbf{x}, t)^2 + |\nabla u(\mathbf{x}, t)|^2) d\mathbf{x} dt \right)^{1/2}. \end{aligned}$$

Moreover, we denote by

$$\partial_t u \in L^2(0, T; H^1(\Omega)^*)$$

the *generalized weak derivative* of $u \in L^2(0, T; H^1(\Omega))$, if there holds

$$\int_0^T u(t) \partial_t \varphi(t) dt = - \int_0^T \partial_t u(t) \varphi(t) dt \quad \forall \varphi \in C_0^\infty(0, T).$$

Analogously, we denote by $\partial_t u \in L^2(0, T; H^{-1}(\Omega))$ the generalized weak derivative of the function $u \in L^2(0, T; H_0^1(\Omega))$. Here, we can define the duality product $\langle \cdot, \cdot \rangle_{X^*, X} : X^* \times X \rightarrow \mathbb{R}$ by using the definition of the generalized weak derivative, i.e.,

$$\int_0^T \langle \partial_t u(t), \varphi(t) \rangle_{X^*, X} dt = \int_0^T \partial_t u(t) \varphi(t) dt,$$

where $\partial_t u \in L^2(0, T; X^*)$, $\varphi \in L^2(0, T; X)$ and $\partial_t \varphi \in L^2(0, T; X^*)$, for which the integration by parts formula

$$\int_0^T \langle \partial_t u(t), \varphi(t) \rangle_{X^*, X} dt = (y(T), \varphi(T))_{L^2(\Omega)} - (y(0), \varphi(0))_{L^2(\Omega)} - \int_0^T \langle \partial_t \varphi(t), u(t) \rangle_{X^*, X} dt$$

holds. The space $W(0, T)$ denotes the linear space of all $u \in L^2(0, T; X)$ having a generalized weak derivative $\partial_t u \in L^2(0, T; X^*)$ and equipped with the norm

$$\|u\|_{W(0,T)} = \left(\int_0^T (\|u(t)\|_X^2 + \|\partial_t u(t)\|_{X^*}^2) dt \right)^{1/2}.$$

In a nutshell, it can be defined as

$$W(0, T) = \{u \in L^2(0, T; X) : \partial_t u \in L^2(0, T; X^*)\},$$

which is a Hilbert space equipped with the inner product

$$(u, v)_{W(0, T)} = \int_0^T (u(t), v(t))_X dt + \int_0^T (\partial_t u, \partial_t v)_{X^*} dt.$$

Note that the space $W(0, T)$ is continuously embedded in $C([0, T]; L^2(\Omega))$, see, e.g., [177, 181]. The chain of dense and continuous embeddings

$$H^1(\Omega) \subset L^2(\Omega) \subset H^1(\Omega)^*$$

as well as

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega)$$

is called a *Gelfand triple* or *evolution triple*, in general, denoted by

$$V \subset H = H^* \subset V^*,$$

where V is a real, separable and reflexive Banach space and H a real and separable Hilbert space, see, e.g., [177].

Relationship between the function spaces

Comparing the norms corresponding to the Sobolev spaces defined in the space-time cylinder and to the Bochner spaces suggests the following relationships between these spaces:

$$\begin{aligned} L^2(Q_T) &\cong L^2(0, T; L^2(\Omega)), \\ H^{1,0}(Q_T) &\cong L^2(0, T; H^1(\Omega)), \\ H_0^{1,0}(Q_T) &\cong L^2(0, T; H_0^1(\Omega)), \end{aligned}$$

where the symbol \cong denotes the equivalence of the corresponding norms. In fact, one can show that the two spaces are isometric and isomorphic, see [79]. Let us define the space $H^1(0, T; L^2(\Omega))$ consisting of all functions $u : (0, T) \rightarrow L^2(\Omega)$ for which $\|u\|_{H^1(0, T; L^2(\Omega))} < \infty$, where

$$\|u\|_{H^1(0, T; L^2(\Omega))} = \left(\int_0^T \|u\|_{L^2(\Omega)}^2 + \|\partial_t u\|_{L^2(\Omega)}^2 dt \right)^{1/2}.$$

The treatment of the norms corresponding to the spaces $H^{1,1}(Q_T)$, $H_0^{1,1}(Q_T)$ and $W(0, T)$ provides that

$$\begin{aligned} H^{1,1}(Q_T) &\cong L^2(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega)), \\ H_0^{1,1}(Q_T) &\cong L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega)), \end{aligned}$$

and the space $W(0, T)$ with $X = H^1(\Omega)$ or $X = H_0^1(\Omega)$ is a larger space than $H^{1,1}(Q_T)$ or $H_0^{1,1}(Q_T)$, respectively.

Finally, we introduce the function spaces $\tilde{H}^{1,0}(Q_T)$ and $\tilde{H}_0^{1,0}(Q_T)$ in order to study parabolic initial-boundary value problems following again Ladyzhenskaya [108] and Ladyzhenskaya et al. [109]. The spaces $\tilde{H}^{1,0}(Q_T)$ and $\tilde{H}_0^{1,0}(Q_T)$ are defined as

$$\tilde{H}^{1,0}(Q_T) = H^{1,0}(Q_T) \cap C([0, T]; L^2(\Omega)) \tag{2.6}$$

and

$$\tilde{H}_0^{1,0}(Q_T) = H_0^{1,0}(Q_T) \cap C([0, T]; L^2(\Omega)), \tag{2.7}$$

respectively, equipped with the norm

$$\|u\|_{\tilde{H}^{1,0}(Q_T)} = \max_{t \in [0, T]} \|u(t)\|_{L^2(\Omega)} + \left(\int_0^T \int_{\Omega} |\nabla u(\mathbf{x}, t)|^2 d\mathbf{x} dt \right)^{1/2}. \quad (2.8)$$

Here, the elements of $H^{1,0}(Q_T)$ and $H_0^{1,0}(Q_T)$ have to be identified with abstract functions belonging to $L^2(0, T; H^1(Q_T))$ and $L^2(0, T; H_0^1(Q_T))$, respectively.

2.2 Fourier series

The aim of this section is to give an overview on Fourier series and on some convergence results for Fourier series. For more details, we refer the reader to [85] and the references therein.

2.2.1 Fourier series expansion

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a periodic function, i.e., $f(0) = f(T)$ with period T and with frequency $\omega = 2\pi/T$. The *Fourier series expansion* of f is denoted by

$$f(t) = f_0^c + \sum_{k=1}^{\infty} [f_k^c \cos(k\omega t) + f_k^s \sin(k\omega t)],$$

where the *Fourier coefficients* of f are defined by

$$\begin{aligned} f_0^c &= \frac{1}{T} \int_0^T f(t) dt, \\ f_k^c &= \frac{2}{T} \int_0^T f(t) \cos(k\omega t) dt, \\ f_k^s &= \frac{2}{T} \int_0^T f(t) \sin(k\omega t) dt, \end{aligned}$$

for $k \in \mathbb{N}$, which are referred as the modes. The Fourier series of a function f is a trigonometric series. The *multiharmonic approximation* of this Fourier series is a trigonometric polynomial, written as

$$f_N(t) = f_0^c + \sum_{k=1}^N [f_k^c \cos(k\omega t) + f_k^s \sin(k\omega t)].$$

The following orthogonalities are valid:

$$\begin{aligned} \frac{1}{T} \int_0^T \cos(0) \cos(0) dt &= 1, \\ \frac{2}{T} \int_0^T \cos(k\omega t) \sin(l\omega t) dt &= 0, \\ \frac{2}{T} \int_0^T \cos(k\omega t) \cos(l\omega t) dt &= \delta_{kl}, \\ \frac{2}{T} \int_0^T \sin(k\omega t) \sin(l\omega t) dt &= \delta_{kl}, \end{aligned} \quad (2.9)$$

where $k, l \in \mathbb{N}$ and δ_{kl} is the Kronecker delta. Due to the representation

$$\cos(k\omega t) = \frac{e^{ik\omega t} + e^{-ik\omega t}}{2}, \quad \sin(k\omega t) = \frac{e^{ik\omega t} - e^{-ik\omega t}}{2i},$$

we derive a complex form for the Fourier series of f , i.e.,

$$\begin{aligned}
 f(t) &= f_0^c + \sum_{k=1}^{\infty} \left[f_k^c \left(\frac{e^{ik\omega t} + e^{-ik\omega t}}{2} \right) + f_k^s \left(\frac{e^{ik\omega t} - e^{-ik\omega t}}{2i} \right) \right] \\
 &= f_0^c + \sum_{k=1}^{\infty} \left[\left(\frac{f_k^c}{2} + \frac{f_k^s}{2i} \right) e^{ik\omega t} + \left(\frac{f_k^c}{2} - \frac{f_k^s}{2i} \right) e^{-ik\omega t} \right] \\
 &= f_0 + \sum_{k=1}^{\infty} [f_k e^{ik\omega t} + f_{-k} e^{-ik\omega t}] \\
 &= \sum_{k=-\infty}^{\infty} f_k e^{ik\omega t},
 \end{aligned}$$

where $f_0 = f_0^c$ and, for $k \in \mathbb{N}$, $f_k = \frac{f_k^c}{2} + \frac{f_k^s}{2i}$ and $f_{-k} = \frac{f_k^c}{2} - \frac{f_k^s}{2i}$. Now, the orthogonalities read as

$$\frac{1}{T} \int_0^T e^{ik\omega t} e^{il\omega t} dt = \begin{cases} 1 & k+l=0, \\ 0 & k+l \neq 0, \end{cases}$$

since $e^{it} = \cos(t) + i \sin(t)$. Moreover, we can rewrite the complex representation of the Fourier series of f as

$$\sum_{k=-\infty}^{\infty} f_k e^{ik\omega t} = \sum_{k=-\infty}^{\infty} f_k (e^{i\omega t})^k = \sum_{k=-\infty}^{\infty} f_k z^k = \tilde{f}(z)$$

with $z = e^{i\omega t}$.

2.2.2 Convergence of Fourier series

In the treatment of Fourier series, the question of convergence arises very early. We start with three rather standard types of convergence results for Fourier series. More precisely, we consider pointwise and uniform convergence as well as norm convergence. This subsection only overviews the standard results. Proofs and more details as well as other references can be found in [85].

Pointwise and uniform convergence

Let $f \in L^1(0, T)$ and $t_0 \in [0, T]$. Assume that the limit

$$\lim_{h \rightarrow 0} [f(t_0 + h) + f(t_0 - h)]$$

exists, then f_N converges pointwise to the one-sided limits at t_0 , i.e.,

$$f_N(t_0) \rightarrow \frac{1}{2} \lim_{h \rightarrow 0} [f(t_0 + h) + f(t_0 - h)].$$

If f is continuous in t_0 , then $f_N(t_0)$ converges to $f(t_0)$. Moreover, if f is continuous in every point of the closed interval $[0, T]$, then f_N converges uniformly to f on $[0, T]$, i.e.,

$$\lim_{N \rightarrow \infty} \sup_{t \in (0, T)} |f(t) - f_N(t)| = 0.$$

Otherwise, if f_N converges pointwise to f in $[0, T]$ and is uniformly bounded, i.e., there exists a constant K such that $|f_N(t)| \leq K$ for all t and for all N , then f_N converges to f in the L^1 -norm. Moreover, uniform convergence is a sufficient condition for interchanging integrals and limits.

Norm convergence

The *Riesz-Fischer theorem* states that a Lebesgue-measurable function f on $(0, T)$ is square integrable if and only if its corresponding Fourier series converges in L^2 . Hence, if $f \in L^2(0, T)$, then

$$\lim_{N \rightarrow \infty} \|f - f_N\|_{L^2(0, T)} = 0.$$

The norm convergence (strong convergence) still holds in L^p -spaces with $1 < p < \infty$, i.e., if $f \in L^p(0, T)$, then f_N converges to f in the L^p -norm.

2.2.3 Fourier series expansions in the space-time cylinder

Let $f \in L^2(Q_T)$ be a time-periodic function. Then f can be expanded into a Fourier series in time, which is denoted by

$$f(\mathbf{x}, t) = f_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [f_k^c(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \sin(k\omega t)], \quad (2.10)$$

where the Fourier coefficients $f_k^c(\mathbf{x})$, $f_k^s(\mathbf{x})$, with $k \in \mathbb{N}$, and $f_0^c(\mathbf{x})$ are all from the space $L^2(\Omega)$ and are defined as

$$\begin{aligned} f_0^c(\mathbf{x}) &= \frac{1}{T} \int_0^T f(\mathbf{x}, t) dt, \\ f_k^c(\mathbf{x}) &= \frac{2}{T} \int_0^T f(\mathbf{x}, t) \cos(k\omega t) dt, \\ f_k^s(\mathbf{x}) &= \frac{2}{T} \int_0^T f(\mathbf{x}, t) \sin(k\omega t) dt. \end{aligned} \quad (2.11)$$

The multiharmonic approximation $f_N(\mathbf{x}, t)$ of $f(\mathbf{x}, t)$ is defined in the same way as for a function $f : \mathbb{R} \rightarrow \mathbb{R}$. Moreover, all orthogonalities (2.9) as well as the known convergence results are also valid for space-time dependent functions, which can be expanded into Fourier series in time. According to the Riesz-Fischer theorem functions $f \in L^2(Q_T)$ converge strongly in the corresponding L^2 -norm.

Remark 2.1. The L^2 -norm of functions $f \in L^2(Q_T)$ is defined in the Fourier space by

$$\|f\|_{L^2(Q_T)}^2 = \int_0^T \int_{\Omega} f(\mathbf{x}, t) d\mathbf{x} dt = T \|f_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} [\|f_k^c\|_{L^2(\Omega)}^2 + \|f_k^s\|_{L^2(\Omega)}^2]$$

due to the orthogonalities (2.9) of the cosine and sine functions and because integrals and sums can be interchanged.

Theorem 2.2. Let $f \in H_0^{1,0}(Q_T)$. Then the Fourier series representation (2.10) of f converges strongly in $H_0^{1,0}(Q_T)$.

Proof. Let us take an arbitrary test function $\varphi(\mathbf{x}) \in [C_0^\infty(\Omega)]^d$. Strong convergence in $L^2(Q_T)$ yields weak convergence in $L^2(Q_T)$. Hence, the limit

$$- \int_0^T \int_{\Omega} f_N(\mathbf{x}, t) \operatorname{div} \varphi(\mathbf{x}) d\mathbf{x} dt \rightarrow - \int_0^T \int_{\Omega} f(\mathbf{x}, t) \operatorname{div} \varphi(\mathbf{x}) d\mathbf{x} dt \quad (2.12)$$

exists, where

$$\begin{aligned} - \int_0^T \int_{\Omega} f_N(\mathbf{x}, t) \operatorname{div} \varphi(\mathbf{x}) d\mathbf{x} dt &= - \int_0^T \int_{\Omega} \left(\sum_{k=0}^N [f_k^c(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \sin(k\omega t)] \right) \operatorname{div} \varphi(\mathbf{x}) d\mathbf{x} dt \\ &= - \int_0^T \int_{\Omega} \left(\sum_{k=0}^N [f_k^c(\mathbf{x}) \operatorname{div} \varphi(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \operatorname{div} \varphi(\mathbf{x}) \sin(k\omega t)] \right) d\mathbf{x} dt. \end{aligned}$$

Integration by parts with respect to the spatial variable \mathbf{x} leads to

$$\begin{aligned} & - \int_0^T \int_{\Omega} \left(\sum_{k=0}^N [f_k^c(\mathbf{x}) \operatorname{div} \boldsymbol{\varphi}(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \operatorname{div} \boldsymbol{\varphi}(\mathbf{x}) \sin(k\omega t)] \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(\sum_{k=0}^N [\nabla f_k^c(\mathbf{x}) \cdot \boldsymbol{\varphi}(\mathbf{x}) \cos(k\omega t) + \nabla f_k^s(\mathbf{x}) \cdot \boldsymbol{\varphi}(\mathbf{x}) \sin(k\omega t)] \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(\sum_{k=0}^N [\nabla f_k^c(\mathbf{x}) \cos(k\omega t) + \nabla f_k^s(\mathbf{x}) \sin(k\omega t)] \right) \cdot \boldsymbol{\varphi}(\mathbf{x}) d\mathbf{x} dt. \end{aligned}$$

Since the limit (2.12) exists, we define the weak limit

$$\nabla f_N \rightharpoonup \nabla f \quad \text{in } [L^2(\Omega)]^d \quad \text{for } N \rightarrow \infty. \quad (2.13)$$

Moreover, since $\nabla f = \nabla_{\mathbf{x}} f \in [L^2(Q_T)]^d$ (as $f \in H_0^{1,0}(Q_T)$), ∇f can be expanded into a Fourier series in time and its Fourier coefficients are given by

$$\begin{aligned} \nabla f_0^c(\mathbf{x}) &:= \frac{1}{T} \int_0^T \nabla f(\mathbf{x}, t) dt, \\ \nabla f_k^c(\mathbf{x}) &:= \frac{2}{T} \int_0^T \nabla f(\mathbf{x}, t) \cos(k\omega t) dt, \\ \nabla f_k^s(\mathbf{x}) &:= \frac{2}{T} \int_0^T \nabla f(\mathbf{x}, t) \sin(k\omega t) dt, \end{aligned}$$

which are all from the space $[L^2(\Omega)]^d$. Hence, the limit (2.13) is also strong and can be defined as

$$\nabla f(\mathbf{x}, t) := \lim_{N \rightarrow \infty} \sum_{k=0}^N [\nabla f_k^c(\mathbf{x}) \cos(k\omega t) + \nabla f_k^s(\mathbf{x}) \sin(k\omega t)],$$

i.e., the gradient of the Fourier series expansion of a function is equal to the Fourier series expansion of the gradient of the function. \square

Theorem 2.3. *Let $f \in H_0^{1,1}(Q_T)$. Then the Fourier series representation (2.10) of f converges strongly in $H_0^{1,1}(Q_T)$.*

Proof. Due to Theorem 2.2, it remains to show that $\partial_t f$ converges strongly in $L^2(0, T)$. Let us take an arbitrary test function $\varphi(t)$ from $C_0^\infty(0, T)$. The convergence in $L^2(Q_T)$ leads to

$$- \int_{\Omega} \int_0^T f_N(\mathbf{x}, t) \partial_t \varphi(t) dt d\mathbf{x} \rightarrow - \int_{\Omega} \int_0^T f(\mathbf{x}, t) \partial_t \varphi(t) dt d\mathbf{x}, \quad (2.14)$$

where

$$\begin{aligned} - \int_{\Omega} \int_0^T f_N(\mathbf{x}, t) \partial_t \varphi(t) dt d\mathbf{x} &= - \int_{\Omega} \int_0^T \left(\sum_{k=0}^N [f_k^c(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \sin(k\omega t)] \right) \partial_t \varphi(t) dt d\mathbf{x} \\ &= - \int_{\Omega} \int_0^T \left(\sum_{k=0}^N [f_k^c(\mathbf{x}) \cos(k\omega t) \partial_t \varphi(t) + f_k^s(\mathbf{x}) \sin(k\omega t) \partial_t \varphi(t)] \right) dt d\mathbf{x}. \end{aligned}$$

Then integration by parts with respect to the time variable t yields

$$\begin{aligned} & - \int_{\Omega} \int_0^T \left(\sum_{k=0}^N [f_k^c(\mathbf{x}) \cos(k\omega t) \partial_t \varphi(t) + f_k^s(\mathbf{x}) \sin(k\omega t) \partial_t \varphi(t)] \right) dt d\mathbf{x} \\ &= \int_{\Omega} \int_0^T \left(\sum_{k=1}^N [-k\omega f_k^c(\mathbf{x}) \sin(k\omega t) \varphi(t) + k\omega f_k^s(\mathbf{x}) \cos(k\omega t) \varphi(t)] \right) dt d\mathbf{x} \\ &= \int_{\Omega} \int_0^T \left(\sum_{k=1}^N k\omega [-f_k^c(\mathbf{x}) \sin(k\omega t) + f_k^s(\mathbf{x}) \cos(k\omega t)] \right) \varphi(t) dt d\mathbf{x} \end{aligned}$$

and since the limit (2.14) exists, we define the weak limit

$$\partial_t f_N(\mathbf{x}, t) \rightharpoonup \partial_t f(\mathbf{x}, t) \quad \text{in } L^2(0, T) \quad \text{for } N \rightarrow \infty. \quad (2.15)$$

Moreover, since $\partial_t f \in L^2(Q_T)$ (as $f \in H_0^{1,1}(Q_T)$), $\partial_t f$ can be expanded into a Fourier series in time. Due to integration by parts with respect to t , the Fourier coefficients of $\partial_t f$ are given by

$$\begin{aligned} (\partial_t f)_k^c(\mathbf{x}) &:= \frac{2}{T} \int_0^T \partial_t f(\mathbf{x}, t) \cos(k\omega t) dt = -\frac{2}{T} \int_0^T f(\mathbf{x}, t) (-k\omega \sin(k\omega t)) dt \\ &= k\omega \frac{2}{T} \int_0^T f(\mathbf{x}, t) \sin(k\omega t) dt = k\omega f_k^s(\mathbf{x}), \\ (\partial_t f)_k^s(\mathbf{x}) &:= \frac{2}{T} \int_0^T \partial_t f(\mathbf{x}, t) \sin(k\omega t) dt = -\frac{2}{T} \int_0^T f(\mathbf{x}, t) (k\omega \cos(k\omega t)) dt \\ &= -k\omega \frac{2}{T} \int_0^T f(\mathbf{x}, t) \cos(k\omega t) dt = -k\omega f_k^c(\mathbf{x}). \end{aligned}$$

Hence, the limit (2.15) is also strong and can be defined as

$$\partial_t f(\mathbf{x}, t) := \lim_{N \rightarrow \infty} \sum_{k=1}^N k\omega [-f_k^c(\mathbf{x}) \sin(k\omega t) + f_k^s(\mathbf{x}) \cos(k\omega t)].$$

□

2.3 Variational problems

In this section, we provide some of the basic results on variational problems in Hilbert spaces. We start with the Friedrichs and the Poincaré inequalities and then present some fundamental results as the well-known Lax-Milgram theorem and the Babuška-Aziz theorem.

Theorem 2.4 (Poincaré inequality). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with $d \in \{1, 2, 3\}$. Then, there exists a constant $C_P > 0$ depending only on Ω such that for all $u \in H^1(\Omega)$ we have that*

$$\|u - u_{\Omega}\|_{L^2(\Omega)} \leq C_P |u|_{H^1(\Omega)}, \quad (2.16)$$

where

$$u_{\Omega} = \frac{1}{|\Omega|} \int_{\Omega} u(\mathbf{x}) d\mathbf{x}$$

is the mean value of u over Ω and with $|\Omega|$ being the Lebesgue measure of the domain Ω .

Proof. See, e.g., [42, 142, 168].

□

If the mean value of u is zero, i.e., $\int_{\Omega} u(\mathbf{x}) \, d\mathbf{x} = 0$, then (2.16) reads as

$$\|u\|_{L^2(\Omega)} \leq C_P |u|_{H^1(\Omega)}.$$

A closely related result to the Poincaré inequality is the Friedrichs inequality.

Theorem 2.5 (Friedrichs inequality). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with $d \in \{1, 2, 3\}$. Then, there exists a constant $C_F > 0$ depending only on Ω such that for all $u \in H_0^1(\Omega)$ we have that*

$$\|u\|_{L^2(\Omega)} \leq C_F |u|_{H^1(\Omega)} = C_F \|\nabla u\|_{L^2(\Omega)}. \quad (2.17)$$

Proof. See, e.g., [42, 168]. □

The Friedrichs inequality is often used in order to prove that the H^1 -seminorm can be estimated from below by the H^1 -norm. Due to

$$\|u\|_{H^1(\Omega)}^2 = \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 \leq (C_F^2 + 1) \|\nabla u\|_{L^2(\Omega)}^2,$$

it follows that

$$|u|_{H^1(\Omega)}^2 = \|\nabla u\|_{L^2(\Omega)}^2 \geq \frac{1}{C_F^2 + 1} \|u\|_{H^1(\Omega)}^2.$$

Remark 2.6. *The proofs of Theorems 2.4 and 2.5 do not tell anything about the dependence of the constants C_F and C_P on the shape of Ω . Only for certain classes of simple domains, explicit constants can be computed. For example, explicit Poincaré constants for star-shaped domains are presented in [42, 174]. Moreover, we can explicitly compute a very simple estimate of the Poincaré constant for convex domains, i.e.,*

$$C_P \leq \frac{\text{diam} \Omega}{\pi},$$

where $\text{diam} \Omega$ denotes the diameter of Ω , see [29, 139]. For further information regarding the Poincaré and Friedrichs inequalities, we refer the reader to the books, e.g., [42, 142, 168] and to the papers, e.g., [2, 128, 143].

Now, we want to present the fundamental existence and uniqueness results of Lax and Milgram as well as of Babuška and Aziz for variational problems. Let $\{V, (\cdot, \cdot)_V\}$ be a Hilbert space with the associated norm

$$\|\cdot\|_V = (\cdot, \cdot)_V^{1/2},$$

and let $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ be a bilinear form and $F : V \rightarrow \mathbb{R}$ be a linear form. Let us consider the following abstract variational problem: Find $u \in V$ such that

$$a(u, v) = \langle F, v \rangle \quad \forall v \in V. \quad (2.18)$$

We can rewrite the variational problem (2.18) as an operator equation by introducing the operator $A : V \rightarrow V^*$, $u \mapsto a(u, \cdot)$, i.e.,

$$\langle Au, v \rangle_{V^*, V} := a(u, v) \quad \forall u, v \in V.$$

Then the operator equation reads as follows: Find $u \in V$ such that

$$Au = F \quad \text{in } V^*. \quad (2.19)$$

The following theorem by Lax and Milgram [114] states which conditions on the bilinear form $a(\cdot, \cdot)$ have to be satisfied in order to guarantee the existence and uniqueness of the solution to the variational problem (2.18).

Theorem 2.7 (Lax-Milgram). *Let V be a Hilbert space and let the bilinear form $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ be elliptic (coercive) and bounded (continuous), i.e.,*

$$\begin{aligned} \underline{c}\|u\|_V^2 &\leq a(u, u) & \forall u \in V, \\ |a(u, v)| &\leq \bar{c}\|u\|_V\|v\|_V & \forall u, v \in V, \end{aligned}$$

with constants $\underline{c}, \bar{c} > 0$. Then, for all bounded linear functionals $F \in V^*$, the variational problem (2.18) has a unique solution $u \in V$, which fulfills the a priori estimate

$$\frac{1}{\bar{c}}\|F\|_{V^*} \leq \|u\|_V \leq \frac{1}{\underline{c}}\|F\|_{V^*}.$$

Proof. The statement of the theorem follows directly from the linearity, ellipticity and boundedness of the bilinear form as well as by the definition of the dual norm. For more details, we refer the reader to the original paper by Lax and Milgram [114], see also [42, 186] and the references therein. \square

We present now an extension of the Lax-Milgram theorem, i.e., the theorem of Babuška and Aziz or also known as the Babuška-Lax-Milgram (or generalized Lax-Milgram) theorem, which is due to Babuška, see [18, 19, 20]. This generalized version of the Lax-Milgram theorem can also be applied to the variational problem (2.18), if the solution u is not from the same space as the test functions v , i.e., $u \in U$ and $v \in V$ with $U \neq V$, and if the bilinear form is not elliptic (coercive), but fulfills the weaker conditions

$$\begin{aligned} \inf_{0 \neq u \in U} \sup_{0 \neq v \in V} \frac{a(u, v)}{\|v\|_V\|u\|_U} &\geq \underline{c} > 0, \\ \inf_{0 \neq v \in V} \sup_{0 \neq u \in U} \frac{a(u, v)}{\|u\|_U\|v\|_V} &\geq \underline{c} > 0, \end{aligned}$$

which are called inf-sup conditions or weakly coercive. Bounded bilinear forms fulfilling the inf-sup conditions are already mentioned in Nečas [132], but their use for variational boundary value problems was presented by Babuška [18, 19] and Babuška-Aziz [20], see [134]. Note that it is obvious that the Babuška-Aziz theorem is valid for $U = V$ as well and that the Babuška-Aziz theorem with $U = V$ is not equivalent to the Lax-Milgram theorem, cf. Remark 2.9. We now state the Babuška-Aziz theorem for $U \neq V$, see, e.g., [20, 134, 149].

Theorem 2.8 (Babuška-Aziz). *Let $\{U, \|\cdot\|_U\}$ and $\{V, \|\cdot\|_V\}$ be Hilbert spaces and let $F : V \rightarrow \mathbb{R}$ be a bounded linear form, i.e., $F \in V^*$. Assume that there exist constants $\underline{c}, \bar{c} > 0$ such that the bilinear form $a(\cdot, \cdot) : U \times V \rightarrow \mathbb{R}$ fulfills the inf-sup and sup-sup conditions*

$$\underline{c}\|u\|_U \leq \sup_{0 \neq v \in V} \frac{a(u, v)}{\|v\|_V} \leq \bar{c}\|u\|_U \quad \forall u \in U, \quad (2.20)$$

$$\underline{c}\|v\|_V \leq \sup_{0 \neq u \in U} \frac{a(u, v)}{\|u\|_U} \leq \bar{c}\|v\|_V \quad \forall v \in V. \quad (2.21)$$

Then, for any $F \in V^*$, the variational problem

$$a(u, v) = \langle F, v \rangle \quad \forall v \in V,$$

has a unique solution $u \in U$, which fulfills the a priori estimate

$$\frac{1}{\bar{c}}\|F\|_{V^*} \leq \|u\|_U \leq \frac{1}{\underline{c}}\|F\|_{V^*}.$$

Proof. See, e.g., [20, 134]. \square

Remark 2.9. The inf-sup condition of (2.20) yields the definition of the inf-sup constant as follows

$$c_{\text{inf-sup}} = \inf_{0 \neq u \in U} \sup_{0 \neq v \in V} \frac{|a(u, v)|}{\|u\|_U \|v\|_V}.$$

In case of $U = V$ and $a(u, v) = a(v, u)$ for all $u, v \in V$, the inf-sup and sup-sup conditions (2.20) and (2.21) in Theorem 2.8 simplify to one inf-sup and sup-sup condition, i.e.,

$$\underline{c} \|u\|_V \leq \sup_{0 \neq v \in V} \frac{a(u, v)}{\|v\|_V} \leq \bar{c} \|u\|_V \quad \forall u \in V, \quad (2.22)$$

and the inf-sup constant is given by

$$c_{\text{inf-sup}} = \inf_{0 \neq u \in V} \sup_{0 \neq v \in V} \frac{|a(u, v)|}{\|u\|_V \|v\|_V} = \underline{c}.$$

Moreover, the relation between the ellipticity constant of the Lax-Milgram theorem, which we denote by \underline{c}_{L-M} , and the inf-sup constant $c_{\text{inf-sup}}$ is given by $\underline{c}_{L-M} \leq c_{\text{inf-sup}}$ and can be proven by

$$\underline{c}_{L-M} \|u\|_V \leq \frac{|a(u, u)|}{\|u\|_V} \leq \sup_{0 \neq v \in V} \frac{|a(u, v)|}{\|v\|_V} \quad \forall 0 \neq u \in V,$$

which implies

$$\underline{c}_{L-M} \leq \inf_{0 \neq u \in V} \sup_{0 \neq v \in V} \frac{|a(u, v)|}{\|u\|_V \|v\|_V} = c_{\text{inf-sup}}.$$

Hence, the bilinear form of a variational problem does not necessarily have to be elliptic (strongly coercive), but has to satisfy the inf-sup condition (weakly coercive) in order to guarantee existence and uniqueness.

2.4 The finite element method

In this section, we consider the finite element method (FEM), which is a numerical procedure for finding approximate solutions of boundary value problems. The finite element method goes back to the paper by Courant [51], where, again, the introduced method is based on the methods by Ritz [155] and Galerkin [62], see also [64]. For more details, we refer the reader to, e.g., [40, 42, 46, 162, 168] or the German books [41, 84, 161, 186].

Let $\Omega \subset \mathbb{R}^d$ be again a bounded Lipschitz domain with $d \in \{1, 2, 3\}$. Moreover, we assume that Ω is a polygonal or polyhedral domain. Such a domain can be subdivided into finitely many non-overlapping elements τ . Domains with a curved boundary are, e.g., discussed in [46, 97, 184, 185].

Definition 2.10. We denote by $\mathcal{T}_h = \{\tau\}$ a mesh or triangulation of the domain $\Omega \subset \mathbb{R}^d$ into finitely many d -dimensional elements τ . Such a triangulation is called admissible if

$$\bar{\Omega} = \bigcup_{\tau \in \mathcal{T}_h} \bar{\tau} \quad \text{and} \quad \tau_i \cap \tau_j = \emptyset \quad \forall \tau_i \neq \tau_j \text{ with } i \neq j \text{ and } \tau_i, \tau_j \in \mathcal{T}_h.$$

Moreover, we denote by $\mathcal{T}_h(\Omega)$ the family of finer and finer triangulations of Ω , cf. [46].

For $d = 1$, $d = 2$ and $d = 3$, the finite elements τ are, e.g., line segments, triangles and tetrahedra, respectively. Hence, for any two elements of an admissible triangulation, the intersection is either empty, an element vertex, edge or, for $d = 3$, a face. Let us denote by

$$h_\tau := \text{diam } \tau$$

the *diameter* of a finite element $\tau \in \mathcal{T}_h$ and by

$$h := \max_{\tau \in \mathcal{T}_h} h_\tau$$

the *mesh size* of \mathcal{T}_h . Moreover, ρ_τ denotes the diameter of the largest ball contained in τ .

Definition 2.11. We call a triangulation \mathcal{T}_h *shape-regular* if there exists a constant $C > 0$, which is independent of h , such that

$$h_\tau \leq C \rho_\tau \quad \forall \tau \in \mathcal{T}_h.$$

Definition 2.12. We call a triangulation \mathcal{T}_h *quasi-uniform* if there exists a constant $C > 0$, which is independent of h , such that

$$h_\tau \geq C h \quad \forall \tau \in \mathcal{T}_h.$$

Remark 2.13. From the definition of shape-regular and quasi-uniform follows that a quasi-uniform triangulation is also shape-regular but not vice versa.

Let us consider the abstract variational problem (2.18). We want to discretize this problem by the so-called *Ritz-Galerkin finite element method*. Following this method, let V_h be a finite dimensional subspace of V . We construct the space V_h by choosing appropriate basis functions φ_i , i.e.,

$$V_h = \text{span}\{\varphi_1, \dots, \varphi_n\},$$

where $n = \dim V_h$. The finite element approximation of u and its nodal parameter vector \underline{u} are given by

$$u_h = \sum_{i=1}^n u_i \varphi_i \in V_h \quad \text{and} \quad \underline{u} = (u_i)_{i=1, \dots, n} \in \mathbb{R}^n,$$

respectively, and they are related via the so-called *Ritz isomorphism*, a one-to-one mapping between them. We arrive at the following discrete variational problem: Find the approximate solution $u_h \in V_h$ such that

$$a(u_h, v_h) = \langle F, v_h \rangle \quad \forall v_h \in V_h. \quad (2.23)$$

Testing with the basis functions φ_i yields a system of linear equations which is equivalent to the discrete variational problem, i.e.,

$$\mathbf{A} \underline{u} = \underline{f}, \quad (2.24)$$

where

$$\mathbf{A} = (a(\varphi_i, \varphi_j))_{i,j=1, \dots, n} \quad \text{and} \quad \underline{f} = (\langle F, \varphi_i \rangle)_{i=1, \dots, n}.$$

In order to prove existence and uniqueness of the square system (2.24) and, hence, of the discrete problem (2.23), only uniqueness has to be guaranteed since the problem is finite dimensional. For that, it has to be shown that it does not exist a nonzero \underline{u} such that $\mathbf{A} \underline{u} = \underline{0}$, or, in other words, the kernel of \mathbf{A} contains only the zero vector $\underline{0}$.

In the following, we are interested in the discrete spaces

$$V_h^p = \{v \in C(\bar{\Omega}) : v|_\tau \in \mathcal{P}_p(\tau) \quad \forall \tau \in \mathcal{T}_h\},$$

i.e., the spaces of continuous, piecewise polynomial functions, where $\mathcal{P}_p(\tau)$ is the space of polynomials up to the order p on the element τ . In particular, we will consider the space of continuous, piecewise linear functions, where $p = 1$,

$$V_h = V_h^1 = \{v \in C(\bar{\Omega}) : v|_\tau \in \mathcal{P}_1(\tau) \quad \forall \tau \in \mathcal{T}_h\},$$

which is a finite dimensional subspace of the Hilbert space $H^1(\Omega)$. Moreover, we can include the homogeneous Dirichlet boundary conditions by defining the discrete space

$$V_{h,0} = \{v \in C_0(\bar{\Omega}) : v|_{\tau} \in \mathcal{P}_1(\tau) \quad \forall \tau \in \mathcal{T}_h\},$$

which can be also defined as

$$V_{h,0} = V_h \cap H_0^1(\Omega).$$

The following result states that the discretization error can be estimated by the best approximation error. More precisely, the approximation $u_h \in V_h \subset V$ to the exact solution $u \in V$ is only worse up to a constant, i.e., it is quasi-optimal, see [18].

Lemma 2.14 (Céa-type estimate). *Let $u \in V$ be the solution of the variational problem (2.18) and let $u_h \in V_h \subset V$ be the solution of (2.23), where the symmetric bilinear form $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ fulfills the inf-sup and sup-sup condition (2.22). Moreover, we assume that there exists a constant $\underline{c}_d > 0$ such that the so-called discrete inf-sup condition*

$$\underline{c}_d \|u_h\|_V \leq \sup_{0 \neq v_h \in V_h} \frac{a(u_h, v_h)}{\|v_h\|_V} \quad \forall u_h \in V_h \quad (2.25)$$

is fulfilled. Then the following discretization error estimate holds:

$$\|u - u_h\|_V \leq \left(1 + \frac{\bar{c}}{\underline{c}_d}\right) \inf_{v_h \in V_h} \|u - v_h\|. \quad (2.26)$$

Proof. Since $V_h \subset V$, we have that

$$a(u, v_h) = \langle F, v_h \rangle \quad \text{and} \quad a(u_h, v_h) = \langle F, v_h \rangle \quad \forall v_h \in V_h.$$

Subtracting yields the Galerkin orthogonality

$$a(u - u_h, v_h) = 0 \quad \forall v_h \in V_h. \quad (2.27)$$

By inserting an arbitrary $v_h \in V_h$ and using triangle inequality, the discrete inf-sup condition (2.25), the sup-sup condition in (2.22) as well as the Galerkin orthogonality (2.27), we obtain the following estimate:

$$\begin{aligned} \|u - u_h\|_V &\leq \|u - v_h\|_V + \|u_h - v_h\|_V \leq \|u - v_h\|_V + \frac{1}{\underline{c}_d} \sup_{0 \neq \tilde{v}_h \in V_h} \frac{a(u_h - v_h, \tilde{v}_h)}{\|\tilde{v}_h\|_V} \\ &\leq \|u - v_h\|_V + \frac{1}{\underline{c}_d} \underbrace{\sup_{0 \neq \tilde{v}_h \in V_h} \frac{a(u_h - u, \tilde{v}_h)}{\|\tilde{v}_h\|_V}}_{=0} + \frac{1}{\underline{c}_d} \sup_{0 \neq \tilde{v}_h \in V_h} \frac{a(u - v_h, \tilde{v}_h)}{\|\tilde{v}_h\|_V} \\ &\leq \|u - v_h\|_V + \frac{\bar{c}}{\underline{c}_d} \|u - v_h\|_V = \left(1 + \frac{\bar{c}}{\underline{c}_d}\right) \|u - v_h\|_V, \end{aligned}$$

which finally provides the discretization error estimate (2.26). \square

Remark 2.15. *If the bilinear form $a(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ satisfies the ellipticity and boundedness assumptions of the Lax-Milgram theorem with constants \underline{c} and \bar{c} , see Theorem 2.7, one can prove another discretization error estimate, i.e.,*

$$\|u - u_h\|_V \leq \frac{\bar{c}}{\underline{c}} \inf_{v_h \in V_h} \|u - v_h\|,$$

which is usually called the Céa lemma or the Céa theorem, see, e.g., [42].

2.5 Parabolic partial differential equations

This section presents some basic results on parabolic partial differential equations, which can be found, e.g., in [108, 109, 177, 181, 182, 188], and is a starting point for further investigations presented in this work. We start with some variational formulations of a parabolic initial-boundary value problem including existence and uniqueness results, and then, present a corresponding parabolic time-periodic boundary value problem (BVP), which will particularly be discussed in Chapter 3.

2.5.1 Parabolic initial-boundary value problems

Let $Q_T := \Omega \times (0, T)$ denote again the space-time cylinder and $\Sigma_T := \Gamma \times (0, T)$ its mantle boundary, where $\Omega \subset \mathbb{R}^d$, $d = \{1, 2, 3\}$, is again a bounded Lipschitz domain, and let $T > 0$ be a fixed final time. We consider the parabolic initial-boundary value problem

$$\begin{aligned} \sigma(\mathbf{x}) \partial_t u(\mathbf{x}, t) - \operatorname{div}(\nu(\mathbf{x}) \nabla u(\mathbf{x}, t)) &= f(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ u(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}) & \mathbf{x} \in \bar{\Omega}, \end{aligned} \quad (2.28)$$

where $f(\mathbf{x}, t)$ and $u_0(\mathbf{x})$ are some given data. Moreover, let us assume that

$$0 < \underline{\sigma} \leq \sigma(\mathbf{x}) \leq \bar{\sigma}, \quad 0 < \underline{\nu} \leq \nu(\mathbf{x}) \leq \bar{\nu}, \quad \text{for } \mathbf{x} \in \Omega. \quad (2.29)$$

Remark 2.16. *In practical applications, as, e.g., in computational electromagnetics, σ corresponds to the (electric) conductivity and ν to the reluctivity, where $\nu = 1/\mu$ with μ being the (magnetic) permeability. Usually, the coefficients σ and ν are piecewise constant due to different materials of which electrical devices are made. A practical example is*

$$\bar{\Omega} = \bar{\Omega}_c \cup \bar{\Omega}_{nc}, \quad \Omega_c \cap \Omega_{nc} = \emptyset,$$

where Ω is a domain consisting of a conducting region Ω_c with $\sigma > 0$ and a non-conducting region Ω_{nc} with $\sigma = 0$, i.e., problem (2.28) is parabolic in Ω_c and elliptic in Ω_{nc} with appropriate interface conditions on $\bar{\Omega}_c \cap \bar{\Omega}_{nc}$.

For the time being, let us consider the case $\sigma(\mathbf{x}) = \nu(\mathbf{x}) = 1$. Then, problem (2.28) reads as follows

$$\begin{aligned} \partial_t u(\mathbf{x}, t) - \Delta u(\mathbf{x}, t) &= f(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ u(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}) & \mathbf{x} \in \bar{\Omega}. \end{aligned} \quad (2.30)$$

In this subsection, we now present well-known existence and uniqueness results for this parabolic initial-boundary value problem, see, e.g., [169, 181, 182]. We start with a first appropriate variational formulation and a corresponding existence and uniqueness statement due to Ladyzhenskaya et al., see [109], see also [108]. Let

$$f \in L^2(Q_T) \cong L^2(0, T; L^2(\Omega)) \quad \text{and} \quad u_0 \in L^2(\Omega).$$

We multiply the first equation of problem (2.30) by a test function

$$v \in C_0^{1,1}(Q_T) = \{v \in C^1(Q_T) : v = 0 \text{ on } \Sigma_T\}$$

and integrate over the space-time cylinder Q_T . So far, we argue formally according to [169], assuming

that u is a classical solution. Then, integration by parts yields

$$\begin{aligned} & \int_0^T \int_{\Omega} (\partial_t u(\mathbf{x}, t) v(\mathbf{x}, t) - \Delta u(\mathbf{x}, t) v(\mathbf{x}, t)) \, d\mathbf{x} \, dt \\ &= \int_{\Omega} (u(\mathbf{x}, T) v(\mathbf{x}, T) - u(\mathbf{x}, 0) v(\mathbf{x}, 0)) \, d\mathbf{x} - \int_0^T \int_{\Omega} u(\mathbf{x}, t) \partial_t v(\mathbf{x}, t) \, d\mathbf{x} \, dt \\ & \quad + \int_0^T \int_{\Omega} \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t) \, d\mathbf{x} \, dt \\ &= \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) \, d\mathbf{x} \, dt \end{aligned}$$

for all $v \in C_0^{1,1}(Q_T)$. Now, we can insert the initial data and require that the test functions vanish for the time $t = T$ as well. Altogether, it is enough to choose test functions $v \in H_0^{1,1}(Q_T)$ with $v(\cdot, T) = 0$. We arrive at the following variational formulation: Find $u \in H_0^{1,0}(Q_T)$ such that

$$\begin{aligned} & \int_0^T \int_{\Omega} (-u(\mathbf{x}, t) \partial_t v(\mathbf{x}, t) + \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t)) \, d\mathbf{x} \, dt \\ &= \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) \, d\mathbf{x} \, dt + \int_{\Omega} u_0(\mathbf{x}) v(\mathbf{x}, 0) \, d\mathbf{x} \end{aligned} \tag{2.31}$$

for all $v \in H_0^{1,1}(Q_T)$ with $v(\mathbf{x}, T) = 0$ for almost all $\mathbf{x} \in \Omega$.

The following theorem provides existence and uniqueness for the variational problem (2.31) and was proven by Ladyzhenskaya et al. [109], see also [108, 169].

Theorem 2.17. *Let $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a bounded Lipschitz domain, and let $Q_T := \Omega \times (0, T)$ be the space-time cylinder with $T > 0$ fixed. Under the assumptions that $f \in L^2(Q_T)$ and $u_0 \in L^2(\Omega)$, the parabolic initial-boundary value problem (2.31) has a unique solution $u \in H_0^{1,0}(Q_T)$ that belongs to $\tilde{H}_0^{1,0}(Q_T)$ and the stability estimate*

$$\max_{t \in [0, T]} \|u(\cdot, t)\|_{L^2(\Omega)} + \|u\|_{H^{1,0}(Q_T)} \leq C (\|f\|_{L^2(Q_T)} + \|u_0\|_{L^2(\Omega)})$$

holds with a constant $C > 0$, which is independent of f and u_0 .

Proof. See, e.g., [109]. □

From Theorem 2.17 follows that the linear mapping $(f, u_0) \mapsto u$ is a continuous operator from $L^2(Q_T) \times L^2(\Omega)$ into $\tilde{H}_0^{1,0}(Q_T) = H_0^{1,0}(Q_T) \cap C([0, T]; L^2(\Omega))$ and hence, into $H_0^{1,0}(Q_T)$ and $H^{1,0}(Q_T)$. This result ensures that the solution u is from the Bochner space $C([0, T]; L^2(\Omega))$, i.e., u is a continuous mapping from $[0, T]$ into $L^2(\Omega)$, see [169].

Remark 2.18. *As it is mentioned in the book by Tröltzsch [169], the variational formulation (2.31) is only conditionally suitable for the study of optimal control problems since the requirements of the solution u and of the test function v are not equal. This is a problem in the case of optimal control problems, since the adjoint state is inserted in the place of v . So, one needs a different approach.*

Let $X = H_0^1(\Omega)$. In the following, we will present a second variational formulation for the parabolic initial-boundary value problem (2.30), which is set in the space

$$W(0, T) = \{u \in L^2(0, T; X) : \partial_t u \in L^2(0, T; X^*)\},$$

as defined in Section 2.1, see also [177, 181]. Moreover, weak solutions u from $H_0^{1,0}(Q_T)$ to the parabolic initial-boundary value problem (2.30) also belong to $W(0, T)$ as it is shown in [169]. Note that there exists a functional $F \in L^2(0, T; X^*)$ such that

$$\langle F(t), v \rangle_{X^*, X} = \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}) \, d\mathbf{x} \quad \forall v \in X,$$

see [169, 182]. By the definition of the generalized weak time derivative together with the variational formulation (2.31), we arrive at the following variational problem: Find $u \in W(0, T)$ such that

$$\int_0^T \langle \partial_t u(t), v(t) \rangle_{X^*, X} dt + \int_0^T \int_{\Omega} \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t) d\mathbf{x} dt = \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) d\mathbf{x} dt \quad (2.32)$$

for all $v \in L^2(0, T; X)$ and $u(0) = u_0$.

Theorem 2.19. *Let $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a bounded Lipschitz domain, and let $Q_T := \Omega \times (0, T)$ be the space-time cylinder with $T > 0$ fixed. Under the assumptions that $f \in L^2(Q_T)$ and $u_0 \in L^2(\Omega)$, the parabolic initial-boundary value problem (2.32) has a unique solution $u \in W(0, T)$ which satisfies the stability estimate*

$$\|u\|_{W(0, T)} \leq C (\|f\|_{L^2(Q_T)} + \|u_0\|_{L^2(\Omega)})$$

with a constant $C > 0$, which is independent of f and u_0 .

Proof. See [169]. □

Theorem 2.19 concludes that the mapping $(f, u_0) \mapsto u$ defines a continuous linear operator from $L^2(Q_T) \times L^2(\Omega)$ into $W(0, T)$ and into $C([0, T]; L^2(\Omega))$ as well, since $W(0, T)$ is continuously embedded in $C([0, T]; L^2(\Omega))$.

2.5.2 Parabolic time-periodic boundary value problems

In this subsection, we want to present the parabolic time-periodic analogon to the parabolic initial-boundary value problem (2.28), which will be studied more detailed in Chapter 3. Instead of the initial condition

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}) \quad \mathbf{x} \in \Omega,$$

we now prescribe the time-periodic condition

$$u(\mathbf{x}, 0) = u(\mathbf{x}, T) \quad \mathbf{x} \in \Omega,$$

and call $T > 0$ the periodicity. Hence, we obtain the parabolic time-periodic problem

$$\begin{aligned} \sigma(\mathbf{x}) \partial_t u(\mathbf{x}, t) - \operatorname{div}(\nu(\mathbf{x}) \nabla u(\mathbf{x}, t)) &= f(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ u(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ u(\mathbf{x}, 0) &= u(\mathbf{x}, T) & \mathbf{x} \in \bar{\Omega}, \end{aligned} \quad (2.33)$$

where, again, f is some given data and we suppose that the assumptions (2.29) on the coefficients σ and ν hold.

Theorem 2.20 presents an existence and uniqueness result for the time-periodic problem (2.33), which can be found in [182]. For that, let $X = H_0^1(\Omega)$, and let us define the space

$$W_{per}(0, T) = \{u \in W(0, T) : u(\mathbf{x}, 0) = u(\mathbf{x}, T) \text{ for almost all } \mathbf{x} \in \Omega\},$$

which includes the time-periodicity conditions. We have the following variational problem: Find $u \in W_{per}(0, T)$ such that

$$\int_0^T \langle \partial_t u(t), v(t) \rangle_{X^*, X} dt + \int_0^T \int_{\Omega} \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t) d\mathbf{x} dt = \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) d\mathbf{x} dt \quad (2.34)$$

for all $v \in L^2(0, T; X)$.

Theorem 2.20. *Let $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a bounded Lipschitz domain, and let $Q_T := \Omega \times (0, T)$ be the space-time cylinder with $T > 0$ fixed. Under the assumption that $f \in L^2(Q_T)$, the parabolic time-periodic problem (2.34) has a unique solution $u \in W_{per}(0, T)$.*

Proof. See [182]. □

Time-periodic conditions occur in many practical applications, as, e.g., in electromagnetics, where we can assume that the source term f is time-periodic and, hence, also the solution u , see, e.g., [69, 70]. A very important tool for the treatment of time-periodic problems is the multiharmonic ansatz, where the source term f as well as the solution u (and the test function v) are expanded into Fourier series and are approximated by corresponding truncated Fourier series, see Chapter 3.

2.6 Optimal control problems

The main focus of this section is to present well-known results on optimal control problems, especially, on parabolic optimal control problems. For further information and a general analysis of optimal control problems, we refer the reader, e.g., to [81, 117, 169]. These general results provide a basis for the future discussion and analysis of parabolic time-periodic optimal control problems (OCPs), which are presented in Chapter 4.

2.6.1 General linear-quadratic optimal control problems

Let H , U , Y and Z be Hilbert spaces and let Y be continuously embedded in H . We denote by $y \in Y$ the *state variable* and by $u \in U$ the *control variable*. Hence, Y and U are also called the *state space* and the *control space*. To begin with, let us consider a general linear-quadratic optimal control problem of the form

$$\begin{aligned} \min_{(y,u) \in Y \times U} \mathcal{J}(y, u) &= \frac{1}{2} \|E_Y y - y_d\|_H^2 + \frac{\lambda}{2} \|u\|_U^2, \\ \text{subject to } Ay &= Bu + g, \quad u \in U_{ad}, \end{aligned} \tag{2.35}$$

where $y_d \in H$ is the given *desired state*, $g \in Z$ is a given *source term*, and $\lambda > 0$ is the *cost or regularization parameter*. The linear continuous operator

$$E_Y : Y \rightarrow H$$

is called the *embedding operator*, which assigns to each $y \in Y$ the same function in H . The equation $Ay = Bu + g$ is called *state equation* and is a partial differential equation in our context, where $A : Y \rightarrow Z$ and $B : U \rightarrow Z$ are continuous linear operators. Moreover, we assume that the *set of admissible controls* $U_{ad} \subset U$ is nonempty, closed and convex. One could also impose constraints on the state, i.e., $y \in Y_{ad} \subset Y$. However, in this work, we mainly consider the case $U_{ad} = U$ (and also $Y_{ad} = Y$), which is of course covered by the theory presented in this section. In order to derive the corresponding *reduced optimal control problem*, we define the continuous linear operators

$$G : U \rightarrow Y, \quad u \mapsto y(u) \quad \text{and} \quad S : U \rightarrow H, \quad u \mapsto y(u)$$

which are called the *control-to-state operator* and the *solution operator*, respectively. Moreover, we set $S = E_Y G$, and so, it follows that

$$Su = E_Y Gu = E_Y y.$$

Under the assumption that the operator A has a bounded inverse, i.e., the state equation has a unique solution, we obtain

$$Gu = A^{-1}(Bu + g)$$

and arrive at the following reduced optimal control problem:

$$\min_{u \in U_{ad}} f(u) = \frac{1}{2} \|S u - y_d\|_H^2 + \frac{\lambda}{2} \|u\|_U^2, \quad (2.36)$$

where f is the so-called *reduced cost functional*.

Theorem 2.21. *Under the assumptions made above, the optimal control problem (2.35) has a unique solution (\bar{y}, \bar{u}) . Moreover, the corresponding reduced optimization problem (2.36) admits the unique solution $\bar{u} \in U_{ad}$.*

Proof. See, e.g., [81], as well as [169]. \square

The following theorem provides the *first-order optimality conditions* of the reduced optimal control problem (2.36):

Theorem 2.22. *The reduced optimization problem (2.36) has a unique solution $\bar{u} \in U_{ad}$ if and only if \bar{u} solves the variational inequality*

$$(S^*(S\bar{u} - y_d) + \lambda\bar{u}, u - \bar{u})_U \geq 0 \quad \forall u \in U_{ad}. \quad (2.37)$$

Proof. See, e.g., [169]. \square

In order to derive the first-order optimality conditions, also called the *optimality system*, for problem (2.35), we introduce the so-called *adjoint state* (or *Lagrange multiplier*) $p \in P = Z^*$ and define the *Lagrange functional* $\mathcal{L} : Y \times U \times P \rightarrow \mathbb{R}$, i.e.,

$$\mathcal{L}(y, u, p) = \mathcal{J}(y, u) + \langle p, A y - B u - g \rangle_{P, Z}. \quad (2.38)$$

Using this Lagrangian based approach, we are able to state the first-order optimality conditions for problem (2.35) in a compact form.

Theorem 2.23. *The optimal control problem (2.35) has an optimal solution (\bar{y}, \bar{u}) if and only if there exists a Lagrange multiplier $\bar{p} \in P$ for the corresponding Lagrange functional (2.38) such that the following optimality conditions hold:*

$$\mathcal{L}_p(\bar{y}, \bar{u}, \bar{p}) = 0, \quad (2.39)$$

$$\mathcal{L}_y(\bar{y}, \bar{u}, \bar{p}) = 0, \quad (2.40)$$

$$\bar{u} \in U_{ad}, \quad \langle \mathcal{L}_u(\bar{y}, \bar{u}, \bar{p}), u - \bar{u} \rangle_{U^*, U} \geq 0 \quad \forall u \in U_{ad}. \quad (2.41)$$

Proof. See, e.g., [81]. \square

The first-order optimality conditions of Theorem 2.22 and Theorem 2.23 are necessary and also sufficient since the minimization functional as well as the reduced functional are convex. The condition (2.39) corresponds to the state equation, (2.41) to the variational inequality and (2.40) is referred to as *adjoint equation*. The optimality system (2.39)-(2.41) of problem (2.35) takes the form

$$\begin{aligned} A\bar{y} &= B\bar{u} + g, \\ A^*\bar{p} &= -E_Y^*(E_Y\bar{y} - y_d), \\ \bar{u} \in U_{ad}, \quad \langle \lambda\bar{u} - B^*\bar{p}, u - \bar{u} \rangle_{U^*, U} &\geq 0 \quad \forall u \in U_{ad}. \end{aligned}$$

In the unconstrained case, i.e., $U_{ad} = U$, the variational inequality (2.41) of the optimality system, simplifies to the equation $\mathcal{L}_u(\bar{y}, \bar{u}, \bar{p}) = 0$. Hence, the optimality system is given by

$$\nabla \mathcal{L}(\bar{y}, \bar{u}, \bar{p}) = 0, \quad \text{or,} \quad \begin{aligned} \mathcal{L}_p(\bar{y}, \bar{u}, \bar{p}) &= 0, \\ \mathcal{L}_y(\bar{y}, \bar{u}, \bar{p}) &= 0, \\ \mathcal{L}_u(\bar{y}, \bar{u}, \bar{p}) &= 0. \end{aligned} \quad (2.42)$$

2.6.2 Linear-quadratic parabolic optimal control problems

Let us consider the optimal control problem

$$\min_{(y,u)} \mathcal{J}(y,u) = \frac{1}{2} \|y - y_d\|_{L^2(Q_T)}^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \quad (2.43)$$

subject to the parabolic initial-boundary value problem

$$\begin{aligned} \partial_t y(\mathbf{x}, t) - \Delta y(\mathbf{x}, t) &= u(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ y(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ y(\mathbf{x}, 0) &= y_0(\mathbf{x}) & \mathbf{x} \in \bar{\Omega}, \end{aligned} \quad (2.44)$$

and the control constraints

$$u_a(\mathbf{x}, t) \leq u(\mathbf{x}, t) \leq u_b(\mathbf{x}, t) \quad \text{for a.e. } (\mathbf{x}, t) \in Q_T. \quad (2.45)$$

Due to Theorem 2.19, the parabolic initial-boundary value problem (2.44) has a unique solution $y \in W(0, T)$, i.e., there exists a linear continuous mapping $(u, y_0) \mapsto y$ from $L^2(Q_T) \times L^2(\Omega)$ into $W(0, T)$. This information leads to the following existence and uniqueness result for the parabolic optimal control problem (2.43)-(2.45):

Theorem 2.24. *Let $\lambda > 0$. Under the assumption that the parabolic initial-boundary value problem (2.44) has a unique solution in $\bar{y} \in W(0, T)$, there exists an optimal control $\bar{u} \in L^2(Q_T)$ solving the optimization problem (2.43)-(2.45).*

Proof. The existence and uniqueness statement follows immediately from applying Theorem 2.19 and Theorem 2.21, see [169]. \square

We now state the optimality system corresponding to problem (2.43)-(2.45). Let us consider an optimal control $\bar{u} \in L^2(Q_T)$ with associated state \bar{y} solving the state equation (2.44). Then, the corresponding adjoint state \bar{p} is the unique weak solution of the adjoint equation

$$\begin{aligned} -\partial_t p(\mathbf{x}, t) - \Delta p(\mathbf{x}, t) &= \bar{y}(\mathbf{x}, t) - y_d(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ p(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ p(\mathbf{x}, T) &= 0 & \mathbf{x} \in \bar{\Omega}. \end{aligned}$$

Moreover, \bar{p} and \bar{u} have to fulfill the variational inequality

$$\int_0^T \int_{\Omega} (\lambda \bar{u}(\mathbf{x}, t) - \bar{p}(\mathbf{x}, t)) (u(\mathbf{x}, t) - \bar{u}(\mathbf{x}, t)) \, d\mathbf{x} \, dt \geq 0 \quad \forall u \in U_{ad},$$

where $U_{ad} = \{u \in L^2(Q_T) : u_a(\mathbf{x}, t) \leq u(\mathbf{x}, t) \leq u_b(\mathbf{x}, t) \text{ for a.e. } (\mathbf{x}, t) \in Q_T\}$.

In case of $U_{ad} = U$, the variational inequality simplifies to the equation

$$u(\mathbf{x}, t) = \lambda^{-1} p(\mathbf{x}, t) \quad (\mathbf{x}, t) \in Q_T,$$

and we obtain the reduced optimality system

$$\begin{aligned} \partial_t y(\mathbf{x}, t) - \Delta y(\mathbf{x}, t) &= \lambda^{-1} p(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ y(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ y(\mathbf{x}, 0) &= y_0(\mathbf{x}) & \mathbf{x} \in \bar{\Omega}, \\ -\partial_t p(\mathbf{x}, t) - \Delta p(\mathbf{x}, t) &= y(\mathbf{x}, t) - y_d(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ p(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ p(\mathbf{x}, T) &= 0 & \mathbf{x} \in \bar{\Omega}. \end{aligned}$$

If, in addition, the state equation is not an initial-boundary value problem but a time-periodic one, the reduced optimality system is given by

$$\begin{aligned} \partial_t y(\mathbf{x}, t) - \Delta y(\mathbf{x}, t) &= \lambda^{-1} p(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ y(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ y(\mathbf{x}, 0) &= y(\mathbf{x}, T) & \mathbf{x} \in \bar{\Omega}, \\ \\ -\partial_t p(\mathbf{x}, t) - \Delta p(\mathbf{x}, t) &= y(\mathbf{x}, t) - y_d(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ p(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ p(\mathbf{x}, 0) &= p(\mathbf{x}, T) & \mathbf{x} \in \bar{\Omega}. \end{aligned}$$

This kind of optimal control problems will particularly be discussed in Chapter 4 including existence and uniqueness results, discretization and numerical methods as well as full discretization error estimates.

2.7 Robust block-diagonal preconditioning for the MINRES method

In the context of simulation and optimal control of parabolic time-periodic problems, there are often arising bilinear forms with a saddle point structure which, after discretization, lead to linear saddle point systems of the form

$$\mathcal{A}u = f, \quad (2.46)$$

where

$$\mathcal{A} = \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (2.47)$$

is a regular, symmetric, but indefinite system matrix. Here, the symmetric matrices A and C are assumed to be positive definite and positive semidefinite, respectively. Therefore, the linear system (2.46) can be solved by a preconditioned *minimal residual (MINRES) method* which was introduced by Paige and Saunders [136]. This method belongs to the class of preconditioned *Krylov subspace methods*, which are very efficient iterative methods for solving large scale linear systems, see, e.g., [156]. A very popular Krylov subspace method is the conjugate gradient (CG) method for problems with symmetric positive definite system matrices, see [78].

The purpose of this section is to provide the fundamental results which are needed to construct robust preconditioned MINRES solvers for the parabolic time-periodic problems presented in Chapters 3 and 4. The construction of efficient preconditioners is subject of discussion in many papers, see, e.g., the survey paper [30]. In particular, we mention the so-called operator preconditioning technique that exploits the mapping property of the underlying operator and that leads to block-diagonal preconditioners for saddle point problems like (2.46), see [27, 82, 124, 173] and the references therein. By applying the preconditioned MINRES method, we aim at minimizing the preconditioned residual

$$r_m = \mathcal{P}^{-1}(f - \mathcal{A}u_m)$$

over the Krylov subspace

$$\mathcal{K}_m(\mathcal{P}^{-1}\mathcal{A}, r_0) := \text{span}\{r_0, (\mathcal{P}^{-1}\mathcal{A})r_0, \dots, (\mathcal{P}^{-1}\mathcal{A})^m r_0\}$$

where the symmetric positive definite (SPD) matrix \mathcal{P} is a preconditioner for \mathcal{A} . Hence, the approximate solution of problem (2.46) follows from solving the minimization problem

$$u_m = \underset{u \in u_0 + \mathcal{K}_m(\mathcal{P}^{-1}\mathcal{A}, \mathcal{P}^{-1}r_0)}{\operatorname{argmin}} \|r_m\|_{\mathcal{P}}$$

by constructing an orthonormal basis for the Krylov subspace using the Lanczos algorithm, where its solution can be calculated by a three-term recurrence relation. The algorithm for the preconditioned MINRES method is presented in Algorithm 1, where the framed steps denote the action of the preconditioner \mathcal{P} .

Data: $\mathcal{A} \in \mathbb{R}^{n \times n}$ regular and symmetric, $\mathcal{P} \in \mathbb{R}^{n \times n}$ SPD, $f \in \mathbb{R}^n$ right-hand side, $u_0 \in \mathbb{R}^n$ initial guess.

Result: approximate solution of (2.46).

Set $r_0 := 0$; Set $w_0 := 0$; Set $w_1 := 0$;

Set $\beta_0 := 1$;

Set $r_1 := f - \mathcal{A}u_0$;

Solve $\mathcal{P}q_1 = r_1$;

Set $\beta_1 := \sqrt{(q_1, r_1)}$;

Set $\eta := \beta_1$;

Set $s_0 := s_1 := 0$; Set $c_0 := c_1 := 1$;

Set $m := 1$;

while *not converged* **do**

Set $q_m := \frac{q_m}{\beta_m}$;

Set $\alpha_m := (\mathcal{A}q_m, q_m)$;

Set $r_{m+1} := \mathcal{A}q_m - \frac{\alpha_m}{\beta_m} r_m - \frac{\beta_m}{\beta_{m-1}} r_{m-1}$;

Solve $\mathcal{P}q_{m+1} = r_{m+1}$;

Set $\beta_{m+1} := \sqrt{(q_{m+1}, r_{m+1})}$;

Set $\gamma_0 := c_m \alpha_m - c_{m-1} s_m \beta_m$;

Set $\gamma_1 := \sqrt{\gamma_0^2 + \beta_{m+1}^2}$;

Set $\gamma_2 := s_m \alpha_m + c_{m-1} c_m \beta_m$;

Set $\gamma_3 := s_{m-1} \beta_m$;

Set $c_{m+1} := \frac{\gamma_0}{\gamma_1}$; Set $s_{m+1} := \frac{\beta_{m+1}}{\gamma_1}$;

Set $w_{m+1} := \frac{1}{\gamma_1} q_m - \frac{\gamma_3}{\gamma_1} w_{m-1} - \frac{\gamma_2}{\gamma_1} w_m$;

Set $u_m := u_{m-1} + c_{m+1} \eta w_{m+1}$;

Set $\eta := -s_{m+1} \eta$;

Set $m := m + 1$;

end

Algorithm 1: Preconditioned minimal residual (preconditioned MINRES) method, cf. [67].

A convergence result for the preconditioned MINRES method can be found in Greenbaum [67]. It states that the convergence rate of the preconditioned MINRES method depends on the condition number of the preconditioned system. This convergence result is summarized in the following theorem in detail.

Theorem 2.25. *The preconditioned MINRES method applied to the system $\mathcal{A}u = f$ with some symmetric and positive definite preconditioner \mathcal{P} , where \mathcal{A} is a regular and symmetric system matrix, converges to the solution of this system for an arbitrary initial guess u_0 . More precisely, the preconditioned residual $r_{2m} = \mathcal{P}^{-1}(f - \mathcal{A}u_{2m})$ after $2m$ iterations can be estimated by the initial residual r_0 as follows*

$$\|r_{2m}\|_{\mathcal{P}} \leq \frac{2q^m}{1 + q^{2m}} \|r_0\|_{\mathcal{P}} \quad \text{with} \quad q = \frac{\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) - 1}{\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) + 1}, \quad (2.48)$$

where $\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) := \|\mathcal{P}^{-1}\mathcal{A}\|_{\mathcal{P}} \|(\mathcal{P}^{-1}\mathcal{A})^{-1}\|_{\mathcal{P}}$ is the condition number of the preconditioned system matrix and $\|\cdot\|_{\mathcal{P}} = (\mathcal{P}\cdot, \cdot)^{1/2}$ is the \mathcal{P} -energy norm.

Proof. Cf. Greenbaum [67]. \square

Hence, the goal is to construct preconditioners for the preconditioned MINRES method such that the condition number $\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A})$ of the preconditioned system $\mathcal{P}^{-1}\mathcal{A}$ is independent of all problem involved parameters and small, more precisely, as close as possible to one.

We start with a result on parameter robust Schur complement preconditioners for saddle point problems of the form (2.46) with the system matrix (2.47).

Theorem 2.26. *Let A and C be symmetric and positive definite matrices and let*

$$S = C + BA^{-1}B^T \quad \text{and} \quad R = A + B^T C^{-1}B$$

be the negative Schur complements. If \mathcal{A} is preconditioned by

$$\mathcal{P}_0 = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \quad \text{or} \quad \mathcal{P}_1 = \begin{pmatrix} R & 0 \\ 0 & C \end{pmatrix}, \quad (2.49)$$

then the eigenvalues of the preconditioned matrices $\mathcal{P}_0^{-1}\mathcal{A}$ and $\mathcal{P}_1^{-1}\mathcal{A}$ are located in the set $(-1, \frac{1-\sqrt{5}}{2}] \cup \{1\} \cup (1, \frac{1+\sqrt{5}}{2}]$.

Proof. See Kuznetsov [107] and Murphy et al. [127]. \square

Theorem 2.26 immediately yields the following norm estimates.

Corollary 2.27. *The inequalities*

$$\underline{c} \|u\|_{\mathcal{P}_0} \leq \|\mathcal{A}u\|_{\mathcal{P}_0^{-1}} \leq \bar{c} \|u\|_{\mathcal{P}_0} \quad \text{and} \quad \underline{c} \|u\|_{\mathcal{P}_1} \leq \|\mathcal{A}u\|_{\mathcal{P}_1^{-1}} \leq \bar{c} \|u\|_{\mathcal{P}_1} \quad (2.50)$$

are valid for all $u \in \mathbb{R}^n$, with $\underline{c} = (\sqrt{5} - 1)/2$ and $\bar{c} = (\sqrt{5} + 1)/2$.

Proof. From Theorem 2.26 follows that the smallest and largest absolute values of the eigenvalues of the preconditioned matrices $\mathcal{P}_0^{-1}\mathcal{A}$ and $\mathcal{P}_1^{-1}\mathcal{A}$ are given by

$$\underline{c} = (\sqrt{5} - 1)/2 \quad \text{and} \quad \bar{c} = (\sqrt{5} + 1)/2.$$

Moreover, we have that

$$\begin{aligned} \frac{\|\mathcal{A}u\|_{\mathcal{P}_j^{-1}}}{\|u\|_{\mathcal{P}_j}} &\leq \sup_{u \neq 0} \frac{\|\mathcal{A}u\|_{\mathcal{P}_j^{-1}}}{\|u\|_{\mathcal{P}_j}} = \sup_{u \neq 0} \frac{(\mathcal{P}_j^{-1}\mathcal{A}u, \mathcal{A}u)^{1/2}}{(\mathcal{P}_j u, u)^{1/2}} = \sup_{u \neq 0} \frac{(\mathcal{P}_j^{-1/2}\mathcal{A}u, \mathcal{P}_j^{-1/2}\mathcal{A}u)^{1/2}}{(\mathcal{P}_j^{1/2}u, \mathcal{P}_j^{1/2}u)^{1/2}} \\ &= \sup_{v \neq 0} \frac{(\mathcal{P}_j^{-1/2}\mathcal{A}\mathcal{P}_j^{-1/2}v, \mathcal{P}_j^{-1/2}\mathcal{A}\mathcal{P}_j^{-1/2}v)^{1/2}}{(v, v)^{1/2}} = \sup_{v \neq 0} \frac{\|\mathcal{P}_j^{-1/2}\mathcal{A}\mathcal{P}_j^{-1/2}v\|}{\|v\|} = \bar{c}, \end{aligned}$$

where we have set $v = \mathcal{P}_j^{1/2}u$ and $j \in \{0, 1\}$. Analogously, we obtain

$$\frac{\|\mathcal{A}u\|_{\mathcal{P}_j^{-1}}}{\|u\|_{\mathcal{P}_j}} \geq \inf_{u \neq 0} \frac{\|\mathcal{A}u\|_{\mathcal{P}_j^{-1}}}{\|u\|_{\mathcal{P}_j}} = \inf_{v \neq 0} \frac{\|\mathcal{P}_j^{-1/2}\mathcal{A}\mathcal{P}_j^{-1/2}v\|}{\|v\|} = \underline{c},$$

which completes the proof. \square

Since

$$\begin{aligned} \|\mathcal{P}_j^{-1}\mathcal{A}\|_{\mathcal{P}_j} &= \sup_{u \neq 0} \frac{\|\mathcal{P}_j^{-1}\mathcal{A}u\|_{\mathcal{P}_j}}{\|u\|_{\mathcal{P}_j}} = \sup_{u \neq 0} \frac{\|\mathcal{A}u\|_{\mathcal{P}_j^{-1}}}{\|u\|_{\mathcal{P}_j}} \leq \bar{c}, \\ \|(\mathcal{P}_j^{-1}\mathcal{A})^{-1}\|_{\mathcal{P}_j} &= \|\mathcal{A}^{-1}\mathcal{P}_j\|_{\mathcal{P}_j} = \sup_{u \neq 0} \frac{\|\mathcal{A}^{-1}\mathcal{P}_j u\|_{\mathcal{P}_j}}{\|u\|_{\mathcal{P}_j}} = \sup_{v \neq 0} \frac{\|v\|_{\mathcal{P}_j}}{\|\mathcal{P}_j^{-1}\mathcal{A}v\|_{\mathcal{P}_j}} = \frac{1}{\inf_{v \neq 0} \frac{\|\mathcal{A}v\|_{\mathcal{P}_j^{-1}}}{\|v\|_{\mathcal{P}_j}}} \leq \frac{1}{\underline{c}}, \end{aligned}$$

the block-diagonal preconditioners yield the condition number estimate

$$\kappa_{\mathcal{P}_j}(\mathcal{P}_j^{-1}\mathcal{A}) = \|\mathcal{P}_j^{-1}\mathcal{A}\|_{\mathcal{P}_j} \|(\mathcal{P}_j^{-1}\mathcal{A})^{-1}\|_{\mathcal{P}_j} \leq \frac{\bar{c}}{\underline{c}} = \frac{\sqrt{5}+1}{\sqrt{5}-1} \approx 2.618, \quad j \in \{0, 1\}.$$

However, in general, the inverse of the Schur complements S and R is hard to be realized in practice. In order to obtain parameter robust convergence rates, we construct block-diagonal preconditioners by the operator interpolation technique presented in Zulehner [187]. The idea is to construct two preconditioners which yield robust convergence rates for the preconditioned MINRES method and apply the operator interpolation theorem, which is based on the construction of intermediate spaces via the so-called real interpolation method. The ideas of the real method (J- and the K-method) are due to Lions and Peetre, see, e.g., [119, 120]. The theory of the real method can also be found, e.g., in Bergh and Löfström [31], see Adams and Fournier [3]. For further information, we refer the reader also to Tartar [166]. Theorem 2.29 presents a matrix version that follows easily from the general operator interpolation theory. A similar notation was used by Zulehner in [187]. This notation goes back to the general theory of matrix means, see, e.g., [35, 113, 144] and was even found before, see [147], in a very different context. In Definition 2.28, we define the geometric mean of two symmetric and positive definite matrices.

Definition 2.28. *Let $A, B \in \mathbb{R}^{n \times n}$ be two symmetric and positive definite matrices. Then the geometric mean of A and B is given by*

$$[A, B]_{1/2} = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}.$$

Moreover, we define, for all $\theta \in [0, 1]$, the symmetric and positive definite matrix $[A, B]_\theta$, i.e.,

$$[A, B]_\theta = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^\theta A^{1/2}.$$

The following theorem presents a finite dimensional matrix version of the operator interpolation theorem and uses the notation presented in Definition 2.28.

Theorem 2.29. *Let $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with*

$$\underline{c}_0 \|u\|_{X_0} \leq \|\mathcal{A}u\|_{Y_0} \leq \bar{c}_0 \|u\|_{X_0} \quad \text{and} \quad \underline{c}_1 \|u\|_{X_1} \leq \|\mathcal{A}u\|_{Y_1} \leq \bar{c}_1 \|u\|_{X_1} \quad \forall u \in \mathbb{R}^n,$$

where the linear vector spaces $X_j = \mathbb{R}^n$ and $Y_j = \mathbb{R}^n$ with $j \in \{0, 1\}$ are equipped with the norms $\|\cdot\|_{X_j}$ and $\|\cdot\|_{Y_j}$ which are associated to the inner products

$$(u, v)_{X_j} = \langle M_j u, v \rangle \quad \text{and} \quad (u, v)_{Y_j} = \langle N_j u, v \rangle$$

given by the symmetric positive definite matrices $M_0, M_1, N_0, N_1 \in \mathbb{R}^{n \times n}$, respectively. Then, for

$$X_\theta = [X_0, X_1]_\theta \quad \text{and} \quad Y_\theta = [Y_0, Y_1]_\theta$$

with $\theta \in [0, 1]$, we have

$$\underline{c}_0^{1-\theta} \underline{c}_1^\theta \|u\|_{X_\theta} \leq \|\mathcal{A}u\|_{Y_\theta} \leq \bar{c}_0^{1-\theta} \bar{c}_1^\theta \|u\|_{X_\theta} \quad \forall u \in \mathbb{R}^n. \quad (2.51)$$

The norms $\|\cdot\|_{X_\theta}$ and $\|\cdot\|_{Y_\theta}$ are the norms associated to the inner products

$$\begin{aligned} (u, v)_{X_\theta} &= \langle M_\theta u, v \rangle \quad \text{with} \quad M_\theta = [M_0, M_1]_\theta = M_0^{1/2} \left(M_0^{-1/2} M_1 M_0^{-1/2} \right)^\theta M_0^{1/2}, \\ (u, v)_{Y_\theta} &= \langle N_\theta u, v \rangle \quad \text{with} \quad N_\theta = [N_0, N_1]_\theta = N_0^{1/2} \left(N_0^{-1/2} N_1 N_0^{-1/2} \right)^\theta N_0^{1/2}. \end{aligned}$$

Proof. The proof follows from the more general version of the space interpolation theorem, as in Adams and Fournier [3], and is based on the general operator interpolation theorem, see also [187]. \square

Hence, from interpolating between the block-diagonal preconditioners \mathcal{P}_0 and \mathcal{P}_1 , we can obtain again parameter independent condition number estimates for all $\theta \in [0, 1]$. We choose $M_0 = \mathcal{P}_0$, $M_1 = \mathcal{P}_1$, $N_0 = \mathcal{P}_0^{-1}$ and $N_1 = \mathcal{P}_1^{-1}$ in Theorem 2.29 and obtain the preconditioners

$$\mathcal{P}_\theta = [P_0, P_1]_\theta = \begin{pmatrix} [A, R]_\theta & 0 \\ 0 & [S, C]_\theta \end{pmatrix}. \quad (2.52)$$

Due to Theorems 2.26 and 2.29, we obtain the estimates

$$\underline{c} \|u\|_{\mathcal{P}_\theta} \leq \|\mathcal{A}u\|_{\mathcal{P}_\theta^{-1}} \leq \bar{c} \|u\|_{\mathcal{P}_\theta} \quad \forall u \in \mathbb{R}^n, \quad (2.53)$$

with the constants $\underline{c} = (\sqrt{5}-1)/2$ and $\bar{c} = (\sqrt{5}+1)/2$. These estimates finally yield a robust estimate of the condition number, i.e.,

$$\kappa_{\mathcal{P}_\theta}(\mathcal{P}_\theta^{-1}\mathcal{A}) \leq \bar{c}/\underline{c} \approx 2.618.$$

The practical implementation of these preconditioners can be done by various methods like (algebraic) multigrid, multilevel or domain decomposition methods, see, e.g., [101, 142, 168, 173]. Since our focus lies on the algebraic multilevel iteration (AMLI) method, we are going to present some basic results regarding this method in the next section.

In this work, we use the special notation $\mathcal{A} \sim \mathcal{B}$ for the spectral equivalence of the matrices \mathcal{A} and \mathcal{B} , which is defined in the following way:

Definition 2.30. *Two symmetric and positive definite matrices \mathcal{A} and \mathcal{B} in $\mathbb{R}^{n \times n}$ are called spectral equivalent, denoted by $\mathcal{A} \sim \mathcal{B}$, if there exist positive constants \underline{c} and \bar{c} which are independent of all involved “bad” parameters such that*

$$\underline{c} u^T \mathcal{A} u \leq u^T \mathcal{B} u \leq \bar{c} u^T \mathcal{A} u \quad \forall u \in \mathbb{R}^n.$$

Of course, we have to specify these “bad” parameters. Beside the discretization parameters, we also have in mind parameters connected with the problem setting, e.g., the regularization or cost parameter in optimal control.

2.8 The AMLI method

Let us consider algebraic systems of linear equations

$$A u = f, \quad (2.54)$$

where $A \in \mathbb{R}^{n \times n}$ is now a sparse, symmetric and positive definite (SPD) matrix. In this final section, we briefly present the algebraic multilevel iteration (AMLI) method in order to solve the linear system (2.54). For more details regarding the AMLI method, we refer the reader to, e.g., [17, 101] and the references therein. We mainly focus on the so-called linear AMLI method, which can be used to define preconditioners for the preconditioned conjugate gradient (PCG) method in order to solve the linear system (2.54). Let us denote by B the preconditioner for the linear system (2.54). The pseudocode of the PCG method is presented in Algorithm 2. Here, we denote by r_m and p_m the residuals and the search directions, respectively.

The linear AMLI method is used to implement the framed preconditioning step in Algorithm 2, i.e.,

$$\text{Solve } B z_m = r_m. \quad (2.55)$$

Its pseudocode is presented in Algorithm 3 at the end of Subsection 2.8.1. Moreover, we present the so-called nonlinear AMLI method in Subsection 2.8.2.

Data: $A \in \mathbb{R}^{n \times n}$ regular SPD, $f \in \mathbb{R}^n$ right-hand side, $u_0 \in \mathbb{R}^n$ initial guess, B preconditioner.

Result: approximate solution of (2.54).

Set $m := 0$;

Set $r_0 := f - Au_0$;

while not converged **do**

Solve $Bz_m = r_m$;

Set $m := m + 1$;

Set $\gamma_{m-1} := (r_{m-1}, z_{m-1})$;

if $m = 1$ **then**

Set $p_m := z_{m-1}$;

else

Set $\beta_m := \frac{\gamma_{m-1}}{\gamma_{m-2}}$;

Set $p_m := z_{m-1} + \beta_m p_{m-1}$;

end

Set $q_m = Ap_m$;

Set $\alpha_m = \frac{\gamma_{m-1}}{(p_m, q_m)}$;

Set $u_m = u_{m-1} + \alpha_m p_m$;

Set $r_m = r_{m-1} - \alpha_m q_m$;

end

Algorithm 2: Preconditioned conjugate gradient (PCG) method, cf. [67, 101].

2.8.1 The linear AMLI method

The classical framework of the so-called linear AMLI method can be found in [14, 15], which is based on a multilevel block factorization and polynomial stabilization.

Let the symmetric and positive definite matrix $A = A^{(L)}$ in (2.54) be obtained in the course of a regular refinement procedure, which defines a sequence of symmetric positive definite matrices starting from a coarsest level system matrix $A^{(0)}$, i.e.,

$$\{A^{(\ell)}\}, \quad A^{(\ell)} \in \mathbb{R}^{n^{(\ell)} \times n^{(\ell)}},$$

where $\ell = 0, \dots, L$, and $n^{(\ell)} > n^{(\ell-1)}$, for $\ell = 1, \dots, L$, see [17]. These matrices are constructed via the finite element method (FEM) for the sequence of nested spaces

$$V^{(0)} \subset V^{(1)} \subset \dots \subset V^{(\ell)} \subset \dots \subset V^{(L)} = V_h, \quad (2.56)$$

corresponding to nested meshes $\mathcal{T}^{(\ell)}$ for $\ell = 0, \dots, L$, where $\mathcal{T}^{(L)} = \mathcal{T}_h$ is the finest mesh. The spaces

$$V^{(\ell)} = \text{span}\{\varphi_1^{(\ell)}, \dots, \varphi_{n^{(\ell)}}^{(\ell)}\}$$

are finite element spaces spanned by the standard nodal basis functions

$$\{\varphi_i^{(\ell)} : i = 1, \dots, n^{(\ell)}\},$$

where we use continuous, piecewise linear conforming finite elements on triangles on a regular triangulation to construct the finite element spaces and their bases, see [41, 46, 84, 161] and Section 2.4. On each level ℓ , we partition the matrix $A^{(\ell)}$ in a two-by-two block form, i.e.,

$$A^{(\ell)} = \begin{pmatrix} A_{11}^{(\ell)} & A_{12}^{(\ell)} \\ A_{21}^{(\ell)} & A_{22}^{(\ell)} \end{pmatrix} \begin{matrix} \} n^{(\ell)} - n^{(\ell-1)} \\ \} n^{(\ell-1)} \end{matrix}, \quad (2.57)$$

where the blocks $A_{22}^{(\ell)}$ and $A_{11}^{(\ell)}$ correspond to the unknowns that are associated with the coarser mesh $\mathcal{T}^{(\ell-1)}$ and to the unknowns that are added in the course of refining the mesh $\mathcal{T}^{(\ell-1)}$ resulting in the mesh $\mathcal{T}^{(\ell)}$, respectively. The Schur complements

$$S^{(\ell)} = A_{22}^{(\ell)} - A_{21}^{(\ell)} (A_{11}^{(\ell)})^{-1} A_{12}^{(\ell)} \quad (2.58)$$

are dense symmetric and positive definite matrices. In the course of designing optimal multilevel methods, it is important to construct a sparse approximation of $S^{(\ell)}$, see [101]. More precisely, $S^{(\ell)}$ has to be spectrally equivalent to $A^{(\ell-1)}$ on all levels $\ell = 1, \dots, L$ with spectral equivalence bounds that neither depend on the level index $\ell - 1$ nor on any problem parameters, see [17].

The efficiency of preconditioners based on two-by-two block factorization strongly depends on the coupling of the diagonal blocks of the two-level matrix via its off-diagonal blocks. A measure for the strength of this coupling is the constant γ in the *strengthened Cauchy-Bunyakowski-Schwarz (CBS)* inequality

$$\left| v_1^T A_{12}^{(\ell)} v_2 \right| \leq \gamma \left(v_1^T A_{11}^{(\ell)} v_1 \right)^{1/2} \left(v_2^T A_{22}^{(\ell)} v_2 \right)^{1/2} \quad \forall v_1 \in V_1^{(\ell)} \quad \forall v_2 \in V_2^{(\ell)}, \quad (2.59)$$

where $V_1^{(\ell)}$ and $V_2^{(\ell)}$ form a splitting of the vector space $V^{(\ell)}$, which is consistent with the partitioning (2.57), cf. [54, 101]. The strengthened CBS inequality refines the usual one, cf. (2.1), by stating the existence of a constant $\gamma \leq 1$. The so-called *CBS constant*, i.e., the smallest γ for which (2.59) holds, maybe called the cosine of the angle between the spaces $V_1^{(\ell)}$ and $V_2^{(\ell)}$ and can be estimated locally, see, e.g., [115] and the references therein. Let the elements of $\mathcal{T}^{(\ell)}$ be uniform refinements of the coarse-grid elements $e \in \mathcal{T}^{(\ell-1)}$. In the following, we call $E \subset \mathcal{T}^{(\ell)}$ a macro element. The global CBS constant can be estimated by the maximum of the local CBS constants on the macro elements $E \subset \mathcal{T}^{(\ell)}$, i.e.,

$$\gamma \leq \max_{E \subset \mathcal{T}^{(\ell)}} \gamma_E \leq 1, \quad (2.60)$$

and can be computed via a simple rule, i.e.,

$$\gamma_E^2 = 1 - \lambda_E^{\min}, \quad (2.61)$$

where λ_E^{\min} is the minimal eigenvalue of the generalized eigenvalue problem

$$S_E v_{E:2} = \lambda A_e v_{E:2}, \quad (2.62)$$

and $v_{E:2} \neq (c, c, \dots, c)^T$, c is a real constant, see, e.g., [101]. The global matrices $A^{(\ell)}$ and $A^{(\ell-1)}$ can be computed via the local matrices A_E and A_e , respectively. The standard FEM assembling can be written in the form

$$\begin{aligned} A^{(\ell)} &= \sum_{E \subset \mathcal{T}^{(\ell)}} R_E^T A_E R_E, \\ A^{(\ell-1)} &= \sum_{e \in \mathcal{T}^{(\ell-1)}} R_e^T A_e R_e, \end{aligned} \quad (2.63)$$

where R_E and R_e are the restriction mappings of a global vector of unknowns at levels ℓ and $\ell - 1$ to the local vectors corresponding to the elements $E \subset \mathcal{T}^{(\ell)}$ and $e \in \mathcal{T}^{(\ell-1)}$, respectively, cf. [101]. Hence, it suffices to consider the local matrices A_E and A_e for analyzing the robustness and optimal complexity of the linear AMLI method for solving problem (2.54).

It is well known that ensuring that the CBS constant γ in (2.59) and (2.60) is strictly less than 1 in general requires a change of basis. Let us consider the two nested finite element spaces

$$V^{(\ell-1)} \subset V^{(\ell)},$$

which correspond to the two consecutive meshes $\mathcal{T}^{(\ell-1)}$ and $\mathcal{T}^{(\ell)}$, respectively. Their standard finite element nodal basis functions are given by

$$\{\varphi_i^{(\ell-1)} : i = 1, \dots, n^{(\ell-1)}\} \quad \text{and} \quad \{\varphi_i^{(\ell)} : i = 1, \dots, n^{(\ell)}\}.$$

We split the $n^{(\ell)}$ meshpoints into the group containing the $n^{(\ell-1)}$ nodes of the coarse mesh $\mathcal{T}^{(\ell-1)}$ and the rest. Then by defining the hierarchical basis functions

$$\{\tilde{\varphi}_i^{(\ell)} : i = 1, \dots, n^{(\ell)}\},$$

the hierarchical matrix $\tilde{A}^{(\ell)}$ as well as $A^{(\ell)}$ (for the latter see (2.57)) are naturally partitioned in a two-by-two block form, i.e.,

$$\tilde{A}^{(\ell)} = \begin{pmatrix} \tilde{A}_{11}^{(\ell)} & \tilde{A}_{12}^{(\ell)} \\ \tilde{A}_{21}^{(\ell)} & \tilde{A}_{22}^{(\ell)} \end{pmatrix} \begin{matrix} \} n^{(\ell)} - n^{(\ell-1)} \\ \} n^{(\ell-1)} \end{matrix},$$

see [101]. The hierarchical matrix $\tilde{A}^{(\ell)}$ is more dense than $A^{(\ell)}$, however, the related CBS constant γ is typically less than 1. The nodal unknown vectors for the standard and for the hierarchical basis functions are related by a transformation matrix of the form

$$J^{(\ell)} = \begin{pmatrix} I & J_{12}^{(\ell)} \\ 0 & I \end{pmatrix}, \quad (2.64)$$

where I is the corresponding identity matrix and 0 the zero matrix. In practical applications, we can work with $A^{(\ell)}$ instead of $\tilde{A}^{(\ell)}$, since

$$\tilde{A}^{(\ell)} = (J^{(\ell)})^T A^{(\ell)} J^{(\ell)}. \quad (2.65)$$

Lemma 2.31 (cf. [54, 101]). *The transformation from the standard to the hierarchical basis does not change the Schur complement, i.e.,*

$$S^{(\ell)} = \tilde{S}^{(\ell)}.$$

Proof. This follows from a straightforward computation:

$$\begin{aligned} \tilde{A}^{(\ell)} &= \begin{pmatrix} \tilde{A}_{11}^{(\ell)} & \tilde{A}_{12}^{(\ell)} \\ \tilde{A}_{21}^{(\ell)} & \tilde{A}_{22}^{(\ell)} \end{pmatrix} \\ &= (J^{(\ell)})^T A^{(\ell)} J^{(\ell)} \\ &= \begin{pmatrix} I & 0 \\ (J_{12}^{(\ell)})^T & I \end{pmatrix} \begin{pmatrix} A_{11}^{(\ell)} & A_{12}^{(\ell)} \\ A_{21}^{(\ell)} & A_{22}^{(\ell)} \end{pmatrix} \begin{pmatrix} I & J_{12}^{(\ell)} \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} A_{11}^{(\ell)} & A_{12}^{(\ell)} \\ (J_{12}^{(\ell)})^T A_{11}^{(\ell)} + A_{21}^{(\ell)} & (J_{12}^{(\ell)})^T A_{12}^{(\ell)} + A_{22}^{(\ell)} \end{pmatrix} \begin{pmatrix} I & J_{12}^{(\ell)} \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} A_{11}^{(\ell)} & A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{12}^{(\ell)} \\ (J_{12}^{(\ell)})^T A_{11}^{(\ell)} + A_{21}^{(\ell)} & (J_{12}^{(\ell)})^T A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{21}^{(\ell)} J_{12}^{(\ell)} + (J_{12}^{(\ell)})^T A_{12}^{(\ell)} + A_{22}^{(\ell)} \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned}
\tilde{S}^{(\ell)} &= \tilde{A}_{22}^{(\ell)} - \tilde{A}_{21}^{(\ell)} (\tilde{A}_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} \\
&= (J_{12}^{(\ell)})^T A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{21}^{(\ell)} J_{12}^{(\ell)} + (J_{12}^{(\ell)})^T A_{12}^{(\ell)} + A_{22}^{(\ell)} \\
&\quad - \left((J_{12}^{(\ell)})^T A_{11}^{(\ell)} + A_{21}^{(\ell)} \right) (A_{11}^{(\ell)})^{-1} \left(A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{12}^{(\ell)} \right) \\
&= (J_{12}^{(\ell)})^T A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{21}^{(\ell)} J_{12}^{(\ell)} + (J_{12}^{(\ell)})^T A_{12}^{(\ell)} + A_{22}^{(\ell)} \\
&\quad - \left((J_{12}^{(\ell)})^T + A_{21}^{(\ell)} (A_{11}^{(\ell)})^{-1} \right) \left(A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{12}^{(\ell)} \right) \\
&= (J_{12}^{(\ell)})^T A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{21}^{(\ell)} J_{12}^{(\ell)} + (J_{12}^{(\ell)})^T A_{12}^{(\ell)} + A_{22}^{(\ell)} \\
&\quad - \left((J_{12}^{(\ell)})^T A_{11}^{(\ell)} J_{12}^{(\ell)} + A_{21}^{(\ell)} J_{12}^{(\ell)} + (J_{12}^{(\ell)})^T A_{12}^{(\ell)} + A_{21}^{(\ell)} (A_{11}^{(\ell)})^{-1} A_{12}^{(\ell)} \right) \\
&= A_{22}^{(\ell)} - A_{21}^{(\ell)} (A_{11}^{(\ell)})^{-1} A_{12}^{(\ell)} \\
&= S^{(\ell)}.
\end{aligned}$$

□

Remark 2.32. *Note that*

$$\tilde{A}_{11}^{(\ell)} = A_{11}^{(\ell)} \quad \text{and} \quad \tilde{A}_{22}^{(\ell)} = A^{(\ell-1)}.$$

Moreover, we can compute the minimal eigenvalue of the generalized eigenvalue problem (2.62) using the Schur complement S_E , from which we obtain the local CBS constant γ_E via the rule (2.61).

Let us consider the following complete factorization of $\tilde{A}^{(\ell)}$:

$$\tilde{A}^{(\ell)} = \begin{pmatrix} A_{11}^{(\ell)} & 0 \\ \tilde{A}_{21}^{(\ell)} & S^{(\ell)} \end{pmatrix} \begin{pmatrix} I & (A_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} \\ 0 & I \end{pmatrix}, \quad (2.66)$$

see [14]. There are also linear AMLI versions using complete factorizations of the form

$$\tilde{A}^{(\ell)} = \begin{pmatrix} I & 0 \\ \tilde{A}_{21}^{(\ell)} (A_{11}^{(\ell)})^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11}^{(\ell)} & 0 \\ 0 & S^{(\ell)} \end{pmatrix} \begin{pmatrix} I & (A_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} \\ 0 & I \end{pmatrix},$$

see [101]. In order to construct uniform AMLI preconditioners $B^{(\ell)}$ for the matrices $A^{(\ell)}$ with an asymptotically optimal order condition number, i.e.,

$$\kappa((B^{(\ell)})^{-1} A^{(\ell)}) = \mathcal{O}(1),$$

and whose application to a vector has optimal computational complexity, we combine hierarchical basis preconditioners with polynomial stabilization techniques. Here, the symbol \mathcal{O} denotes the big O notation and is a member of the so-called Landau notation, also called the Bachmann-Landau or asymptotic notation, where $\mathcal{O}(1)$ means constant.

Remember that $A^{(\ell)}$ and $\tilde{A}^{(\ell)}$ are related via (2.65). Here, we mainly present the so-called multiplicative form of the AMLI method in the spirit of [14, 15], cf. [101]. Let us start at the coarsest level on which a complete factorization of the matrix $A^{(0)}$ is performed. We define

$$B^{(0)} := A^{(0)}.$$

Then the preconditioner $B^{(\ell)}$ at level ℓ according to (2.66) is defined by

$$B^{(\ell)} := \begin{pmatrix} A_{11}^{(\ell)} & 0 \\ \tilde{A}_{21}^{(\ell)} & Z^{(\ell-1)} \end{pmatrix} \begin{pmatrix} I & (A_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} \\ 0 & I \end{pmatrix}, \quad (2.67)$$

where $Z^{(\ell-1)}$ denotes the Schur complement approximation. More precisely, we use the following approximation of the inverse of the Schur complements:

$$(Z^{(\ell-1)})^{-1} := \left(I - P^{(\ell)} \left((B^{(\ell-1)})^{-1} A^{(\ell-1)} \right) \right) (A^{(\ell-1)})^{-1}. \quad (2.68)$$

Here, $P^{(\ell)}(t) = P_{v_\ell}(t)$, is a *polynomial of degree*

$$1 \leq v_\ell \leq v := \max_{0 \leq \ell \leq L} v_\ell,$$

which has to satisfy the conditions

$$0 \leq P^{(\ell)}(t) < 1, \quad \text{for } 0 < t \leq 1, \quad \text{and} \quad P^{(\ell)}(0) = 1,$$

for all $\ell = 1, \dots, L$. The Schur complement approximation (2.68) is equivalent to

$$(Z^{(\ell-1)})^{-1} := (B^{(\ell-1)})^{-1} Q^{(\ell)} \left(A^{(\ell-1)} (B^{(\ell-1)})^{-1} \right) \quad (2.69)$$

with

$$Q^{(\ell)}(t) = \frac{1 - P^{(\ell)}(t)}{t},$$

see [101]. The polynomial $Q^{(\ell)}(t) = Q_{v_\ell-1}(t)$ is of degree $v_\ell - 1$ and can also be written as

$$Q^{(\ell)}(t) = q_0^{(\ell)} + q_1^{(\ell)} t + \dots + q_{v_\ell-1}^{(\ell)} t^{v_\ell-1}.$$

Remark 2.33. *Using the polynomial degree $v_\ell = 1$ at all intermediate levels and choosing*

$$P^{(\ell)}(t) = P_1(t) = 1 - t,$$

we obtain the preconditioner

$$B^{(\ell)} := \begin{pmatrix} A_{11}^{(\ell)} & 0 \\ \tilde{A}_{21}^{(\ell)} & B^{(\ell-1)} \end{pmatrix} \begin{pmatrix} I & (A_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} \\ 0 & I \end{pmatrix},$$

which corresponds to the so-called (linear) AMLI V-cycle method, see [14].

More details regarding a proper choice of stabilization polynomials are provided in Chapter 5 including condition number bounds for

$$\kappa((B^{(\ell)})^{-1} A^{(\ell)})$$

following the results and ideas in [11, 14, 15]. Moreover, one can replace $A_{11}^{(\ell)}$ in the definition of the AMLI preconditioner (2.67) by an approximation $C_{11}^{(\ell)}$, i.e.,

$$C_{11}^{(\ell)} \approx A_{11}^{(\ell)},$$

by which we mean that

$$\underline{c} v_1^T C_{11}^{(\ell)} v_1 \leq v_1^T A_{11}^{(\ell)} v_1 \leq \bar{c} v_1^T C_{11}^{(\ell)} v_1, \quad (2.70)$$

with some positive constants \underline{c} and \bar{c} . We present an additive preconditioner $C_{11}^{(\ell)}$ for $A_{11}^{(\ell)}$ in the context of heterogeneous reaction-diffusion type problems in Chapter 5.

The *optimality conditions* for the multiplicative variant (2.67) of the AMLI method are given by

$$\frac{1}{\sqrt{1-\gamma^2}} < v < \varrho. \quad (2.71)$$

If (2.71) are fulfilled, then the AMLI preconditioner $B^{(\ell)}$ defined in (2.67) is spectrally equivalent to $A^{(\ell)}$. Here, the parameter ϱ stands for the *refinement factor*. More precisely, in case of an m -refinement we have $\varrho = m^2$, which means that we subdivide one element into m^2 congruent elements in one refinement step. More results concerning the refinement factor can be found in Chapter 5.

Remark 2.34. For the additive variant of the AMLI method, the optimality conditions read as follows

$$\sqrt{\frac{1+\gamma}{1-\gamma}} < v < \varrho.$$

In Algorithm 3, we provide the pseudocode for the linear AMLI method in order to implement step (2.55), rewritten as

$$B^{(\ell)} z^{(\ell)} = r^{(\ell)} \quad (2.72)$$

at a level $\ell \in \{1, \dots, L\}$, using the PCG method for solving problem (2.54), see [15]. Here, $B^{(\ell)}$ denotes the algebraic multilevel preconditioner defined by (2.67) using the approximation $C_{11}^{(\ell)}$ for the pivot block $A_{11}^{(\ell)}$ according to (2.70), i.e.,

$$B^{(\ell)} := \begin{pmatrix} C_{11}^{(\ell)} & 0 \\ \tilde{A}_{21}^{(\ell)} & Z^{(\ell-1)} \end{pmatrix} \begin{pmatrix} I & (C_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} \\ 0 & I \end{pmatrix}, \quad (2.73)$$

and $r^{(\ell)}$ is some given right-hand side. We mention that $r^{(\ell)}$ and $z^{(\ell)}$ can be represented as

$$r^{(\ell)} = \begin{pmatrix} r_1^{(\ell)} \\ r_2^{(\ell)} \end{pmatrix} \quad \text{and} \quad z^{(\ell)} = \begin{pmatrix} z_1^{(\ell)} \\ z_2^{(\ell)} \end{pmatrix},$$

respectively, due to our space splitting.

Remark 2.35. The AMLI preconditioner (2.73) using the approximation $C_{11}^{(\ell)}$ for the pivot block of $A^{(\ell)}$ has to fulfill the same optimality conditions (2.71) as (2.67). The AMLI preconditioner (2.73) is spectrally equivalent to $A^{(\ell)}$ if

$$\frac{1}{\sqrt{1-\gamma^2}} < v,$$

where γ is the CBS constant and v the degree of the stabilization polynomial, see [15].

2.8.2 The nonlinear AMLI method

The linear AMLI method depends on a proper choice of the stabilization polynomials which are used to construct the approximations of the inverse of the Schur complements. Variable-step AMLI methods that result in nonlinear preconditioners have been introduced in [16] and further analyzed in [98, 133]. In contrast to the linear AMLI method, the stabilization in the nonlinear AMLI method is achieved by performing a few inner iterations of a flexible Krylov subspace method on each or on certain levels ℓ of the multilevel cycle. Hence, the nonlinear AMLI algorithm is parameter-free and uses inner iterations in order to implement the approximations of the inverse of the Schur complements, see [101]. The nonlinear AMLI methods have also been combined with additive Schur complement approximations to obtain fully parameter-robust preconditioners for elliptic problems with highly varying coefficients [99], and problems with a highly anisotropic diffusion tensor [100]. Although the nonlinear AMLI methods have considerable advantages from a practical point of view, the focus of Chapter 5 is on the construction of optimal *linear* AMLI methods for reaction-diffusion type problems in the (classical) setting of hierarchical bases as presented in Subsection 2.8.2 and, most important, we prove the robustness and optimal complexity of our linear AMLI method. Moreover, we present numerical results using the linear and nonlinear AMLI methods in Chapter 7.

Algorithm 4 presents the pseudocode of the nonlinear AMLI method for solving the linear system (2.54), i.e.,

$$A^{(L)} u^{(L)} = f^{(L)}.$$

Data: $r^{(L)} \in \mathbb{R}^n$ right-hand side.
Result: solution $z^{(L)} \in \mathbb{R}^n$ of (2.72).
for $1 \leq \ell \leq L$ **do**
 | Set $\sigma_\ell := 0$;
end
Set $\ell := L$;
forward:
Set $\sigma_\ell := \sigma_\ell + 1$;
if $\sigma_\ell = 1$ **then**
 | Set $z^{(\ell)} := 0$;
 | Set $w^{(\ell)} := q_{v_\ell-1} r^{(\ell)}$;
else
 | Set $w^{(\ell)} := q_{v_\ell-\sigma_\ell} r^{(\ell)} + A^{(\ell)} z^{(\ell)}$;
end
Set $z_1^{(\ell)} := (C_{11}^{(\ell)})^{-1} w_1^{(\ell)}$;
Set $r^{(\ell-1)} := w_2^{(\ell)} - \tilde{A}_{21}^{(\ell)} z_1^{(\ell)}$;
Set $r^{(\ell-1)} := (J^{(\ell-1)})^T r^{(\ell-1)}$;
Set $\ell := \ell - 1$;
if $\ell > 0$ **then**
 | **goto forward**
end
Solve $A^{(0)} z^{(0)} = r^{(0)}$;
backward:
Set $\ell := \ell + 1$;
Set $z_2^{(\ell)} := z^{(\ell-1)}$;
Set $z_1^{(\ell)} := z_1^{(\ell)} - (C_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} z_2^{(\ell)}$;
Set $z^{(\ell)} := J^{(\ell)} z^{(\ell)}$;
if $\sigma_\ell < v_\ell$ **then**
 | **goto forward**
end
Set $\sigma_\ell := 0$;
if $\ell < L$ **then**
 | **goto backward**
end

Algorithm 3: Linear algebraic multilevel iteration (linear AMLI) method, cf. [15, 101].

We denote by $f^{(\ell)}$, $r^{(\ell)}$ and $C_{11}^{(\ell)}$ the current right-hand side, residual and the preconditioner for $A_{11}^{(\ell)}$ at level ℓ , respectively. Moreover, v_ℓ and σ_ℓ now denote the number of recursive calls and the counter for the number of visits at level ℓ , respectively, and

$$p_{(j)}^{(\ell)} = \begin{pmatrix} p_{1,(j)}^{(\ell)} \\ p_{2,(j)}^{(\ell)} \end{pmatrix}$$

is the j -th search direction at level ℓ for $1 \leq j \leq \sigma_\ell$. Here, we implement the algorithm for applying the two-level hierarchical basis transformation (2.65) at level L . More details regarding the nonlinear AMLI method can be found in [101].

Data: $A^{(L)} \in \mathbb{R}^{n \times n}$ regular SPD, $f^{(L)} \in \mathbb{R}^n$ right-hand side.

Result: approximation of the solution $u^{(L)} \in \mathbb{R}^n$ of (2.54).

for $1 \leq \ell \leq L$ **do**

 | Set $\sigma_\ell := 0$; Set $u^{(\ell)} := 0$;

end

Set $\ell := L$; Set $f^{(L)} := (J^{(L)})^T f^{(L)}$; Set $r^{(L)} := f^{(L)}$;

while not converged do

forward:

 Set $\sigma_\ell := \sigma_\ell + 1$;

if $\sigma_\ell = 1$ && $\ell < L$ **then**

 | Set $u^{(\ell)} := 0$; Set $r^{(\ell)} := f^{(\ell)}$;

end

 Set $p_{1(\sigma_\ell)}^{(\ell)} := (C_{11}^{(\ell)})^{-1} r_1^{(\ell)}$; Set $f^{(\ell-1)} := r_2^{(\ell)} - \tilde{A}_{21}^{(\ell)} p_{1(\sigma_\ell)}^{(\ell)}$;

 Set $\ell := \ell - 1$;

if $\ell > 0$ **then**

if $\sigma_\ell = 0$ **then**

 | $f^{(\ell)} := (J^{(\ell)})^T f^{(\ell)}$;

end

goto forward

end

 Solve $A^{(0)} u^{(0)} = f^{(0)}$;

backward:

 Set $\sigma_\ell := 0$; Set $\ell := \ell + 1$; Set $p_{2(\sigma_\ell)}^{(\ell)} := u^{(\ell-1)}$;

 Set $p_{1(\sigma_\ell)}^{(\ell)} := p_{1(\sigma_\ell)}^{(\ell)} - (C_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} p_{2(\sigma_\ell)}^{(\ell)}$;

if $v_\ell = 1$ **then**

 | Set $u^{(\ell)} := p_{(\sigma_\ell)}^{(\ell)}$;

else

 Set $q_{(\sigma_\ell)}^{(\ell)} := \tilde{A}^{(\ell)} p_{(\sigma_\ell)}^{(\ell)}$; Set $\gamma_{(\sigma_\ell)}^{(\ell)} := (q_{(\sigma_\ell)}^{(\ell)}, p_{(\sigma_\ell)}^{(\ell)})$;

for $1 \leq j \leq \sigma_\ell - 1$ **do**

 | Set $\beta := \frac{(q_{(\sigma_\ell)}^{(\ell)}, p_{(j)}^{(\ell)})}{\gamma_{(j)}^{(\ell)}}$; Set $p_{(\sigma_\ell)}^{(\ell)} := p_{(\sigma_\ell)}^{(\ell)} - \beta p_{(j)}^{(\ell)}$; Set $q_{(\sigma_\ell)}^{(\ell)} := q_{(\sigma_\ell)}^{(\ell)} - \beta q_{(j)}^{(\ell)}$;

end

 Set $\alpha := \frac{(r^{(\ell)}, p_{(\sigma_\ell)}^{(\ell)})}{\gamma_{(\sigma_\ell)}^{(\ell)}}$; Set $u^{(\ell)} := u^{(\ell)} + \alpha p_{(\sigma_\ell)}^{(\ell)}$; Set $r^{(\ell)} := r^{(\ell)} - \alpha q_{(\sigma_\ell)}^{(\ell)}$;

end

if $\sigma_\ell < v_\ell$ && $\ell < L$ **then**

 | **goto forward**

end

if $\ell < L$ **then**

 | Set $u^{(\ell)} := J^{(\ell)} u^{(\ell)}$;

 | **goto backward**

end

if $\sigma_L = v_L$ **then**

 | Set $\sigma_L := 0$;

end

end

Set $u^{(L)} := J^{(L)} u^{(L)}$;

Algorithm 4: Nonlinear algebraic multilevel iteration (nonlinear AMLI) method, cf. [101].

Chapter 3

Multiharmonic finite element analysis of parabolic time-periodic boundary value problems

3.1 A parabolic time-periodic boundary value problem

Let $Q_T := \Omega \times (0, T)$ denote the space-time cylinder and $\Sigma_T := \Gamma \times (0, T)$ its mantle boundary, where $\Omega \subset \mathbb{R}^d$, $d = \{1, 2, 3\}$, is a bounded Lipschitz domain, and $(0, T)$ is a given time interval. We consider the parabolic time-periodic boundary value problem (2.33), i.e.,

$$\sigma(\mathbf{x}) \partial_t u(\mathbf{x}, t) - \operatorname{div}(\nu(\mathbf{x}) \nabla u(\mathbf{x}, t)) = f(\mathbf{x}, t) \quad (\mathbf{x}, t) \in Q_T, \quad (3.1)$$

$$u(\mathbf{x}, t) = 0 \quad (\mathbf{x}, t) \in \Sigma_T, \quad (3.2)$$

$$u(\mathbf{x}, 0) = u(\mathbf{x}, T) \quad \mathbf{x} \in \bar{\Omega}, \quad (3.3)$$

where $f(\mathbf{x}, t)$ is some given data, and $\sigma(\cdot)$ and $\nu(\cdot)$ satisfy the assumptions (2.29), i.e.,

$$0 < \underline{\sigma} \leq \sigma(\mathbf{x}) \leq \bar{\sigma}, \quad 0 < \underline{\nu} \leq \nu(\mathbf{x}) \leq \bar{\nu}, \quad \mathbf{x} \in \Omega.$$

In order to study the parabolic time-periodic problem (3.1)-(3.3), we will use an approach inspired by Ladyzhenskaya et al., see [108, 109]. Hence, we consider the Sobolev spaces $H^{1,0}(Q_T)$ and $H^{1,1}(Q_T)$ defined in (2.2) and (2.4), respectively, which are equipped with the norms (2.3) and (2.5), respectively. One could imagine that the space variables $\mathbf{x} = (x^1, \dots, x^d)$ and the time variable t form a $d + 1$ dimensional “domain” Q_T and so t can be treated simply as an additional variable of the space-time cylinder Q_T . This approach provides the clue to require that the time derivative is also from the space $L^2(Q_T)$. Moreover, we consider the Sobolev spaces

$$H_0^{1,0}(Q_T) = \{u \in H^{1,0}(Q_T) : u = 0 \text{ on } \Sigma_T\},$$

and

$$H_{0,per}^{1,1}(Q_T) = \{u \in H_0^{1,1}(Q_T) : u(\mathbf{x}, 0) = u(\mathbf{x}, T) \text{ for almost all } \mathbf{x} \in \Omega\},$$

which include the boundary and time-periodicity conditions. In order to derive the space-time variational formulation of the parabolic time-periodic problem (3.1)-(3.3), we multiply the parabolic partial differential equation (3.1) by a test function $v \in H_{0,per}^{1,1}(Q_T)$, integrate over the space-time cylinder Q_T , and after integration by parts with respect to the space variables, we obtain the following space-time variational formulation of the parabolic time-periodic problem (3.1)-(3.3): Given the

right-hand side $f \in L^2(Q_T)$, find $u \in H_{0,per}^{1,1}(Q_T)$ such that

$$\int_0^T \int_{\Omega} (\sigma(\mathbf{x}) \partial_t u(\mathbf{x}, t) v(\mathbf{x}, t) + \nu(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t)) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad (3.4)$$

for all test functions $v \in H_{0,per}^{1,1}(Q_T)$. Moreover, we can formulate a second space-time variational formulation which is weaker than the space-time variational problem (3.4) by applying, in addition, integration by parts with respect to the time variable. Since the test functions are time-periodic, i.e., $v(\mathbf{x}, 0) = v(\mathbf{x}, T)$, we obtain that

$$\begin{aligned} \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) \, d\mathbf{x} \, dt &= \int_0^T \int_{\Omega} (\sigma(\mathbf{x}) \partial_t u(\mathbf{x}, t) v(\mathbf{x}, t) + \nu(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t)) \, d\mathbf{x} \, dt \\ &= \int_{\Omega} (\sigma(\mathbf{x}) u(\mathbf{x}, T) v(\mathbf{x}, T) - \sigma(\mathbf{x}) u(\mathbf{x}, 0) v(\mathbf{x}, 0)) \, d\mathbf{x} - \int_0^T \int_{\Omega} \sigma(\mathbf{x}) u(\mathbf{x}, t) \partial_t v(\mathbf{x}, t) \, d\mathbf{x} \, dt \\ &\quad + \int_0^T \int_{\Omega} \nu(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t) \, d\mathbf{x} \, dt \\ &= \int_{\Omega} \sigma(\mathbf{x}) (u(\mathbf{x}, T) - u(\mathbf{x}, 0)) v(\mathbf{x}, 0) \, d\mathbf{x} - \int_0^T \int_{\Omega} \sigma(\mathbf{x}) u(\mathbf{x}, t) \partial_t v(\mathbf{x}, t) \, d\mathbf{x} \, dt \\ &\quad + \int_0^T \int_{\Omega} \nu(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t) \, d\mathbf{x} \, dt. \end{aligned}$$

Hence, we see that the time-periodicity condition $u(\mathbf{x}, 0) = u(\mathbf{x}, T)$ of the solution can be incorporated in a weak sense. Thus, we arrive at the following variational formulation: Given $f \in L^2(Q_T)$, find $u \in H_0^{1,0}(Q_T)$ such that

$$\int_0^T \int_{\Omega} (-\sigma(\mathbf{x}) u(\mathbf{x}, t) \partial_t v(\mathbf{x}, t) + \nu(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t)) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad (3.5)$$

for all $v \in H_{0,per}^{1,1}(Q_T)$, and the periodicity condition is incorporated in a weak sense.

Note that the space-time variational problem (3.5) is the time-periodic analogon to the variational formulation (2.31) for initial-boundary value problems, where $\sigma(\cdot)$ and $\nu(\cdot)$ are not only constant but more general, i.e., fulfill the assumptions (2.29).

Since all functions are at least from $L^2(Q_T)$, we can expand the functions u , v and f into Fourier series in time. In particular, this approach makes sense due to the time-periodicity condition (for u and v). The Fourier series expansion in time (2.10), e.g., for v , is given by

$$v(\mathbf{x}, t) = v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [v_k^c(\mathbf{x}) \cos(k\omega t) + v_k^s(\mathbf{x}) \sin(k\omega t)]$$

with the Fourier coefficients

$$\begin{aligned} v_0^c(\mathbf{x}) &= \frac{1}{T} \int_0^T v(\mathbf{x}, t) \, dt, \\ v_k^c(\mathbf{x}) &= \frac{2}{T} \int_0^T v(\mathbf{x}, t) \cos(k\omega t) \, dt, \\ v_k^s(\mathbf{x}) &= \frac{2}{T} \int_0^T v(\mathbf{x}, t) \sin(k\omega t) \, dt, \end{aligned}$$

where T and $\omega = 2\pi/T$ denote the periodicity and the frequency, respectively. Then, the time derivative of v is given by

$$\partial_t v(\mathbf{x}, t) = \sum_{k=1}^{\infty} [k\omega v_k^s(\mathbf{x}) \cos(k\omega t) - k\omega v_k^c(\mathbf{x}) \sin(k\omega t)].$$

In the following, we will use the notation

$$\mathbf{v}_k = (v_k^c, v_k^s)^T, \quad \mathbf{v}_k^\perp = (-v_k^s, v_k^c)^T \quad \text{and} \quad \nabla \mathbf{v}_k = ((\nabla v_k^c)^T, (\nabla v_k^s)^T)^T,$$

and the definition

$$\begin{aligned} v^\perp(\mathbf{x}, t) &:= -v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [-v_k^c(\mathbf{x}) \sin(k\omega t) + v_k^s(\mathbf{x}) \cos(k\omega t)] \\ &= -v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} \underbrace{(v_k^s(\mathbf{x}), -v_k^c(\mathbf{x}))}_{=(-\mathbf{v}_k^\perp)^T} \cdot \begin{pmatrix} \cos(k\omega t) \\ \sin(k\omega t) \end{pmatrix} \\ &= -v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} (v_k^c(\mathbf{x}), v_k^s(\mathbf{x})) \cdot \begin{pmatrix} -\sin(k\omega t) \\ \cos(k\omega t) \end{pmatrix}. \end{aligned}$$

Note that

$$\begin{aligned} (v^\perp)^\perp(\mathbf{x}, t) &= -v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} (v_k^c(\mathbf{x}), v_k^s(\mathbf{x})) \cdot \begin{pmatrix} -\cos(k\omega t) \\ -\sin(k\omega t) \end{pmatrix} \\ &= -v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [-v_k^c(\mathbf{x}) \cos(k\omega t) - v_k^s(\mathbf{x}) \sin(k\omega t)] = -v(\mathbf{x}, t). \end{aligned}$$

Inserting the Fourier series ansatz into (3.5) and exploiting the orthogonality of the functions $\cos(k\omega t)$ and $\sin(k\omega t)$ with respect to the inner product $(\cdot, \cdot)_{L^2(0, T)}$, i.e., the orthogonalities (2.9), we arrive at the following variational formulation corresponding to every single mode $k \in \mathbb{N}$: Given $\mathbf{f}_k \in (L^2(\Omega))^2$, find $\mathbf{u}_k \in \mathbb{V} := V \times V = (H_0^1(\Omega))^2$ such that

$$\int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{v}_k^\perp(\mathbf{x})) \, d\mathbf{x} = \int_{\Omega} \mathbf{f}_k(\mathbf{x}) \cdot \mathbf{v}_k(\mathbf{x}) \, d\mathbf{x} \quad (3.6)$$

for all $\mathbf{v}_k \in \mathbb{V}$. In the case $k = 0$, we obtain the following variational formulation: Given $f_0^c \in L^2(\Omega)$, find $u_0^c \in V = H_0^1(\Omega)$ such that

$$\int_{\Omega} \nu(\mathbf{x}) \nabla u_0^c(\mathbf{x}) \cdot \nabla v_0^c(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f_0^c(\mathbf{x}) v_0^c(\mathbf{x}) \, d\mathbf{x} \quad (3.7)$$

for all $v_0^c \in V$. The space $\mathbb{V} = (H_0^1(\Omega))^2$ for the Fourier coefficients is equipped with the norm

$$\|\mathbf{u}_k\|_{H^1(\Omega)}^2 = \|\mathbf{u}_k\|_{L^2(\Omega)}^2 + \|\nabla \mathbf{u}_k\|_{L^2(\Omega)}^2.$$

Note that the following relation is valid:

$$\|\mathbf{u}_k^\perp\|_{L^2(\Omega)}^2 = \int_{\Omega} \mathbf{u}_k^\perp \cdot \mathbf{u}_k^\perp \, d\mathbf{x} = \int_{\Omega} ((-u_k^s)^2 + (u_k^c)^2) \, d\mathbf{x} = \int_{\Omega} \mathbf{u}_k \cdot \mathbf{u}_k \, d\mathbf{x} = \|\mathbf{u}_k\|_{L^2(\Omega)}^2.$$

Theorem 3.1. *The variational problems (3.6) and (3.7) have a unique solution.*

Proof. Let us start with the variational problem (3.6) and define the bilinear form

$$a_k(\mathbf{u}_k, \mathbf{v}_k) = \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{v}_k^\perp(\mathbf{x})) \, d\mathbf{x}. \quad (3.8)$$

In order to prove existence and uniqueness of the variational problem (3.6), we have to verify the assumptions of Theorem 2.8, i.e., the Babuška-Aziz theorem. Using triangle and Cauchy-Schwarz

inequalities together with (2.29), i.e., the boundedness of the coefficients ν and σ , we obtain the estimate

$$\begin{aligned} |a_k(\mathbf{u}_k, \mathbf{v}_k)| &\leq \left| \int_{\Omega} \nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) d\mathbf{x} \right| + \left| \int_{\Omega} k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{v}_k^{\perp}(\mathbf{x}) d\mathbf{x} \right| \\ &\leq \bar{\nu} \int_{\Omega} |\nabla \mathbf{u}_k(\mathbf{x})| |\nabla \mathbf{v}_k(\mathbf{x})| d\mathbf{x} + k\omega \bar{\sigma} \int_{\Omega} |\mathbf{u}_k(\mathbf{x})| |\mathbf{v}_k^{\perp}(\mathbf{x})| d\mathbf{x} \\ &\leq \bar{\nu} \|\nabla \mathbf{u}_k\|_{L^2(\Omega)} \|\nabla \mathbf{v}_k\|_{L^2(\Omega)} + k\omega \bar{\sigma} \|\mathbf{u}_k\|_{L^2(\Omega)} \|\mathbf{v}_k^{\perp}\|_{L^2(\Omega)} \\ &\leq \max\{\bar{\nu}, k\omega \bar{\sigma}\} (\|\nabla \mathbf{u}_k\|_{L^2(\Omega)} \|\nabla \mathbf{v}_k\|_{L^2(\Omega)} + \|\mathbf{u}_k\|_{L^2(\Omega)} \|\mathbf{v}_k^{\perp}\|_{L^2(\Omega)}) \\ &\leq \bar{c} \|\mathbf{u}_k\|_{H^1(\Omega)} \|\mathbf{v}_k\|_{H^1(\Omega)} \end{aligned}$$

for all $\mathbf{u}_k, \mathbf{v}_k \in \mathbb{V} = (H_0^1(\Omega))^2$, where $\bar{c} = \max\{\bar{\nu}, k\omega \bar{\sigma}\}$ is the sup-sup constant in the conditions (2.20) and (2.21). Now, we prove the inf-sup conditions for the bilinear form $a_k(\cdot, \cdot)$. By choosing the test function $\mathbf{v}_k = \mathbf{u}_k - \mathbf{u}_k^{\perp}$, we obtain

$$a_k(\mathbf{u}_k, \mathbf{u}_k) = \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{u}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{u}_k^{\perp}(\mathbf{x})) d\mathbf{x} = \int_{\Omega} \nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{u}_k(\mathbf{x}) d\mathbf{x}$$

and

$$\begin{aligned} a_k(\mathbf{u}_k, -\mathbf{u}_k^{\perp}) &= \int_{\Omega} (-\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{u}_k^{\perp}(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{u}_k(\mathbf{x})) d\mathbf{x} \\ &= \int_{\Omega} k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{u}_k(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

where we have used that $(-\mathbf{u}_k^{\perp})^{\perp} = \mathbf{u}_k$. This yields the estimate

$$\begin{aligned} a_k(\mathbf{u}_k, \mathbf{u}_k - \mathbf{u}_k^{\perp}) &= \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{u}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{u}_k(\mathbf{x})) d\mathbf{x} \\ &\geq \min\{\underline{\nu}, k\omega \underline{\sigma}\} \int_{\Omega} (\nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{u}_k(\mathbf{x}) + \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{u}_k(\mathbf{x})) d\mathbf{x} = \underline{c} \|\mathbf{u}_k\|_{H^1(\Omega)}^2 \end{aligned}$$

with $\underline{c} = \min\{\underline{\nu}, k\omega \underline{\sigma}\}$. Moreover, by choosing $\mathbf{v}_k = \mathbf{u}_k + \mathbf{u}_k^{\perp}$, we obtain

$$\begin{aligned} a_k(\mathbf{u}_k^{\perp}, \mathbf{u}_k) &= \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k^{\perp}(\mathbf{x}) \cdot \nabla \mathbf{u}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k^{\perp}(\mathbf{x}) \cdot \mathbf{u}_k^{\perp}(\mathbf{x})) d\mathbf{x} \\ &= \int_{\Omega} k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{u}_k(\mathbf{x}) d\mathbf{x}, \\ a_k(\mathbf{u}_k + \mathbf{u}_k^{\perp}, \mathbf{u}_k) &= \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{u}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{u}_k(\mathbf{x})) d\mathbf{x} \geq \underline{c} \|\mathbf{u}_k\|_{H^1(\Omega)}^2. \end{aligned}$$

Finally, both inf-sup conditions in (2.20) and (2.21) of the Babuška-Aziz theorem are fulfilled with the inf-sup constant $\underline{c} = \min\{\underline{\nu}, k\omega \underline{\sigma}\}$.

Now, let us consider the variational problem (3.7) for the case $k = 0$. We define the bilinear form

$$a_0(u_0^c, v_0^c) = \int_{\Omega} \nu(\mathbf{x}) \nabla u_0^c(\mathbf{x}) \cdot \nabla v_0^c(\mathbf{x}) d\mathbf{x}. \quad (3.9)$$

The assumption that the coefficient ν is uniformly bounded (2.29), together with the Cauchy-Schwarz inequality yields the boundedness of the bilinear form $a_0(\cdot, \cdot)$, i.e.,

$$|a_0(u_0^c, v_0^c)| \leq \bar{\nu} \|u_0^c\|_{H^1(\Omega)} \|v_0^c\|_{H^1(\Omega)}.$$

Moreover, the Friedrichs inequality (2.5) yields the ellipticity of $a_0(u_0^c, v_0^c)$, i.e.,

$$a_0(u_0^c, u_0^c) = \int_{\Omega} \nu(\mathbf{x}) \nabla u_0^c(\mathbf{x}) \cdot \nabla u_0^c(\mathbf{x}) d\mathbf{x} \geq \underline{\nu} \|\nabla u_0^c\|_{L^2(\Omega)}^2 \geq \underline{\nu} \frac{1}{C_F^2 + 1} \|u_0^c\|_{H^1(\Omega)}^2.$$

Altogether, Theorem 2.7, i.e., the Lax-Milgram theorem, yields the existence and uniqueness of a solution of the variational problem (3.7). \square

In order to show existence and uniqueness of the space-time variational problem (3.5) and later of (3.4), we firstly prove the existence of a unique solution of a third space-time variational formulation of our parabolic time-periodic boundary value problem (3.1)-(3.3), which is again weaker than the first variational formulation (3.4). De facto, it will turn out in the existence and uniqueness proof of the second variational problem (3.5) that the third one is equivalent to the second one under the assumption that the given data f is from $L^2(Q_T)$. We now have to define additional, special function spaces for deriving the new variational formulation.

Definition 3.2. *The function spaces $H^{0,\frac{1}{2}}(Q_T)$ and $H^{1,\frac{1}{2}}(Q_T)$ are defined by*

$$H^{0,\frac{1}{2}}(Q_T) = \{u \in L^2(Q_T) : \|\partial_t^{1/2}u\|_{L^2(Q_T)} < \infty\} \quad \text{and}$$

$$H^{1,\frac{1}{2}}(Q_T) = \{u \in H^{1,0}(Q_T) : \|\partial_t^{1/2}u\|_{L^2(Q_T)} < \infty\},$$

respectively, where $\|\partial_t^{1/2}u\|_{L^2(Q_T)}$ is defined in the Fourier space by the relation

$$\|\partial_t^{1/2}u\|_{L^2(Q_T)}^2 := |u|_{H^{0,\frac{1}{2}}(Q_T)}^2 := \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{u}_k\|_{L^2(\Omega)}^2.$$

Let us also define the corresponding inner product

$$(\partial_t^{1/2}u, \partial_t^{1/2}v)_{L^2(Q_T)} := \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\mathbf{u}_k, \mathbf{v}_k)_{L^2(\Omega)},$$

that is a special case of the σ -weighted inner product

$$(\sigma \partial_t^{1/2}u, \partial_t^{1/2}v)_{L^2(Q_T)} := \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\sigma \mathbf{u}_k, \mathbf{v}_k)_{L^2(\Omega)}.$$

for $\sigma = 1$. The space $H_0^{1,\frac{1}{2}}(Q_T)$ is given by

$$H_0^{1,\frac{1}{2}}(Q_T) = \{u \in H^{1,\frac{1}{2}}(Q_T) : u = 0 \text{ on } \Sigma_T\}.$$

The seminorm and the norm of the space $H^{1,\frac{1}{2}}(Q_T)$ are defined by the relations

$$|u|_{H^{1,\frac{1}{2}}}^2 := T \|\nabla u_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} [k\omega \|\mathbf{u}_k\|_{L^2(\Omega)}^2 + \|\nabla \mathbf{u}_k\|_{L^2(\Omega)}^2] \quad \text{and}$$

$$\|u\|_{H^{1,\frac{1}{2}}}^2 := T (\|u_0^c\|_{L^2(\Omega)}^2 + \|\nabla u_0^c\|_{L^2(\Omega)}^2) + \frac{T}{2} \sum_{k=1}^{\infty} [(1+k\omega) \|\mathbf{u}_k\|_{L^2(\Omega)}^2 + \|\nabla \mathbf{u}_k\|_{L^2(\Omega)}^2],$$

respectively.

Furthermore, the following identities can be shown:

Lemma 3.3. *The identities*

$$(\partial_t^{1/2}u, \partial_t^{1/2}v)_{L^2(Q_T)} = (\partial_t u, v^\perp)_{L^2(Q_T)} \quad \text{and} \quad (\partial_t^{1/2}u, \partial_t^{1/2}v^\perp)_{L^2(Q_T)} = (\partial_t u, v)_{L^2(Q_T)} \quad (3.10)$$

are valid for all $u \in H_{per}^{0,1}(Q_T)$ and $v \in H^{0,\frac{1}{2}}(Q_T)$.

Proof. Due to the definition of the left hand sides and inserting the Fourier expansions

$$\partial_t u(\mathbf{x}, t) := \sum_{k=1}^{\infty} [k\omega u_k^s(\mathbf{x}) \cos(k\omega t) - k\omega u_k^c(\mathbf{x}) \sin(k\omega t)]$$

as well as

$$v^\perp(\mathbf{x}, t) := -v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [v_k^s(\mathbf{x}) \cos(k\omega t) - v_k^c(\mathbf{x}) \sin(k\omega t)],$$

into the inner products, we obtain

$$\begin{aligned} (\partial_t^{1/2} u, \partial_t^{1/2} v)_{L^2(Q_T)} &= \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\mathbf{u}_k, \mathbf{v}_k)_{L^2(\Omega)} = \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\mathbf{u}_k^\perp, \mathbf{v}_k^\perp)_{L^2(\Omega)} \\ &= \frac{T}{2} \sum_{k=1}^{\infty} k\omega (-\mathbf{u}_k^\perp, -\mathbf{v}_k^\perp)_{L^2(\Omega)} = (\partial_t u, v^\perp)_{L^2(Q_T)} \end{aligned}$$

and

$$(\partial_t u, v)_{L^2(Q_T)} = \frac{T}{2} \sum_{k=1}^{\infty} k\omega (-\mathbf{u}_k^\perp, \mathbf{v}_k)_{L^2(\Omega)} = \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\mathbf{u}_k, \mathbf{v}_k^\perp)_{L^2(\Omega)} = (\partial_t^{1/2} u, \partial_t^{1/2} v^\perp)_{L^2(Q_T)}.$$

□

Furthermore, we obtain the identity

$$\begin{aligned} \|v^\perp\|_{L^2(Q_T)}^2 &= \int_0^T \int_\Omega \left(-v_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [v_k^s(\mathbf{x}) \cos(k\omega t) - v_k^c(\mathbf{x}) \sin(k\omega t)] \right)^2 d\mathbf{x} dt \\ &= T \int_\Omega v_0^c(\mathbf{x})^2 d\mathbf{x} + \frac{T}{2} \int_\Omega \sum_{k=1}^{\infty} [v_k^s(\mathbf{x})^2 + v_k^c(\mathbf{x})^2] d\mathbf{x} \\ &= T \|v_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{u}_k\|_{L^2(\Omega)}^2 = \|v\|_{L^2(Q_T)}^2, \end{aligned}$$

and the orthogonality relations

$$(\partial_t u, u)_{L^2(Q_T)} = 0 \quad \text{and} \quad (u^\perp, u)_{L^2(Q_T)} = 0 \quad (3.11)$$

for all $u \in H_{per}^{0,1}(Q_T)$, as well as

$$(\partial_t^{1/2} u, \partial_t^{1/2} u^\perp)_{L^2(Q_T)} = 0 \quad \text{and} \quad (\nabla u, \nabla u^\perp)_{L^2(Q_T)} = 0 \quad (3.12)$$

for all $u \in H^{1, \frac{1}{2}}(Q_T)$, e.g., we prove in detail

$$(\partial_t^{1/2} u, \partial_t^{1/2} u^\perp)_{L^2(Q_T)} = \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\mathbf{u}_k, \mathbf{u}_k^\perp)_{L^2(\Omega)} = 0.$$

Remark 3.4. *Indeed, all the identities and orthogonality relations are also valid for their σ - and ν -weighted counterparts, e.g.,*

$$(\sigma \partial_t^{1/2} u, \partial_t^{1/2} u^\perp)_{L^2(Q_T)} = 0 \quad \text{and} \quad (\nu \nabla u, \nabla u^\perp)_{L^2(Q_T)} = 0 \quad (3.13)$$

for all $u \in H^{1, \frac{1}{2}}(Q_T)$.

In the Fourier space, the Cauchy-Schwarz inequality follows from the usual Cauchy-Schwarz inequality (2.1) for functions in $L^2(\Omega)$, e.g.,

$$\begin{aligned}
|(u, v)| &= \left| T(u_0^c, v_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^{\infty} (\mathbf{u}_k, \mathbf{v}_k)_{L^2(\Omega)} \right| \\
&\leq T \|u_0^c\|_{L^2(\Omega)} \|v_0^c\|_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{u}_k\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} \\
&\leq \left(T \|u_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{u}_k\|_{L^2(\Omega)}^2 \right)^{1/2} \left(T \|v_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{v}_k\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&= \|u\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)}
\end{aligned} \tag{3.14}$$

for all $u, v \in L^2(Q_T)$, where $(\cdot, \cdot) = (\cdot, \cdot)_{L^2(Q_T)}$, and

$$\begin{aligned}
|(\partial_t^{1/2} u, \partial_t^{1/2} v)| &= \left| \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\mathbf{u}_k, \mathbf{v}_k)_{L^2(\Omega)} \right| \leq \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{u}_k\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} \\
&\leq \left(\frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{u}_k\|_{L^2(\Omega)}^2 \right)^{1/2} \left(\frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{v}_k\|_{L^2(\Omega)}^2 \right)^{1/2} \\
&= \|\partial_t^{1/2} u\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)}
\end{aligned} \tag{3.15}$$

for all $u, v \in H^{0, \frac{1}{2}}(Q_T)$.

Now, we are in the position to state a very general variational formulation of our parabolic time-periodic boundary value problem (3.1)-(3.3):

Given $f \in L^2(Q_T)$, find $u \in H_0^{1, \frac{1}{2}}(Q_T)$ such that

$$\begin{aligned}
\int_0^T \int_{\Omega} \left(\sigma(\mathbf{x}) \partial_t^{1/2} u(\mathbf{x}, t) \partial_t^{1/2} v^\perp(\mathbf{x}, t) + \nu(\mathbf{x}) \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}, t) \right) d\mathbf{x} dt \\
= \int_0^T \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}, t) d\mathbf{x} dt
\end{aligned} \tag{3.16}$$

for all test functions $v \in H_0^{1, \frac{1}{2}}(Q_T)$, where all functions are given in their Fourier series expansion in time, i.e., everything has to be understood in the sense of Definition 3.2, e.g., inserting the Fourier series ansatz in the variational formulation (3.5) and using all the definitions and identities before. The following lemma, i.e., Lemma 3.5, provides the existence of a unique solution of the variational problem (3.16) and serves as vehicle for the existence and uniqueness proof of the space-time variational problem (3.5) and for discussing the existence of a unique solution of problem (3.4), see [112]. Moreover, all formulations in the sense of Definition 3.2 will be the basis for the construction of preconditioners and the discretization error analysis.

Lemma 3.5. *The space-time bilinear form*

$$a(u, v) = \int_0^T \int_{\Omega} \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} v^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla v \right) d\mathbf{x} dt \tag{3.17}$$

fulfills the following inf-sup and sup-sup condition:

$$\mu_1 \|u\|_{H^{1, \frac{1}{2}}(Q_T)} \leq \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u, v)}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}} \leq \mu_2 \|u\|_{H^{1, \frac{1}{2}}(Q_T)} \tag{3.18}$$

for all $u \in H_0^{1, \frac{1}{2}}(Q_T)$ with positive constants $\mu_1 = \min\{\frac{\nu}{C_F^2+1}, \underline{\sigma}\}$ and $\mu_2 = \max\{\bar{\sigma}, \bar{\nu}\}$.

Proof. We start with the proof of the sup-sup condition. Using (2.29), i.e., the boundedness of the coefficients σ and ν , as well as the triangle and the Cauchy-Schwarz inequalities (3.14) and (3.15), we obtain the estimate

$$\begin{aligned} |a(u, v)| &= \left| \int_0^T \int_{\Omega} \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} v^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla v \right) d\mathbf{x} dt \right| \\ &\leq \bar{\sigma} \int_0^T \int_{\Omega} \left| \partial_t^{1/2} u \right| \left| \partial_t^{1/2} v^\perp \right| d\mathbf{x} dt + \bar{\nu} \int_0^T \int_{\Omega} |\nabla u| |\nabla v| d\mathbf{x} dt \\ &\leq \bar{\sigma} \left\| \partial_t^{1/2} u \right\|_{L^2(Q_T)} \left\| \partial_t^{1/2} v^\perp \right\|_{L^2(Q_T)} + \bar{\nu} \|\nabla u\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)}. \end{aligned}$$

Since

$$\left\| \partial_t^{1/2} v^\perp \right\|_{L^2(Q_T)}^2 = \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{v}_k^\perp\|_{L^2(\Omega)}^2 = \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{v}_k\|_{L^2(\Omega)}^2 = \left\| \partial_t^{1/2} v \right\|_{L^2(Q_T)}^2,$$

we finally prove the sup-sup condition by

$$\begin{aligned} |a(u, v)| &\leq \bar{\sigma} \left\| \partial_t^{1/2} u \right\|_{L^2(Q_T)} \left\| \partial_t^{1/2} v^\perp \right\|_{L^2(Q_T)} + \bar{\nu} \|\nabla u\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &= \bar{\sigma} \left\| \partial_t^{1/2} u \right\|_{L^2(Q_T)} \left\| \partial_t^{1/2} v \right\|_{L^2(Q_T)} + \bar{\nu} \|\nabla u\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\leq \max\{\bar{\sigma}, \bar{\nu}\} \|u\|_{H^{1, \frac{1}{2}}} \|v\|_{H^{1, \frac{1}{2}}} \\ &\leq \mu_2 \|u\|_{H^{1, \frac{1}{2}}} \|v\|_{H^{1, \frac{1}{2}}} \end{aligned}$$

with the constant $\mu_2 = \max\{\bar{\sigma}, \bar{\nu}\}$. Next, we prove the inf-sup condition by choosing the test function $v = u - u^\perp$ and using the σ - and ν -weighted orthogonality relations (3.13) as well as Friedrichs inequality (2.17), see Theorem 2.5, which we write in the Fourier space by

$$\begin{aligned} \|\nabla u\|_{L^2(Q_T)}^2 &= \int_0^T \int_{\Omega} |\nabla u|^2 d\mathbf{x} dt = T \|\nabla u_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\nabla \mathbf{u}_k\|_{L^2(\Omega)}^2 \\ &\geq \frac{1}{C_F^2} \left(T \|u_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{u}_k\|_{L^2(\Omega)}^2 \right) = \frac{1}{C_F^2} \|u\|_{L^2(Q_T)}^2. \end{aligned} \tag{3.19}$$

So, it follows that

$$\|u\|_{H^{1,0}(Q_T)}^2 = \|u\|_{L^2(Q_T)}^2 + \|\nabla u\|_{L^2(Q_T)}^2 \leq (C_F^2 + 1) \|\nabla u\|_{L^2(Q_T)}^2.$$

Hence, we get

$$\begin{aligned} a(u, u) &= \int_0^T \int_{\Omega} \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} u^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla u \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \nu(\mathbf{x}) \nabla u \cdot \nabla u d\mathbf{x} dt \\ &\geq \underline{\nu} \int_0^T \int_{\Omega} |\nabla u|^2 d\mathbf{x} dt \\ &\geq \underline{\nu} \frac{1}{C_F^2 + 1} \|u\|_{H^{1,0}(Q_T)}^2 \end{aligned}$$

and, since $(u^\perp)^\perp = -u$ and $(-u^\perp)^\perp = u$,

$$\begin{aligned} a(u, -u^\perp) &= \int_0^T \int_\Omega \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} u - \nu(\mathbf{x}) \nabla u \cdot \nabla u^\perp \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} u d\mathbf{x} dt \\ &\geq \underline{\sigma} \|\partial_t^{1/2} u\|_{L^2(Q_T)}^2. \end{aligned}$$

Altogether, we have

$$a(u, u - u^\perp) \geq \min\left\{ \frac{\underline{\nu}}{C_F^2 + 1}, \underline{\sigma} \right\} \left(\|u\|_{H^{1,0}(Q_T)}^2 + \|\partial_t^{1/2} u\|_{L^2(Q_T)}^2 \right) = \mu_1 \|u\|_{H^{1, \frac{1}{2}}(Q_T)}^2$$

with the constant $\mu_1 = \min\left\{ \frac{\underline{\nu}}{C_F^2 + 1}, \underline{\sigma} \right\}$. \square

Theorem 3.6. *The space-time variational problem (3.16) has a unique solution.*

Proof. The proof immediately follows from Lemma 3.5 by applying Theorem 2.8, i.e., the Babuška-Aziz theorem. \square

Now, we are in the position to prove existence and uniqueness of the second space-time variational problem (3.5), where we reuse some ideas known from the analysis of parabolic initial-boundary value problems in [108, 109], and adapt them to the time-periodic case, see [112].

Theorem 3.7. *The space-time variational problem (3.5) has a unique solution.*

Proof. We start with the uniqueness proof. Let us assume that there are two different solutions $u_1, u_2 \in H_0^{1,0}(Q_T)$ of the problem (3.5). We expand these two solutions into Fourier series in time, i.e.,

$$u_i(\mathbf{x}, t) = u_{i0}^c(\mathbf{x}) + \sum_{k=1}^{\infty} [u_{ik}^c(\mathbf{x}) \cos(k\omega t) + u_{ik}^s(\mathbf{x}) \sin(k\omega t)],$$

whose unique Fourier coefficients are given by

$$\begin{aligned} u_{i0}^c(\mathbf{x}) &= \frac{1}{T} \int_0^T u_i(\mathbf{x}, t) dt, \\ u_{ik}^c(\mathbf{x}) &= \frac{2}{T} \int_0^T u_i(\mathbf{x}, t) \cos(k\omega t) dt \quad \text{and} \quad u_{ik}^s(\mathbf{x}) = \frac{2}{T} \int_0^T u_i(\mathbf{x}, t) \sin(k\omega t) dt, \end{aligned}$$

for $i = 1, 2$, see Theorem 2.2. Since $u_1 \neq u_2$, we have $u_{1k}^j(\mathbf{x}) \neq u_{2k}^j(\mathbf{x})$ for at least one $k \in \mathbb{N}_0$ and $j \in \{c, s\}$, and therefore $\mathbf{u}_{1k} \neq \mathbf{u}_{2k}$. Let us fix such a k for which $u_{1k}^j \neq u_{2k}^j$. For this k , we define the difference $w_k^j := u_{2k}^j - u_{1k}^j$ and $\mathbf{w}_k = (w_k^c, w_k^s)^T \in \mathbb{V} = V \times V = (H_0^1(\Omega))^2$. After inserting the Fourier series ansatz for u_1 and u_2 into the variational problem (3.5), see also Remark 2.1, the whole system decouples, and we arrive at variational problems for the Fourier coefficients with respect to every single mode k , analogously as in (3.6) and (3.7). For this index k , the difference $\mathbf{w}_k \in \mathbb{V}$ satisfies the variational equation

$$a_k(\mathbf{w}_k, \mathbf{v}_k) := \int_\Omega k\omega \sigma(\mathbf{x}) \mathbf{w}_k \cdot \mathbf{v}_k^\perp d\mathbf{x} + \int_\Omega \nu(\mathbf{x}) \nabla \mathbf{w}_k \cdot \nabla \mathbf{v}_k d\mathbf{x} = 0$$

for all $\mathbf{v}_k \in \mathbb{V}$. We now choose the test function $\mathbf{v}_k = \mathbf{w}_k - \mathbf{w}_k^\perp$. Hence, we obtain

$$\int_\Omega k\omega \sigma(\mathbf{w}_k \cdot \mathbf{w}_k^\perp - \mathbf{w}_k \cdot (\mathbf{w}_k^\perp)^\perp) d\mathbf{x} + \int_\Omega \nu(\nabla \mathbf{w}_k \cdot \nabla \mathbf{w}_k - \nabla \mathbf{w}_k \cdot \nabla \mathbf{w}_k^\perp) d\mathbf{x} = 0.$$

From $(\mathbf{w}_k^\perp)^\perp = -\mathbf{w}_k$ and $\mathbf{w}_k \cdot \mathbf{w}_k^\perp = 0$, it follows that

$$\int_{\Omega} (k\omega \sigma(\mathbf{x}) \mathbf{w}_k \cdot \mathbf{w}_k + \nu(\mathbf{x}) \nabla \mathbf{w}_k \cdot \nabla \mathbf{w}_k) \, d\mathbf{x} = 0.$$

Hence, the inequalities

$$\begin{aligned} 0 &= a_k(\mathbf{w}_k, \mathbf{w}_k - \mathbf{w}_k^\perp) \geq k\omega \underline{\sigma} \int_{\Omega} \mathbf{w}_k \cdot \mathbf{w}_k \, d\mathbf{x} + \underline{\nu} \int_{\Omega} \nabla \mathbf{w}_k \cdot \nabla \mathbf{w}_k \, d\mathbf{x} \\ &\geq \min\{k\omega \underline{\sigma}, \underline{\nu}\} \|\mathbf{w}_k\|_{H^1(\Omega)}^2 \end{aligned}$$

immediately yield $\mathbf{w}_k = 0$ in \mathbb{V} , which is in contradiction to our assumption at the beginning. For $k = 0$, Friedrichs inequality (2.17) gives the same result. Thus, the uniqueness of a solution $u \in H_0^{1,0}(Q_T)$ of the variational problem (3.5) is proven. Now we come to the existence proof. Let

$$u_N(\mathbf{x}, t) = u_0^c(\mathbf{x}) + \sum_{k=1}^N [u_k^c(\mathbf{x}) \cos(k\omega t) + u_k^s(\mathbf{x}) \sin(k\omega t)] \in H_0^{1,0}(Q_T)$$

be the solution of the variational problem

$$\int_0^T \int_{\Omega} (-\sigma(\mathbf{x}) u_N \partial_t v_N + \nu(\mathbf{x}) \nabla u_N \cdot \nabla v_N) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} f v_N \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} f_N v_N \, d\mathbf{x} \, dt$$

for all truncated Fourier series v_N of functions v from $H_{0,per}^{1,1}(Q_T)$. Existence and uniqueness are ensured by the orthogonalities of the cosine and sine functions and by Theorem 3.1. Moreover, u_N solves the variational problems (3.4) and (3.16) for all test functions v_N such as defined above. Choosing the test function $v_N = u_N - u_N^\perp$ and using Lemma 3.5 as well as applying the Cauchy-Schwarz inequality, we obtain the estimates

$$\begin{aligned} \mu_1 \|u_N\|_{H^{1,\frac{1}{2}}(Q_T)}^2 &\leq a(u_N, u_N - u_N^\perp) = \int_0^T \int_{\Omega} f (u_N - u_N^\perp) \, d\mathbf{x} \, dt \\ &\leq \|f\|_{L^2(Q_T)} (\|u_N\|_{L^2(Q_T)} + \|u_N^\perp\|_{L^2(Q_T)}) = 2\|f\|_{L^2(Q_T)} \|u_N\|_{L^2(Q_T)} \\ &\leq 2\|f\|_{L^2(Q_T)} \|u_N\|_{H^{1,\frac{1}{2}}(Q_T)}. \end{aligned}$$

Hence, u_N is bounded in $H^{1,\frac{1}{2}}(Q_T)$, i.e.,

$$\|u_N\|_{H^{1,\frac{1}{2}}(Q_T)} \leq 2\mu_1^{-1} \|f\|_{L^2(Q_T)} < \infty.$$

From this estimate it follows that there exists a function u from $H^{1,\frac{1}{2}}(Q_T)$ such that u_N (without loss of generality) weakly converges to u in $H^{1,\frac{1}{2}}(Q_T)$, and therefore also in $H_0^{1,0}(Q_T)$. Note that since the space $H^{1,\frac{1}{2}}(Q_T)$ is a reflexive Banach space, boundedness implies that it is weakly compact. It remains to show that u solves our variational problem (3.5). Let us choose $M \in \mathbb{N}$ arbitrarily and $N \in \mathbb{N}$ with $N \leq M$, where $v_M \rightarrow v$ in $H_{0,per}^{1,1}(Q_T)$ and $f_N \rightarrow f$ in $L^2(Q_T)$. Inserting u_N, f_N and the arbitrary test function v_M into the variational problem (3.5) yields

$$\int_0^T \int_{\Omega} (-\sigma(\mathbf{x}) u_N \partial_t v_M + \nu(\mathbf{x}) \nabla u_N \cdot \nabla v_M) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} f_N v_M \, d\mathbf{x} \, dt.$$

Since the test function $v_M \rightarrow v$ in $H_{0,per}^{1,1}(Q_T)$ for $M \rightarrow \infty$, we arrive at the identity

$$\int_0^T \int_{\Omega} (-\sigma(\mathbf{x}) u_N \partial_t v + \nu(\mathbf{x}) \nabla u_N \cdot \nabla v) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} f_N v \, d\mathbf{x} \, dt.$$

Now, we pass N to the limit. Since $f_N \rightarrow f$ in $L^2(Q_T)$ and $u_N \rightarrow u$ in $H_0^{1,0}(Q_T)$, we finally get

$$\int_0^T \int_{\Omega} (-\sigma(\mathbf{x})u \partial_t v + \nu(\mathbf{x})\nabla u \cdot \nabla v) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} f v \, d\mathbf{x} \, dt,$$

which means that $u \in H_0^{1,0}(Q_T)$ solves the variational problem (3.5). \square

Remark 3.8. From the proof of Theorem 3.7 follows that u is even from $H^{1,\frac{1}{2}}(Q_T)$ assuming that $f \in L^2(Q_T)$. In this case, the unique solution u of the space-time variational problem (3.5) is also the unique solution of the variational problem (3.16), i.e., the variational problems (3.5) and (3.16) are equivalent.

Remark 3.9. If $f \in L^2(Q_T)$ and we assume that

$$\operatorname{div}(\nu \nabla u) \in L^2(Q_T),$$

then it follows very easily from

$$\int_0^T \int_{\Omega} \sigma(\mathbf{x}) \partial_t u(\mathbf{x}, t) v(\mathbf{x}, t) \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} (f(\mathbf{x}, t) + \operatorname{div}(\nu(\mathbf{x}) \nabla u(\mathbf{x}, t))) v(\mathbf{x}, t) \, d\mathbf{x} \, dt \quad (3.20)$$

for all test functions $v \in H_{0,per}^{1,1}(Q_T)$ that

$$\sigma \partial_t u \in L^2(Q_T),$$

and, hence,

$$u \in H_{0,per}^{1,1}(Q_T).$$

Due to Theorem 2.3, the Fourier series of u converges strongly in $u \in H_0^{1,1}(Q_T)$. Hence, by choosing the test functions

$$v_k^c(\mathbf{x}) \cos(k\omega t) \quad \forall k = 0, 1, \dots \quad \text{and} \quad v_k^s(\mathbf{x}) \sin(k\omega t) \quad \forall k = 1, 2, \dots$$

with $v_k^j(\mathbf{x}) \in H_0^1(\Omega)$, $j \in \{c, s\}$, and due to the orthogonalities of the cosine and sine functions, we arrive at the following problems for all modes $k = 1, 2, \dots$:

$$\int_{\Omega} k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{v}_k^\perp(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} (\mathbf{f}_k(\mathbf{x}) + \operatorname{div}(\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}))) \cdot \mathbf{v}_k(\mathbf{x}) \, d\mathbf{x}. \quad (3.21)$$

In the case $k = 0$, we obtain the problem

$$- \int_{\Omega} \operatorname{div}(\nu(\mathbf{x}) \nabla u_0^c(\mathbf{x})) v_0^c(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f_0^c(\mathbf{x}) v_0^c(\mathbf{x}) \, d\mathbf{x}. \quad (3.22)$$

Contrariwise, if, in addition to $f_k^j \in L^2(\Omega)$, we have that

$$\operatorname{div}(\nu \nabla u_k^j) \in L^2(\Omega)$$

for all Fourier coefficients with $k = 0, 1, \dots$ and $j = c$ and for all Fourier coefficients with $k = 1, 2, \dots$ and $j = s$, then, together with Theorem 2.3 and the equations (3.21) and (3.22), it follows the strong convergence of the Fourier series of u in $H_{0,per}^{1,1}(Q_T)$ of problem (3.20). Moreover, u is the unique solution of the first space-time variational formulation (3.4) of the parabolic time-periodic boundary value problem (3.1)-(3.3).

Remark 3.10. Under classical regularity assumptions imposed on u , e.g., $u \in C^{2,1}(\overline{Q}_T)$, and on the data f , σ and ν , e.g., $f \in C(\overline{Q}_T)$, $\sigma \in C(\overline{\Omega})$ and $\nu \in C^1(\overline{\Omega})$, it follows that u is the unique solution of the parabolic time-periodic boundary value problem (3.1)-(3.3) in the classical sense.

3.2 Multiharmonic finite element discretization

In order to solve the parabolic time-periodic problem (3.1)-(3.3), we numerically solve the variational problem (3.16) by a multiharmonic finite element discretization. We choose the test functions

$$v_k^c(\mathbf{x}) \cos(k\omega t) \quad \forall k = 0, \dots, N \quad \text{and} \quad v_k^s(\mathbf{x}) \sin(k\omega t) \quad \forall k = 1, \dots, N \quad (3.23)$$

with $v_k^j(\mathbf{x}) \in H_0^1(\Omega)$, $j \in \{c, s\}$, in the space-time variational problem (3.16). Due to the orthogonalities of the cosine and sine functions, we arrive at the variational problems (3.6) corresponding to every mode $k = 1, 2, \dots, N$, i.e., given $\mathbf{f}_k \in (L^2(\Omega))^2$, find $\mathbf{u}_k \in \mathbb{V} := V \times V = (H_0^1(\Omega))^2$ such that

$$\int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{v}_k^\perp(\mathbf{x})) \, d\mathbf{x} = \int_{\Omega} \mathbf{f}_k(\mathbf{x}) \cdot \mathbf{v}_k(\mathbf{x}) \, d\mathbf{x}$$

for all $\mathbf{v}_k \in \mathbb{V}$, and, for $k = 0$, we obtain again the variational problem (3.7), i.e., given $f_0^c \in L^2(\Omega)$, find $u_0^c \in V = H_0^1(\Omega)$ such that

$$\int_{\Omega} \nu(\mathbf{x}) \nabla u_0^c(\mathbf{x}) \cdot \nabla v_0^c(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f_0^c(\mathbf{x}) v_0^c(\mathbf{x}) \, d\mathbf{x}$$

for all $v_0^c \in V$. An equivalent approach for deriving the variational problems which correspond to every mode $k = 0, \dots, N$ is to approximate the data f by truncating its Fourier series expansion. Hence, we arrive at

$$f(\mathbf{x}, t) \approx f_0^c(\mathbf{x}) + \sum_{k=1}^N [f_k^c(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \sin(k\omega t)] = f_N(\mathbf{x}, t), \quad (3.24)$$

where its Fourier coefficients are given by

$$\begin{aligned} f_0^c(\mathbf{x}) &= \frac{1}{T} \int_0^T f(\mathbf{x}, t) \, dt, \\ f_k^c(\mathbf{x}) &= \frac{2}{T} \int_0^T f(\mathbf{x}, t) \cos(k\omega t) \, dt, \\ f_k^s(\mathbf{x}) &= \frac{2}{T} \int_0^T f(\mathbf{x}, t) \sin(k\omega t) \, dt. \end{aligned}$$

Remark 3.11. *In general the Fourier coefficients have to be computed numerically, but we consider only the case where we can compute the Fourier coefficients exactly.*

Remark 3.12. *In the case that f has a multiharmonic representation, the analysis is simplified since we do not have to consider the discretization error due to truncation of the Fourier series expansions.*

We insert the truncated Fourier series expansion (3.24) of f and the Fourier series ansatz of the solution u and of the test function v into the space-time variational formulation (3.16). From the orthogonality of the functions $\cos(k\omega t)$ and $\sin(k\omega t)$ it follows that it is sufficient to consider only the truncated Fourier series of u and v , i.e.,

$$\begin{aligned} u(\mathbf{x}, t) &\approx u_0^c(\mathbf{x}) + \sum_{k=1}^N [u_k^c(\mathbf{x}) \cos(k\omega t) + u_k^s(\mathbf{x}) \sin(k\omega t)] = u_N(\mathbf{x}, t), \\ v(\mathbf{x}, t) &\approx v_0^c(\mathbf{x}) + \sum_{k=1}^N [v_k^c(\mathbf{x}) \cos(k\omega t) + v_k^s(\mathbf{x}) \sin(k\omega t)] = v_N(\mathbf{x}, t). \end{aligned}$$

We finally arrive at the same variational problems as before, i.e., problem (3.6) corresponding to every mode $k = 1, 2, \dots, N$ and (3.7) for $k = 0$.

Now, we approximate the unknown Fourier coefficients

$$\mathbf{u}_k = (u_k^c, u_k^s)^T \in \mathbb{V}$$

by finite element functions

$$\mathbf{u}_{kh} = (u_{kh}^c, u_{kh}^s)^T \in \mathbb{V}_h = V_h \times V_h \subset \mathbb{V}.$$

The space $\mathbb{V}_h = V_h \times V_h$ is a finite element space, where

$$V_h = \text{span}\{\varphi_1, \dots, \varphi_n\}$$

with the standard nodal basis $\{\varphi_i(\mathbf{x}) = \varphi_{ih}(\mathbf{x}) : i = 1, 2, \dots, n_h\}$ and h denotes the usual discretization parameter such that $n = n_h = \dim V_h = O(h^{-d})$. In this work, we will use continuous, piecewise linear functions on the finite elements on a regular triangulation \mathcal{T}_h to construct the finite element subspace V_h and its basis, see, e.g., [41, 46, 84, 161] and Section 2.4. This yields the following linear system arising from the variational problem (3.6) for $k = 1, 2, \dots, N$:

$$\begin{pmatrix} K_{h,\nu} & -k\omega M_{h,\sigma} \\ k\omega M_{h,\sigma} & K_{h,\nu} \end{pmatrix} \begin{pmatrix} \underline{u}_k^c \\ \underline{u}_k^s \end{pmatrix} = \begin{pmatrix} \underline{f}_k^c \\ \underline{f}_k^s \end{pmatrix}. \quad (3.25)$$

Let us assume that the parameter σ is positive. We rewrite the linear system (3.25) in a symmetric form. Hence, we obtain the saddle point system as follows

$$\begin{pmatrix} k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} \end{pmatrix} \begin{pmatrix} \underline{u}_k^s \\ \underline{u}_k^c \end{pmatrix} = \begin{pmatrix} -\underline{f}_k^c \\ -\underline{f}_k^s \end{pmatrix}, \quad (3.26)$$

which has to be solved with respect to the nodal parameter vector

$$\underline{u}_k^j = (u_{k,i}^j)_{i=1,\dots,n} \in \mathbb{R}^n$$

of the finite element approximation

$$u_{kh}^j(\mathbf{x}) = \sum_{i=1}^n u_{k,i}^j \varphi_i(\mathbf{x})$$

to the unknown Fourier coefficients $u_k^j(\mathbf{x})$ with $j \in \{c, s\}$. The matrices $K_{h,\nu}$ and $M_{h,\sigma}$ correspond to the weighted stiffness matrix and weighted mass matrix, respectively. Their entries are computed by the formulas

$$K_{h,\nu}^{ij} = \int_{\Omega} \nu \nabla \varphi_i \cdot \nabla \varphi_j \, d\mathbf{x} \quad \text{and} \quad M_{h,\sigma}^{ij} = \int_{\Omega} \sigma \varphi_i \varphi_j \, d\mathbf{x}$$

with $i, j = 1, \dots, n$, whereas

$$\underline{f}_k^c = \left[\int_{\Omega} f_k^c \varphi_j \, d\mathbf{x} \right]_{j=1,\dots,n} \quad \text{and} \quad \underline{f}_k^s = \left[\int_{\Omega} f_k^s \varphi_j \, d\mathbf{x} \right]_{j=1,\dots,n}.$$

In the case $k = 0$, we obtain the following linear system arising from the variational problem (3.7):

$$K_{h,\nu} \underline{u}_0^c = \underline{f}_0^c. \quad (3.27)$$

From the solutions of the linear systems (3.26) and (3.27), we can easily reconstruct the multiharmonic finite element approximation

$$u_{Nh}(\mathbf{x}, t) = u_{0h}^c(\mathbf{x}) + \sum_{k=1}^N [u_{kh}^c(\mathbf{x}) \cos(k\omega t) + u_{kh}^s(\mathbf{x}) \sin(k\omega t)] \quad (3.28)$$

to the exact solution $u(\mathbf{x}, t)$. We will present an a priori error analysis for the complete discretization error between the unknown solution u and its multiharmonic finite element approximation u_{Nh} in Section 3.4.

3.3 Block-diagonal preconditioned MINRES solver

The aim of this section is to construct robust block-diagonal preconditioners for the MINRES method in order to solve the saddle point problem (3.26) for all $k = 1, 2, \dots, N$ by applying the interpolation theory presented in Section 2.7. A convergence result for the preconditioned MINRES method is stated in Theorem 2.25.

Remark 3.13. *In this work, we assume that the parameter σ is strictly positive, i.e.,*

$$0 < \underline{\sigma} \leq \sigma(\mathbf{x}) \leq \bar{\sigma}, \quad \mathbf{x} \in \Omega.$$

If σ is only non-negative but not strictly positive, then the linear system (3.25) decouples into linear systems of the form (3.27) for $\sigma = 0$. Hence, we have to solve the following problems:

$$\begin{aligned} K_{h,\nu} \underline{u}_k^c &= \underline{f}_k^c & \forall k = 0, 1, \dots, N, \\ K_{h,\nu} \underline{u}_k^s &= \underline{f}_k^s & \forall k = 1, 2, \dots, N. \end{aligned}$$

Here, the system matrix is given by

$$\mathcal{A} = \begin{pmatrix} k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} \end{pmatrix}.$$

Since $(k\omega M_{h,\sigma})$ is symmetric and positive definite, we can build the two Schur complements S and R from Theorem 2.26, i.e.,

$$S = R = k\omega M_{h,\sigma} + \frac{1}{k\omega} K_{h,\nu} M_{h,\sigma}^{-1} K_{h,\nu},$$

which yield the two Schur complement preconditioners (2.49) for \mathcal{A} , i.e.,

$$\mathcal{P}_0 = \begin{pmatrix} k\omega M_{h,\sigma} & 0 \\ 0 & S \end{pmatrix} \quad \text{and} \quad \mathcal{P}_1 = \begin{pmatrix} R & 0 \\ 0 & k\omega M_{h,\sigma} \end{pmatrix}.$$

As mentioned in Section 2.7, it is hard to work with these Schur complements in practice and, therefore, we construct block-diagonal preconditioners \mathcal{P}_θ by interpolating between \mathcal{P}_0 and \mathcal{P}_1 as presented in (2.52). By choosing the parameter $\theta = 1/2$, we obtain the preconditioner

$$\mathcal{P}_{1/2} = \begin{pmatrix} [k\omega M_{h,\sigma}, R]_{1/2} & 0 \\ 0 & [S, k\omega M_{h,\sigma}]_{1/2} \end{pmatrix}$$

with

$$\begin{aligned} [k\omega M_{h,\sigma}, R]_{1/2} &= [R, k\omega M_{h,\sigma}]_{1/2} = (k\omega M_{h,\sigma})^{1/2} \left((k\omega M_{h,\sigma})^{-1/2} R (k\omega M_{h,\sigma})^{-1/2} \right)^{1/2} (k\omega M_{h,\sigma})^{1/2} \\ &= \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(M_{h,\sigma}^{-1/2} R M_{h,\sigma}^{-1/2} \right)^{1/2} M_{h,\sigma}^{1/2} = \sqrt{k\omega} [M_{h,\sigma}, R]_{1/2} \end{aligned}$$

and, due to $S = R$,

$$[S, k\omega M_{h,\sigma}]_{1/2} = [k\omega M_{h,\sigma}, R]_{1/2} = \sqrt{k\omega} [M_{h,\sigma}, R]_{1/2}.$$

The diagonal entries of $\mathcal{P}_{1/2}$ can be estimated from above and below by using the inequality

$$\frac{1}{\sqrt{2}} (\sqrt{a} I + \sqrt{b} X^{1/2}) \leq (a I + b X)^{1/2} \leq \sqrt{a} I + \sqrt{b} X^{1/2}, \quad (3.29)$$

which has to be understood in the sense of Definition 2.30. Here, a and b are some arbitrary positive real numbers and I and X denote the identity matrix and an arbitrary symmetric positive definite matrix, both in $\mathbb{R}^{n \times n}$, respectively. We obtain the estimates in the following way:

$$\begin{aligned}
[k\omega M_{h,\sigma}, R]_{1/2} &= [S, k\omega M_{h,\sigma}]_{1/2} = \sqrt{k\omega} [M_{h,\sigma}, R]_{1/2} = \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(M_{h,\sigma}^{-1/2} R M_{h,\sigma}^{-1/2} \right)^{1/2} M_{h,\sigma}^{1/2} \\
&= \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(M_{h,\sigma}^{-1/2} \left(k\omega M_{h,\sigma} + \frac{1}{k\omega} K_{h,\nu} M_{h,\sigma}^{-1} K_{h,\nu} \right) M_{h,\sigma}^{-1/2} \right)^{1/2} M_{h,\sigma}^{1/2} \\
&= \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(k\omega I + \frac{1}{k\omega} M_{h,\sigma}^{-1/2} K_{h,\nu} M_{h,\sigma}^{-1} K_{h,\nu} M_{h,\sigma}^{-1/2} \right)^{1/2} M_{h,\sigma}^{1/2} \\
&\leq \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(\sqrt{k\omega} I + \frac{1}{\sqrt{k\omega}} \left(M_{h,\sigma}^{-1/2} K_{h,\nu} M_{h,\sigma}^{-1} K_{h,\nu} M_{h,\sigma}^{-1/2} \right)^{1/2} \right) M_{h,\sigma}^{1/2} \\
&= k\omega M_{h,\sigma} + M_{h,\sigma}^{1/2} \left(M_{h,\sigma}^{-1/2} K_{h,\nu} M_{h,\sigma}^{-1/2} \right) M_{h,\sigma}^{1/2} = k\omega M_{h,\sigma} + K_{h,\nu},
\end{aligned}$$

$$\begin{aligned}
[k\omega M_{h,\sigma}, R]_{1/2} &= [S, k\omega M_{h,\sigma}]_{1/2} = \sqrt{k\omega} [M_{h,\sigma}, R]_{1/2} = \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(M_{h,\sigma}^{-1/2} R M_{h,\sigma}^{-1/2} \right)^{1/2} M_{h,\sigma}^{1/2} \\
&= \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(k\omega I + \frac{1}{k\omega} M_{h,\sigma}^{-1/2} K_{h,\nu} M_{h,\sigma}^{-1} K_{h,\nu} M_{h,\sigma}^{-1/2} \right)^{1/2} M_{h,\sigma}^{1/2} \\
&\geq \sqrt{k\omega} M_{h,\sigma}^{1/2} \left(\frac{1}{\sqrt{2}} \left(\sqrt{k\omega} I + \frac{1}{\sqrt{k\omega}} \left(M_{h,\sigma}^{-1/2} K_{h,\nu} M_{h,\sigma}^{-1} K_{h,\nu} M_{h,\sigma}^{-1/2} \right)^{1/2} \right) \right) M_{h,\sigma}^{1/2} \\
&= \frac{1}{\sqrt{2}} \left(k\omega M_{h,\sigma} + M_{h,\sigma}^{1/2} \left(M_{h,\sigma}^{-1/2} K_{h,\nu} M_{h,\sigma}^{-1/2} \right) M_{h,\sigma}^{1/2} \right) = \frac{1}{\sqrt{2}} (k\omega M_{h,\sigma} + K_{h,\nu}).
\end{aligned}$$

Altogether, we arrive at the spectral equivalence

$$[k\omega M_{h,\sigma}, R]_{1/2} = [S, k\omega M_{h,\sigma}]_{1/2} \sim k\omega M_{h,\sigma} + K_{h,\nu}$$

with spectral equivalence bounds $\underline{c} = 1/\sqrt{2}$ and $\bar{c} = 1$ according to Definition 2.30. Thus, we have obtained a new block-diagonal preconditioner for a MINRES solver of problem (3.26) which is given by

$$\mathcal{P} = \begin{pmatrix} k\omega M_{h,\sigma} + K_{h,\nu} & 0 \\ 0 & k\omega M_{h,\sigma} + K_{h,\nu} \end{pmatrix}, \quad (3.30)$$

and this block-diagonal preconditioner yields the robust condition number estimate

$$\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) := \|\mathcal{P}^{-1}\mathcal{A}\|_{\mathcal{P}} \|\mathcal{A}^{-1}\mathcal{P}\|_{\mathcal{P}} \leq \bar{c}/\underline{c} = \sqrt{2} \approx 1.414. \quad (3.31)$$

An alternative approach for obtaining robust norm estimates for the preconditioned system matrix $\mathcal{P}^{-1}\mathcal{A}$ is to verify the inf-sup and sup-sup conditions in Theorem 2.8, i.e., the Babuška-Aziz theorem. More precisely, these norm estimates are equivalent to the inf-sup and sup-sup conditions in the Babuška-Aziz theorem, and, at the same time, provide existence, uniqueness, as well as a priori and a posteriori error estimates. For instance, the assumptions of the Babuška-Aziz theorem yield discretization error estimates which we are going to present in Section 3.4.

Let us now verify the inf-sup and sup-sup conditions of the Babuška-Aziz theorem. We return to the variational formulation (3.6) for each mode $k = 1, 2, \dots, N$, where the corresponding bilinear form is defined in (3.8), i.e.,

$$a_k(\mathbf{u}_k, \mathbf{v}_k) = \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{v}_k^{\perp}(\mathbf{x})) \, d\mathbf{x}.$$

Hence, the variational problem (3.6) reads as follows: Given $\mathbf{f}_k \in (L^2(\Omega))^2$, find $\mathbf{u}_k \in \mathbb{V} = (H_0^1(\Omega))^2$ such that

$$a_k(\mathbf{u}_k, \mathbf{v}_k) = \int_{\Omega} \mathbf{f}_k(\mathbf{x}) \cdot \mathbf{v}_k(\mathbf{x}) \, d\mathbf{x} \quad \forall \mathbf{v}_k \in \mathbb{V}.$$

Note that we have reformulated the discretized version of the variational problem (3.6) such that it has a symmetric form, resulting in problem (3.26). This discretized problem simply corresponds to the variational problem (3.6) multiplied by -1 .

The candidate \mathcal{P} for a parameter robust preconditioner yields the following definition of an inner product and of its associated norm. We define a non-standard (weighted) inner product in $\mathbb{V} = (H_0^1(\Omega))^2$ by

$$(\mathbf{u}_k, \mathbf{v}_k)_{\mathcal{P}} = (\nu \nabla \mathbf{u}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{v}_k)_{L^2(\Omega)}. \quad (3.32)$$

The associated norm is then given by

$$\|\mathbf{u}_k\|_{\mathcal{P}}^2 = (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}, \quad (3.33)$$

which differs from the standard H^1 -norms. Now, we are prepared to verify the assumptions of the Babuška-Aziz theorem, i.e., the inf-sup and sup-sup conditions.

Theorem 3.14. *Let the bilinear form $a_k(\cdot, \cdot)$ be defined as in (3.8). The following inequalities are valid:*

$$\underline{c} \|\mathbf{u}_k\|_{\mathcal{P}} \leq \sup_{0 \neq \mathbf{v}_k \in \mathbb{V}} \frac{a_k(\mathbf{u}_k, \mathbf{v}_k)}{\|\mathbf{v}_k\|_{\mathcal{P}}} \leq \bar{c} \|\mathbf{u}_k\|_{\mathcal{P}} \quad (3.34)$$

for all $\mathbf{u}_k \in \mathbb{V}$ with constants $\underline{c} = 1/\sqrt{2}$ and $\bar{c} = 1$.

Proof. We start with the proof of the inequality from above. Due to the triangle inequality, it follows that

$$\begin{aligned} |a_k(\mathbf{u}_k, \mathbf{v}_k)| &\leq \left| \int_{\Omega} \nu \nabla \mathbf{u}_k \cdot \nabla \mathbf{v}_k \, d\mathbf{x} \right| + \left| \int_{\Omega} k\omega \sigma \mathbf{u}_k \cdot \mathbf{v}_k^{\perp} \, d\mathbf{x} \right| \\ &= |(\nu \nabla \mathbf{u}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)}| + |k\omega (\sigma \mathbf{u}_k, \mathbf{v}_k^{\perp})_{L^2(\Omega)}|. \end{aligned}$$

Applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} |a_k(\mathbf{u}_k, \mathbf{v}_k)| &\leq (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)}^{1/2} (\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)}^{1/2} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}^{1/2} (\sigma \mathbf{v}_k^{\perp}, \mathbf{v}_k^{\perp})_{L^2(\Omega)}^{1/2} \\ &= (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)}^{1/2} (\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)}^{1/2} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}^{1/2} (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)}^{1/2} \\ &\leq \left((\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)} \right)^{1/2} \left((\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)} \right. \\ &\quad \left. + k\omega (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)} \right)^{1/2} \\ &= \|\mathbf{u}_k\|_{\mathcal{P}} \|\mathbf{v}_k\|_{\mathcal{P}}. \end{aligned}$$

Hence, we have proved the upper bound with $\bar{c} = 1$. Now, we want to show the estimate from below. With the choice

$$\mathbf{v}_k = \mathbf{u}_k - \mathbf{u}_k^{\perp}$$

and together with the σ - and ν -weighted orthogonality relations for the modes k (3.13), we obtain

$$a_k(\mathbf{u}_k, \mathbf{v}_k) = (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)} = \|\mathbf{u}_k\|_{\mathcal{P}}^2.$$

By using the fact that

$$\|\mathbf{v}_k\|_{\mathcal{P}}^2 = \|\mathbf{u}_k - \mathbf{u}_k^\perp\|_{\mathcal{P}}^2 = 2\|\mathbf{u}_k\|_{\mathcal{P}}^2,$$

we arrive at the estimate of the supremum from below, i.e.,

$$\sup_{0 \neq \mathbf{v}_k \in \mathbb{V}} \frac{a_k(\mathbf{u}_k, \mathbf{v}_k)}{\|\mathbf{v}_k\|_{\mathcal{P}}} \geq \frac{a_k(\mathbf{u}_k, \mathbf{u}_k - \mathbf{u}_k^\perp)}{\|\mathbf{u}_k - \mathbf{u}_k^\perp\|_{\mathcal{P}}} = \frac{\|\mathbf{u}_k\|_{\mathcal{P}}^2}{\sqrt{2}\|\mathbf{u}_k\|_{\mathcal{P}}} = \frac{1}{\sqrt{2}}\|\mathbf{u}_k\|_{\mathcal{P}}.$$

Hence, we get $\underline{c} = 1/\sqrt{2}$, which finally completes the proof of the theorem. \square

Remark 3.15. *Since we have verified the assumptions of the Babuška-Aziz theorem in Theorem 3.14, we immediately obtain the existence of a unique solution to variational problem (3.6), which gives us an additional existence and uniqueness result to the one obtained in Theorem 3.1.*

Due to the supremum, the discrete version of the inf-sup condition, i.e., the left inequality in Theorem 3.14, does in general not follow from the continuous version. However, in our case, we can repeat the proof step-by-step, and, finally, we arrive at the same inequalities in the discrete case, where \mathbb{V} is replaced by \mathbb{V}_h with the same constants \underline{c} and \bar{c} . Therefore, in matrix-vector notation, we have proved the inequalities

$$\underline{c}\|\underline{u}_k\|_{\mathcal{P}} \leq \sup_{\underline{v}_k \in \mathbb{R}^{2n}} \frac{(\mathcal{A}\underline{u}_k, \underline{v}_k)}{\|\underline{v}_k\|_{\mathcal{P}}} \leq \bar{c}\|\underline{u}_k\|_{\mathcal{P}} \quad \forall \underline{u}_k = (\underline{u}_k^c, \underline{u}_k^s)^T \in \mathbb{R}^{2n} \quad (3.35)$$

implying the condition number estimate (3.31), i.e.,

$$\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) := \|\mathcal{P}^{-1}\mathcal{A}\|_{\mathcal{P}} \|\mathcal{A}^{-1}\mathcal{P}\|_{\mathcal{P}} \leq \bar{c}/\underline{c} = \sqrt{2} \approx 1.414.$$

So, these estimates are exactly the same as the ones obtained by interpolation theory. Theorem 2.25 yields a robust convergence rate of the preconditioned MINRES method with

$$q = \frac{\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) - 1}{\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) + 1} \leq \frac{\sqrt{2} - 1}{\sqrt{2} + 1} \approx 0.172$$

of the factor q defining the residual reduction $2q^m/(1+q^{2m})$ after $2m$ MINRES iterations. Altogether, for every mode $k = 1, 2, \dots, N$, we have determined a preconditioner such that the corresponding system can be solved by the preconditioned MINRES method with a robust convergence rate.

In practical applications, the diagonal blocks $(k\omega M_{h,\sigma} + K_{h,\nu})$ of the preconditioner \mathcal{P} in (3.30) of the discretized problem (3.26) for $k = 1, 2, \dots, N$ have to be replaced by diagonal blocks \tilde{D} , which are spectrally equivalent to these weighted sums of mass and stiffness matrices, i.e., $(k\omega M_{h,\sigma} + K_{h,\nu})$, and which are robust, symmetric positive definite and more cost efficient. The construction of such robust and efficient preconditioners can be done by (algebraic) multigrid, multilevel or domain decomposition methods, see, e.g., [101, 142, 168, 173]. In this work, we consider the construction of practical preconditioners via algebraic multilevel iteration methods. Moreover, a detailed proof for the robustness and optimality of an algebraic multilevel preconditioner for weighted sums of mass and stiffness matrices will be presented in Chapter 5. The spectral equivalence of the diagonal blocks implies the spectral equivalence of the new preconditioner $\tilde{\mathcal{P}}$ to the preconditioner \mathcal{P} with the same parameter independent constants \underline{c}_D and \bar{c}_D . Hence, the condition number $\kappa_{\tilde{\mathcal{P}}}(\tilde{\mathcal{P}}^{-1}\mathcal{A})$ can be estimated by

$$\kappa_{\tilde{\mathcal{P}}}(\tilde{\mathcal{P}}^{-1}\mathcal{A}) \leq \kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) (\bar{c}_D/\underline{c}_D),$$

where $\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) \leq \sqrt{2}$ and $\underline{c}_D \tilde{D} \leq (k\omega M_{h,\sigma} + K_{h,\nu}) \leq \bar{c}_D \tilde{D}$. All proofs on this topic including the computation of the constants \underline{c}_D and \bar{c}_D are presented in Chapter 5. Altogether, the new practical block-diagonal preconditioner $\tilde{\mathcal{P}}$ yields again parameter independent convergence rates.

3.4 Discretization error analysis

For the complete error analysis, we have to define some seminorms and norms in certain function spaces inspired by the \mathcal{P} -norm in the same spirit as we did this in the definitions (3.32) and (3.33). Let us define the function spaces

$$V_0 := H_0^{1, \frac{1}{2}}(Q_T) \quad \text{and} \quad V_1 := (H^{0, \frac{1}{2}})_0^1(Q_T) \cap H_{per}^{0,1}(Q_T), \quad (3.36)$$

where

$$(H^{0, \frac{1}{2}})_0^1(Q_T) := \{u \in H^{0, \frac{1}{2}}(Q_T) : \nabla u \in H^{0, \frac{1}{2}}(Q_T), u = 0 \text{ on } \Sigma_T\},$$

and let us equip these spaces with the corresponding seminorms

$$\begin{aligned} |u|_{V_0}^2 &= (\nu \nabla u, \nabla u)_{L^2(Q_T)} + (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)}, \\ |u|_{V_1}^2 &= (\nu \nabla u, \nabla u)_{L^2(Q_T)} + (\nu \partial_t^{1/2} \nabla u, \partial_t^{1/2} \nabla u)_{L^2(Q_T)} + (\sigma \partial_t u, \partial_t u)_{L^2(Q_T)}, \end{aligned}$$

which are again defined in the Fourier space according to Definition 3.2, i.e.,

$$\begin{aligned} |u|_{V_0}^2 &= T (\nu \nabla u_0^c, \nabla u_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^{\infty} [(\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}], \\ |u|_{V_1}^2 &= T (\nu \nabla u_0^c, \nabla u_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^{\infty} [(1 + k\omega)(\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + (k\omega)^2 (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}]. \end{aligned}$$

Due to the Friedrichs inequality, the seminorms $|\cdot|_{V_0}$ and $|\cdot|_{V_1}$ are equivalent to the norms in these spaces. Hence, in the following, we denote by $\|\cdot\|_{V_0}$ and $\|\cdot\|_{V_1}$, in fact, the corresponding seminorms, i.e., we define

$$\|u\|_{V_0} := |u|_{V_0} \quad \text{and} \quad \|u\|_{V_1} := |u|_{V_1}.$$

Note that the \mathcal{P} -norm defined in (3.33) corresponds to the V_0 -norm for a single mode $k = 1, 2, \dots, N$, i.e.,

$$\|u\|_{V_0}^2 = T (\nu \nabla u_0^c, \nabla u_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{u}_k\|_{\mathcal{P}}^2. \quad (3.37)$$

For the mode $k = 0$, the discretized system of the variational problem (3.7) is given by (3.27). Together with (3.37), it follows that, in the case $k = 0$, natural choices for the \mathcal{P} -inner product and the \mathcal{P} -norm are

$$(u_0^c, v_0^c)_{\mathcal{P}} = (\nu \nabla u_0^c, \nabla v_0^c)_{L^2(\Omega)} \quad \text{and} \quad \|u_0^c\|_{\mathcal{P}}^2 = (\nu \nabla u_0^c, \nabla u_0^c)_{L^2(\Omega)}, \quad (3.38)$$

respectively. Moreover, we write $\|\mathbf{u}_k\|_{V_0} = \|\mathbf{u}_k\|_{\mathcal{P}}$ in the case $k = 1, 2, \dots, N$ and $\|u_0^c\|_{V_0} = \|u_0^c\|_{\mathcal{P}}$ for $k = 0$. So finally, the V_0 -norm can be written as

$$\|u\|_{V_0}^2 = T \|u_0^c\|_{V_0}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{u}_k\|_{V_0}^2.$$

Now, we are prepared for the analysis of the complete discretization error between the exact solution of the variational problem (3.6) and its multiharmonic finite element approximation, given by

$$\|u - u_{Nh}\|_{V_0}, \quad (3.39)$$

where the exact solution u can be represented as Fourier series, i.e.,

$$u(\mathbf{x}, t) = u_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [u_k^c(\mathbf{x}) \cos(k\omega t) + u_k^s(\mathbf{x}) \sin(k\omega t)],$$

and its multiharmonic finite element approximation u_{Nh} is given by (3.28), i.e.,

$$u_{Nh}(\mathbf{x}, t) = u_{0h}^c(\mathbf{x}) + \sum_{k=1}^N [u_{kh}^c(\mathbf{x}) \cos(k\omega t) + u_{kh}^s(\mathbf{x}) \sin(k\omega t)].$$

Using the triangle inequality, we can split the discretization error (3.39) into two parts, a discretization error in the truncation index N and a discretization error in the parameter h arising from the finite element discretization, i.e.,

$$\|u - u_{Nh}\|_{V_0} \leq \|u - u_N\|_{V_0} + \|u_N - u_{Nh}\|_{V_0}.$$

3.4.1 Discretization error with respect to the truncation index

The following theorem provides an estimate for the discretization error due to truncation of the Fourier series at the mode N under weak regularity assumptions.

Theorem 3.16. *Let us assume that $u \in V_1$. Then the discretization error due to truncation of the Fourier series can be estimated by*

$$\|u - u_{Nh}\|_{V_0} \leq c_0 N^{-1/2} \|u\|_{V_1}, \quad (3.40)$$

where c_0 is a constant depending only on the frequency ω .

Proof. Under the assumption that $u \in V_1 = (H^{0, \frac{1}{2}})_0^1(Q_T) \cap H_{per}^{0,1}(Q_T)$, we obtain the following estimate:

$$\begin{aligned} \|u - u_N\|_{V_0}^2 &= \frac{T}{2} \sum_{k=N+1}^{\infty} [(\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}] \\ &\leq \frac{T}{2} \sum_{k=N+1}^{\infty} \left[\frac{1+k\omega}{k\omega} (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + \frac{(k\omega)^2}{k\omega} (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)} \right] \\ &\leq \frac{1}{(N+1)\omega} \frac{T}{2} \sum_{k=N+1}^{\infty} [(1+k\omega) (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + (k\omega)^2 (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}] \\ &\leq c_0^2 \frac{1}{N} \|u\|_{V_1}^2, \end{aligned}$$

where $c_0 = c_0(\omega) = 1/\sqrt{\omega}$. □

3.4.2 Discretization error with respect to the finite element discretization parameter

The discretization error between the multiharmonic approximation of the exact solution and its multiharmonic finite element approximation can be reduced to the discretization error between the unknown Fourier coefficients and their finite element approximations due to the identity

$$\|u_N - u_{Nh}\|_{V_0}^2 = T \|u_0^c - u_{0h}^c\|_{V_0}^2 + \frac{T}{2} \sum_{k=1}^N \|\mathbf{u}_k - \mathbf{u}_{kh}\|_{V_0}^2.$$

Remember that our decoupled variational problems for $k = 1, \dots, N$ are given by (3.6), i.e., find $\mathbf{u}_k \in \mathbb{V} = V \times V$ such that

$$a_k(\mathbf{u}_k, \mathbf{v}_k) = \int_{\Omega} \mathbf{f}_k \cdot \mathbf{v}_k \, d\mathbf{x} =: \langle F_k, \mathbf{v}_k \rangle \quad \forall \mathbf{v}_k \in \mathbb{V}.$$

The variational problem for $k = 0$ is given by (3.7), i.e., find $u_0^c \in V$ such that

$$a_0(u_0^c, v_0^c) = \int_{\Omega} f_0^c(\mathbf{x}) v_0^c(\mathbf{x}) \, d\mathbf{x} =: \langle F_0, v_0^c \rangle \quad \forall v_0^c \in V,$$

where the bilinear form $a_0(\cdot, \cdot)$ is defined in (3.9). Since the Fourier coefficients of the given data f can be computed exactly, the corresponding discrete problems are given by: Find $\mathbf{u}_{kh} \in \mathbb{V}_h$ such that

$$a_k(\mathbf{u}_{kh}, \mathbf{v}_{kh}) = \langle F_k, \mathbf{v}_{kh} \rangle \quad \forall \mathbf{v}_{kh} \in \mathbb{V}_h,$$

which is equivalent to solving the linear system (3.26). Since $\mathbb{V}_h \subset \mathbb{V}$, we have the Galerkin orthogonality

$$a_k(\mathbf{u}_k - \mathbf{u}_{kh}, \mathbf{v}_{kh}) = 0 \quad \forall \mathbf{v}_{kh} \in \mathbb{V}_h. \quad (3.41)$$

For the case $k = 0$, we analogously obtain the following discrete problem: Find $u_{0h}^c \in V_h$ such that

$$a_0(u_{0h}^c, v_{0h}^c) = \langle F_0, v_{0h}^c \rangle \quad \forall v_{0h}^c \in V_h,$$

and the Galerkin orthogonality

$$a_0(u_0^c - u_{0h}^c, v_{0h}^c) = 0 \quad \forall v_{0h}^c \in V_h.$$

The following theorem provides an estimate for the discretization error between the unknown Fourier coefficients and their finite element approximations. To begin with, we prove that the discretization error of the Fourier coefficients can be estimated by the best approximation error. Afterwards we estimate the best approximation error by the interpolation error provided the Fourier coefficients are sufficiently smooth, see [42].

Theorem 3.17. *Under the assumption that $\mathbf{u}_k \in (H^2(\Omega))^2$ for all $k = 1, \dots, N$, the discretization error for the Fourier coefficients can be estimated by*

$$\|\mathbf{u}_k - \mathbf{u}_{kh}\|_{V_0} \leq c_1 c_{par}(k, \omega, \bar{\nu}, \bar{\sigma}, h) h |\mathbf{u}_k|_{H^2(\Omega)}, \quad (3.42)$$

where $c_{par}^2(k, \omega, \bar{\nu}, \bar{\sigma}, h) = \bar{\nu} c_{1,2}^2 + k \omega \bar{\sigma} c_{0,2}^2 h^2$ with constants $c_{0,2}$ and $c_{1,2}$ from the approximation theorem, c_1 is a positive constant, and $|\cdot|_{H^2(\Omega)}$ is the $H^2(\Omega)$ -seminorm. Moreover, if $u_0^c \in H^2(\Omega)$, then

$$\|u_0^c - u_{0h}^c\|_{V_0} \leq \sqrt{\bar{\nu}} c_{1,2} h |u_0^c|_{H^2(\Omega)} \quad (3.43)$$

with the constant $c_{1,2}$ coming from the approximation theorem.

Proof. Let us start with the case $k = 1, \dots, N$. Inserting an arbitrary $\mathbf{v}_{kh} \in \mathbb{V}_h$ and using triangle inequality, the discrete version of the inf-sup condition in (3.34) of the Babuška-Aziz theorem with the constant $\underline{c} = 1/\sqrt{2}$ as well as the sup-sup condition with constant $\bar{c} = 1$ together with the Galerkin orthogonality (3.41), we obtain the following Céa-type estimate according to (2.26):

$$\begin{aligned} \|\mathbf{u}_k - \mathbf{u}_{kh}\|_{V_0} &\leq \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0} + \|\mathbf{u}_{kh} - \mathbf{v}_{kh}\|_{V_0} \leq \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0} + \frac{1}{\underline{c}} \sup_{0 \neq \tilde{\mathbf{v}}_{kh} \in \mathbb{V}_h} \frac{a_k(\mathbf{u}_{kh} - \mathbf{v}_{kh}, \tilde{\mathbf{v}}_{kh})}{\|\tilde{\mathbf{v}}_{kh}\|_{V_0}} \\ &\leq \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0} + \underbrace{\frac{1}{\underline{c}} \sup_{0 \neq \tilde{\mathbf{v}}_{kh} \in \mathbb{V}_h} \frac{a_k(\mathbf{u}_{kh} - \mathbf{u}_k, \tilde{\mathbf{v}}_{kh})}{\|\tilde{\mathbf{v}}_{kh}\|_{V_0}}}_{=0} + \frac{1}{\underline{c}} \sup_{0 \neq \tilde{\mathbf{v}}_{kh} \in \mathbb{V}_h} \frac{a_k(\mathbf{u}_k - \mathbf{v}_{kh}, \tilde{\mathbf{v}}_{kh})}{\|\tilde{\mathbf{v}}_{kh}\|_{V_0}} \\ &\leq \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0} + \frac{1}{\underline{c}} \cdot \bar{c} \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0} \leq \underbrace{\left(1 + \frac{\bar{c}}{\underline{c}}\right)}_{=: c_1} \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0}. \end{aligned}$$

So, we can estimate the discretization error by the best approximation error, i.e.,

$$\|\mathbf{u}_k - \mathbf{u}_{kh}\|_{V_0} \leq c_1 \inf_{\mathbf{v}_{kh} \in \mathbb{V}_h} \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0}. \quad (3.44)$$

Thus, the best approximation error can be estimated by the interpolation error, i.e.,

$$\inf_{\mathbf{v}_{kh} \in \mathbb{V}_h} \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0} \leq \|\mathbf{u}_k - I_h \mathbf{u}_k\|_{V_0}, \quad (3.45)$$

where $I_h : \mathbb{V} \rightarrow \mathbb{V}_h$ is some interpolation operator. The V_0 -norm is bounded by

$$\begin{aligned} \|\mathbf{u}_k\|_{V_0}^2 &= (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)} \\ &\leq \bar{\nu} |\mathbf{u}_k|_{H^1(\Omega)}^2 + k\omega \bar{\sigma} \|\mathbf{u}_k\|_{L^2(\Omega)}^2. \end{aligned}$$

Under the assumption that the Fourier coefficients are from $H^2(\Omega)$, the interpolation error can be estimated by

$$\begin{aligned} \|\mathbf{u}_k - I_h \mathbf{u}_k\|_{V_0}^2 &= \|(I - I_h) \mathbf{u}_k\|_{V_0}^2 \leq \bar{\nu} |(I - I_h) \mathbf{u}_k|_{H^1(\Omega)}^2 + k\omega \bar{\sigma} \|(I - I_h) \mathbf{u}_k\|_{L^2(\Omega)}^2 \\ &\leq \bar{\nu} c_{1,2}^2 h^2 |\mathbf{u}_k|_{H^2(\Omega)}^2 + k\omega \bar{\sigma} c_{0,2}^2 h^4 |\mathbf{u}_k|_{H^2(\Omega)}^2 \\ &= \underbrace{(\bar{\nu} c_{1,2}^2 + k\omega \bar{\sigma} c_{0,2}^2 h^2)}_{=: c_{par}^2(k, \omega, \bar{\nu}, \bar{\sigma}, h)} h^2 |\mathbf{u}_k|_{H^2(\Omega)}^2, \end{aligned}$$

where $c_{0,2}$ and $c_{1,2}$ are constants coming from applying the approximation theorem from finite element discretization theory, see, e.g., [42, 46]. Thus, we have

$$\|\mathbf{u}_k - I_h \mathbf{u}_k\|_{V_0} \leq c_{par}(k, \omega, \bar{\nu}, \bar{\sigma}, h) h |\mathbf{u}_k|_{H^2(\Omega)}.$$

Altogether, the discretization error for the Fourier coefficients corresponding to the modes $k = 1, \dots, N$ can be estimated by

$$\begin{aligned} \|\mathbf{u}_k - \mathbf{u}_{kh}\|_{V_0} &\leq c_1 \inf_{\mathbf{v}_{kh} \in \mathbb{V}_h} \|\mathbf{u}_k - \mathbf{v}_{kh}\|_{V_0} \\ &\leq c_1 \|\mathbf{u}_k - I_h \mathbf{u}_k\|_{V_0} \leq c_1 c_{par}(k, \omega, \bar{\nu}, \bar{\sigma}, h) h |\mathbf{u}_k|_{H^2(\Omega)} \end{aligned} \quad (3.46)$$

with the constant $c_1 = 1 + \sqrt{2}$.

The error for the case $k = 0$ can be similarly estimated. The bilinear form $a_0(\cdot, \cdot)$ is bounded and elliptic in the V_0 -norm with boundedness and ellipticity constants $\bar{c} = \underline{c} = 1$, since, e.g., $a_0^c(u_0^c, u_0^c) = \|u_0^c\|_{V_0}^2$. Hence, we can estimate the discretization error by the best approximation error, i.e.,

$$\|u_0^c - u_{0h}^c\|_{V_0} \leq \inf_{v_{0h}^c \in V_h} \|u_0^c - v_{0h}^c\|_{V_0},$$

due to the C ea lemma, see Lemma 2.14 and Remark 2.15. Thus, the best approximation error can be estimated by the interpolation error, i.e.,

$$\inf_{v_{0h}^c \in V_h} \|u_0^c - v_{0h}^c\|_{V_0} \leq \|u_0^c - I_h u_0^c\|_{V_0},$$

where $I_h : V \rightarrow V_h$ is again some interpolation operator. Under the assumption that $u_0^c \in H^2(\Omega)$, the interpolation error can be estimated by

$$\|u_0^c - I_h u_0^c\|_{V_0}^2 = \|(I - I_h) u_0^c\|_{V_0}^2 \leq \bar{\nu} |(I - I_h) u_0^c|_{H^1(\Omega)}^2 \leq \bar{\nu} c_{1,2}^2 h^2 |u_0^c|_{H^2(\Omega)}^2,$$

where $c_{1,2}$ is again a constant coming from applying the approximation theorem from finite element discretization theory, see, e.g., [42, 46]. Altogether, we have

$$\|u_0^c - u_{0h}^c\|_{V_0} \leq \sqrt{\bar{\nu}} c_{1,2} h |u_0^c|_{H^2(\Omega)},$$

which completes the proof. \square

Let us now define the $H^{2,0}(Q_T)$ -seminorm in the Fourier space as follows

$$|u|_{H^2} = \left(T |u_0^c|_{H^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} |\mathbf{u}_k|_{H^2(\Omega)}^2 \right)^{1/2}, \quad (3.47)$$

where the H^2 -seminorm for the Fourier coefficients is the usual $H^2(\Omega)$ -seminorm as used in Theorem 3.17. The following theorem provides the estimate for the complete discretization error with respect to the spatial discretization parameter h .

Theorem 3.18. *Under the assumptions of Theorem 3.17, the discretization error in h can be estimated as follows*

$$\|u_N - u_{Nh}\|_{V_0} \leq c_1 c_{par}(N, \omega, \bar{\nu}, \bar{\sigma}, h) h |u_N|_{H^2}, \quad (3.48)$$

where $c_{par}^2(N, \omega, \bar{\nu}, \bar{\sigma}, h) = \bar{\nu} c_{1,2}^2 + N \omega \bar{\sigma} c_{0,2}^2 h^2$ with constants $c_{0,2}$ and $c_{1,2}$ from the approximation theorem, and c_1 is a positive constant. The H^2 -seminorm of u_N is given by

$$|u_N|_{H^2}^2 = T |u_0^c|_{H^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N |\mathbf{u}_k|_{H^2(\Omega)}^2.$$

Proof. Due to Theorem 3.17, we obtain the estimate

$$\begin{aligned} \|u_N - u_{Nh}\|_{V_0}^2 &= T \|u_0^c - u_{0h}^c\|_{V_0}^2 + \frac{T}{2} \sum_{k=1}^N \|\mathbf{u}_k - \mathbf{u}_{kh}\|_{V_0}^2 \\ &\leq T \bar{\nu} c_{1,2}^2 h^2 |u_0^c|_{H^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N c_1^2 c_{par}^2(k, \omega, \bar{\nu}, \bar{\sigma}, h) h^2 |\mathbf{u}_k|_{H^2(\Omega)}^2 \\ &\leq T \bar{\nu} c_{1,2}^2 h^2 |u_0^c|_{H^2(\Omega)}^2 + c_1^2 c_{par}^2(N, \omega, \bar{\nu}, \bar{\sigma}, h) h^2 \frac{T}{2} \sum_{k=1}^N |\mathbf{u}_k|_{H^2(\Omega)}^2 \\ &\leq c_1^2 c_{par}^2(N, \omega, \bar{\nu}, \bar{\sigma}, h) h^2 \left(T |u_0^c|_{H^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N |\mathbf{u}_k|_{H^2(\Omega)}^2 \right), \end{aligned}$$

which completes the proof. \square

3.4.3 Complete discretization error

Now, we are in the position to state the final discretization error estimate.

Theorem 3.19. *Let us assume that $u \in V_1 \cap H^{2,0}(Q_T)$. Then the complete discretization error arising from the multiharmonic finite element discretization can be estimated as follows*

$$\|u - u_{Nh}\|_{V_0} \leq c_0 N^{-1/2} \|u\|_{V_1} + c_1 c_{par}(N, \omega, \bar{\nu}, \bar{\sigma}, h) h |u|_{H^2},$$

where c_0 and c_1 come from Theorem 3.16 and Theorem 3.18, respectively, and $c_{par}^2(N, \omega, \bar{\nu}, \bar{\sigma}, h) = \bar{\nu} c_{1,2}^2 + N \omega \bar{\sigma} c_{0,2}^2 h^2$ with constants $c_{0,2}$ and $c_{1,2}$ from the approximation theorem.

Proof. Applying the triangle inequality and using Theorems 3.16 and 3.18 yield the estimates

$$\|u - u_{Nh}\|_{V_0} \leq \|u - u_N\|_{V_0} + \|u_N - u_{Nh}\|_{V_0} \leq c_0 N^{-1/2} \|u\|_{V_1} + c_1 c_{par}(N, \omega, \bar{\nu}, \bar{\sigma}, h) h |u_N|_{H^2},$$

where the seminorm $|u_N|_{H^2}$ can trivially be estimated by (3.47). \square

Remark 3.20. *The convergence rate with respect to the spatial discretization parameter h reduces from h to h^s with some $s \in (0, 1)$, if*

$$u \in V_1 \cap H^{1+s,0}(Q_T).$$

In order to get h^s with $s > 1$, we need higher order elements. On the other side the convergence with respect to N will improve, if u is smoother with respect to the time variable. More precisely, the factor $N^{-1/2}$ improves to $N^{-\ell/2}$ provided that

$$u \in V_\ell \cap H^{1+s,0}(Q_T),$$

where $V_\ell := (H^{0,\ell/2})_0^1(Q_T) \cap H_{per}^{0,(\ell+1)/2}(Q_T)$ and with some $\ell > 1$.

Remark 3.21. *If the given right-hand side f has a multiharmonic representation, then the solution u has a multiharmonic representation as well and the complete discretization error reduces to a discretization error in the spatial variable h and can be estimated as in Theorem 3.18.*

Chapter 4

Multiharmonic finite element analysis of parabolic time-periodic optimal control problems

This chapter considers optimal control problems, where parabolic time-periodic partial differential equations of the form (3.1)-(3.3) appear in their constraints. The multiharmonic finite element analysis of these parabolic time-periodic partial differential equations including existence and uniqueness results has already been presented in Chapter 3. Following these ideas and results, we now discuss the existence and uniqueness of parabolic time-periodic optimal control problems as well as present block-diagonal preconditioned MINRES solvers and full a priori error estimates.

4.1 A parabolic time-periodic optimal control problem

Let us denote the state of our optimal control problem by y and the control by u . The spatial domain $\Omega \subset \mathbb{R}^d$ is assumed to be a bounded Lipschitz domain with the boundary $\Gamma := \partial\Omega$, where $d = \{1, 2, 3\}$. Moreover, the space-time cylinder is again denoted by $Q_T := \Omega \times (0, T)$ and its mantle boundary by $\Sigma_T := \Gamma \times (0, T)$. We consider the following parabolic time-periodic optimal control problem:

$$\min_{y,u} \mathcal{J}(y, u) := \frac{1}{2} \int_0^T \int_{\Omega} (y(\mathbf{x}, t) - y_d(\mathbf{x}, t))^2 dx dt + \frac{\lambda}{2} \int_0^T \int_{\Omega} (u(\mathbf{x}, t))^2 dx dt \quad (4.1)$$

subject to the parabolic time-periodic boundary value problem (3.1)-(3.3), i.e.,

$$\begin{aligned} \sigma(\mathbf{x}) \partial_t y(\mathbf{x}, t) - \operatorname{div}(\nu(\mathbf{x}) \nabla y(\mathbf{x}, t)) &= u(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ y(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ y(\mathbf{x}, 0) &= y(\mathbf{x}, T) & \mathbf{x} \in \bar{\Omega}, \end{aligned} \quad (4.2)$$

with uniformly bounded coefficients $\sigma(\cdot)$ and $\nu(\cdot)$ satisfying the assumptions (2.29). The desired state y_d is the given target that we try to reach via a suitable control u . The positive regularization parameter λ provides a weighting of the cost of the control in the cost functional $\mathcal{J}(\cdot, \cdot)$.

In this section, we discuss existence and uniqueness of the parabolic time-periodic optimal control problem (4.1)-(4.2) as well as formulate its optimality system that we are going to discretize by the multiharmonic finite element method.

In Chapter 3, we have already proved the existence and uniqueness of a solution of certain variational formulations, i.e., (3.4), (3.5) and (3.16), of the parabolic time-periodic boundary value problem (4.2). In particular, remember the existence and uniqueness result of Theorem 3.6, more precisely, that the

variational problem (3.16) has a unique solution due to Lemma 3.5, which exactly are the inf-sup and sup-sup conditions of the Babuška-Aziz theorem for this variational problem.

Theorem 4.1. *The parabolic time-periodic optimal control problem (4.1)-(4.2) has a unique solution $(\bar{y}, \bar{u}) \in H_0^{1, \frac{1}{2}}(Q_T) \times L^2(Q_T)$.*

Proof. Theorem 3.6 implies the existence of a linear and continuous solution operator which uniquely assigns a state

$$y \in H_0^{1, \frac{1}{2}}(Q_T)$$

to every control $u \in L^2(Q_T)$, where the space $H_0^{1, \frac{1}{2}}(Q_T)$ is compactly embedded in $L^2(Q_T)$. With the solution operator, the optimal control problem (4.1)-(4.2) can be rewritten as a reduced (weak) minimization problem. Under the assumptions that $y_d \in L^2(Q_T)$ and $\lambda > 0$, Theorem 2.21 finally yields the existence and uniqueness result. \square

We want to formulate now the optimality system. Its solution is equivalent to the solution of the original optimal control problem (4.1)-(4.2). We denote the Lagrange multiplier by p , which is also referred as the adjoint state. We choose the following Lagrange functional for our minimization problem:

$$\mathcal{L}(y, u, p) := \mathcal{J}(y, u) - \int_0^T \int_{\Omega} (\sigma \partial_t y - \operatorname{div}(\nu \nabla y) - u)p \, dx \, dt. \quad (4.3)$$

The optimality system is given by (2.42), i.e.,

$$\begin{aligned} \mathcal{L}_p(y, u, p) &= 0, \\ \mathcal{L}_y(y, u, p) &= 0, \\ \mathcal{L}_u(y, u, p) &= 0, \end{aligned} \quad (4.4)$$

and characterizes a stationary point (y, u, p) of the Lagrange functional (4.3). Using the second condition, we can eliminate the control u from the optimality system (4.4), i.e.,

$$u = -\lambda^{-1}p \quad \text{in } Q_T. \quad (4.5)$$

From (4.5) it appears very natural to choose y , p and also u from the same space. Moreover, we arrive at a reduced optimality system, written in its classical formulation as

$$\begin{aligned} \sigma(\mathbf{x}) \partial_t y(\mathbf{x}, t) - \operatorname{div}(\nu(\mathbf{x}) \nabla y(\mathbf{x}, t)) &= -\lambda^{-1}p(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ y(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ y(\mathbf{x}, 0) &= y(\mathbf{x}, T) & \mathbf{x} \in \bar{\Omega}, \\ -\sigma(\mathbf{x}) \partial_t p(\mathbf{x}, t) - \operatorname{div}(\nu(\mathbf{x}) \nabla p(\mathbf{x}, t)) &= y(\mathbf{x}, t) - y_d(\mathbf{x}, t) & (\mathbf{x}, t) \in Q_T, \\ p(\mathbf{x}, t) &= 0 & (\mathbf{x}, t) \in \Sigma_T, \\ p(\mathbf{x}, T) &= p(\mathbf{x}, 0) & \mathbf{x} \in \bar{\Omega}. \end{aligned} \quad (4.6)$$

The space-time variational formulation of (4.6) is obtained in the same way as for the parabolic time-periodic partial differential equation (3.16) in Chapter 3 and is stated as the following: Given the desired state $y_d \in L^2(Q_T)$, find y and p from $H_0^{1, \frac{1}{2}}(Q_T)$ such that

$$\begin{aligned} \int_0^T \int_{\Omega} \left(y v - \nu(\mathbf{x}) \nabla p \cdot \nabla v + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} v^\perp \right) dx \, dt &= \int_0^T \int_{\Omega} y_d v \, dx \, dt, \\ \int_0^T \int_{\Omega} \left(\nu(\mathbf{x}) \nabla y \cdot \nabla q + \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} q^\perp + \lambda^{-1} p q \right) dx \, dt &= 0, \end{aligned} \quad (4.7)$$

for all test functions $v, q \in H_0^{1, \frac{1}{2}}(Q_T)$, where all functions are given in their Fourier series expansion in time according to Definition 3.2.

Remark 4.2. In Chapter 6, we show that problem (4.7) has a unique solution based on the fact that the corresponding inf-sup and sup-sup conditions are fulfilled and, hence, the Babuška-Aziz theorem can be applied, see Lemma 6.23 as well as Lemma 6.25. These lemmas provide additional existence and uniqueness results to the one obtained in Theorem 4.1.

4.2 Multiharmonic finite element discretization

In order to solve the optimal control problem (4.1)-(4.2), we discretize the optimality system (4.7) by the multiharmonic finite element method. Again, choosing test functions (3.23) yields variational problems, which correspond to every mode $k = 0, \dots, N$. On the other hand, we can use the equivalent approach and approximate the desired state y_d by truncating its Fourier series expansion, i.e.,

$$y_d(\mathbf{x}, t) \approx y_{d0}^c(\mathbf{x}) + \sum_{k=1}^N [y_{dk}^c(\mathbf{x}) \cos(k\omega t) + y_{dk}^s(\mathbf{x}) \sin(k\omega t)] = y_{dN}(\mathbf{x}, t), \quad (4.8)$$

where its Fourier coefficients are given by

$$\begin{aligned} y_{d0}^c(\mathbf{x}) &= \frac{1}{T} \int_0^T y_d(\mathbf{x}, t) dt, \\ y_{dk}^c(\mathbf{x}) &= \frac{2}{T} \int_0^T y_d(\mathbf{x}, t) \cos(k\omega t) dt, \\ y_{dk}^s(\mathbf{x}) &= \frac{2}{T} \int_0^T y_d(\mathbf{x}, t) \sin(k\omega t) dt. \end{aligned}$$

We mention here again that in this work we consider only the case where we can compute the Fourier coefficients exactly.

We insert the truncated desired state (4.8) and the Fourier series ansatz of the state y and the adjoint state p into the space-time variational formulation (4.7). From the orthogonality of the functions $\cos(k\omega t)$ and $\sin(k\omega t)$ it follows that it is sufficient to consider only the truncated Fourier series of y and p , i.e.,

$$\begin{aligned} y(\mathbf{x}, t) &\approx y_0^c(\mathbf{x}) + \sum_{k=1}^N [y_k^c(\mathbf{x}) \cos(k\omega t) + y_k^s(\mathbf{x}) \sin(k\omega t)] = y_N(\mathbf{x}, t), \\ p(\mathbf{x}, t) &\approx p_0^c(\mathbf{x}) + \sum_{k=1}^N [p_k^c(\mathbf{x}) \cos(k\omega t) + p_k^s(\mathbf{x}) \sin(k\omega t)] = p_N(\mathbf{x}, t), \end{aligned}$$

and we arrive at the following system which has to be solved for every mode $k = 1, 2, \dots, N$: Find $\mathbf{y}_k, \mathbf{p}_k \in \mathbb{V} = V \times V = (H_0^1(\Omega))^2$ such that

$$\begin{aligned} \int_{\Omega} (\mathbf{y}_k \cdot \mathbf{v}_k - \nu(\mathbf{x}) \nabla \mathbf{p}_k \cdot \nabla \mathbf{v}_k + k\omega \sigma(\mathbf{x}) \mathbf{p}_k \cdot \mathbf{v}_k^\perp) d\mathbf{x} &= \int_{\Omega} \mathbf{y}_{dk} \cdot \mathbf{v}_k d\mathbf{x}, \\ \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{y}_k \cdot \nabla \mathbf{q}_k + k\omega \sigma(\mathbf{x}) \mathbf{y}_k \cdot \mathbf{q}_k^\perp + \lambda^{-1} \mathbf{p}_k \cdot \mathbf{q}_k) d\mathbf{x} &= 0, \end{aligned} \quad (4.9)$$

for all test functions $\mathbf{v}_k, \mathbf{q}_k \in \mathbb{V}$. In the case of $k = 0$, we obtain the following optimality system: Find $y_0^c, p_0^c \in V = H_0^1(\Omega)$ such that

$$\begin{aligned} \int_{\Omega} (y_0^c \cdot v_0^c - \nu(\mathbf{x}) \nabla p_0^c \cdot \nabla v_0^c) d\mathbf{x} &= \int_{\Omega} y_{d0}^c \cdot v_0^c d\mathbf{x}, \\ \int_{\Omega} (\nu(\mathbf{x}) \nabla y_0^c \cdot \nabla q_0^c + \lambda^{-1} p_0^c \cdot q_0^c) d\mathbf{x} &= 0, \end{aligned} \quad (4.10)$$

for all test functions $v_0^c, q_0^c \in V$. Analogously to Chapter 3, we approximate the unknown Fourier coefficients

$$\mathbf{y}_k = (y_k^c, y_k^s)^T, \mathbf{p}_k = (p_k^c, p_k^s)^T \in \mathbb{V}$$

by finite element functions

$$\mathbf{y}_{kh} = (y_{kh}^c, y_{kh}^s)^T, \mathbf{p}_{kh} = (p_{kh}^c, p_{kh}^s)^T \in \mathbb{V}_h = V_h \times V_h \subset \mathbb{V},$$

where $\mathbb{V}_h = V_h \times V_h$ is a finite element space with

$$V_h = \text{span}\{\varphi_1, \dots, \varphi_n\}$$

and $\{\varphi_i(x) = \varphi_{ih}(x) : i = 1, 2, \dots, n_h\}$ is the standard nodal basis. Again, we denote by h the usual discretization parameter such that $n = n_h = \dim V_h = O(h^{-d})$ and use continuous, piecewise linear finite elements on the finite elements on a regular triangulation \mathcal{T}_h to construct the finite element subspace V_h and its basis, see, e.g., [41, 46, 84, 161] as well as Sections 2.4 and 3.2. This leads to a linear system arising from the variational formulation (4.9), i.e.,

$$\begin{pmatrix} M_h & 0 & -K_{h,\nu} & k\omega M_{h,\sigma} \\ 0 & M_h & -k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} & -\lambda^{-1} M_h & 0 \\ k\omega M_{h,\sigma} & -K_{h,\nu} & 0 & -\lambda^{-1} M_h \end{pmatrix} \begin{pmatrix} \underline{y}_k^c \\ \underline{y}_k^s \\ \underline{p}_k^c \\ \underline{p}_k^s \end{pmatrix} = \begin{pmatrix} \underline{y}_{dk}^c \\ \underline{y}_{dk}^s \\ 0 \\ 0 \end{pmatrix} \quad (4.11)$$

for $k = 1, 2, \dots, N$, which has to be solved with respect to the nodal parameter vectors

$$\underline{y}_k^j = (y_{k,i}^j)_{i=1,\dots,n} \in \mathbb{R}^n \quad \text{and} \quad \underline{p}_k^j = (p_{k,i}^j)_{i=1,\dots,n} \in \mathbb{R}^n$$

of the finite element approximations

$$y_{kh}^j(x) = \sum_{i=1}^n y_{k,i}^j \varphi_i(x) \quad \text{and} \quad p_{kh}^j(x) = \sum_{i=1}^n p_{k,i}^j \varphi_i(x)$$

to the unknown Fourier coefficients $y_k^j(x)$ and $p_k^j(x)$ with $j \in \{c, s\}$. The matrices M_h , $M_{h,\sigma}$ and $K_{h,\nu}$ correspond to the mass matrix, the weighted mass matrix and the stiffness matrix, respectively. Their entries are computed by the following formulas:

$$M_h^{ij} = \int_{\Omega} \varphi_i \varphi_j \, d\mathbf{x}, \quad M_{h,\sigma}^{ij} = \int_{\Omega} \sigma \varphi_i \varphi_j \, d\mathbf{x}, \quad K_{h,\nu}^{ij} = \int_{\Omega} \nu \nabla \varphi_i \cdot \nabla \varphi_j \, d\mathbf{x},$$

with $i, j = 1, \dots, n$, whereas

$$\underline{y}_{dk}^c = \left[\int_{\Omega} y_{dk}^c \varphi_j \, d\mathbf{x} \right]_{j=1,\dots,n} \quad \text{and} \quad \underline{y}_{dk}^s = \left[\int_{\Omega} y_{dk}^s \varphi_j \, d\mathbf{x} \right]_{j=1,\dots,n}.$$

In the case $k = 0$, we obtain the following linear system arising from the variational problem (4.10):

$$\begin{pmatrix} M_h & -K_{h,\nu} \\ -K_{h,\nu} & -\lambda^{-1} M_h \end{pmatrix} \begin{pmatrix} \underline{y}_0^c \\ \underline{p}_0^c \end{pmatrix} = \begin{pmatrix} \underline{y}_{d0}^c \\ 0 \end{pmatrix}. \quad (4.12)$$

Finally, we can easily reconstruct the multiharmonic finite element approximations

$$y_{Nh}(\mathbf{x}, t) = y_{0h}^c(\mathbf{x}) + \sum_{k=1}^N [y_{kh}^c(\mathbf{x}) \cos(k\omega t) + y_{kh}^s(\mathbf{x}) \sin(k\omega t)] \quad (4.13)$$

and

$$p_{Nh}(\mathbf{x}, t) = p_{0h}^c(\mathbf{x}) + \sum_{k=1}^N [p_{kh}^c(\mathbf{x}) \cos(k\omega t) + p_{kh}^s(\mathbf{x}) \sin(k\omega t)] \quad (4.14)$$

to the state $y(\mathbf{x}, t)$ and the adjoint state $p(\mathbf{x}, t)$ from the solutions of the linear systems (4.11) and (4.12). Analogously to Chapter 3, we will present an a priori error analysis for the complete discretization error between the unknown solution (y, p) of the reduced optimality system (4.7) and its multiharmonic finite element approximation (y_{Nh}, p_{Nh}) , see Section 4.4.

Remark 4.3. *We can also insert the Fourier series ansatz immediately into the optimal control problem (4.1)-(4.2) - before formulating its optimality system. So after inserting the Fourier series ansatz (2.10) for all functions, we obtain the following optimal control problems for every mode $k = 1, 2, \dots$:*

$$\min_{\mathbf{y}_k, \mathbf{u}_k} \mathcal{J}_k(\mathbf{y}_k, \mathbf{u}_k) := \frac{1}{2} \int_{\Omega} (\mathbf{y}_k(\mathbf{x}) - \mathbf{y}_{dk}(\mathbf{x}))^2 d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} (\mathbf{u}_k(\mathbf{x}))^2 d\mathbf{x} \quad (4.15)$$

subject to the boundary value problem

$$\begin{aligned} -k\omega \sigma(\mathbf{x}) \mathbf{y}_k^\perp(\mathbf{x}) - \operatorname{div}(\nu(\mathbf{x}) \nabla \mathbf{y}_k(\mathbf{x}, t)) &= \mathbf{u}_k(\mathbf{x}) & \mathbf{x} \in \Omega, \\ \mathbf{y}_k(\mathbf{x}) &= 0 & \mathbf{x} \in \Gamma. \end{aligned} \quad (4.16)$$

The variational problem of (4.16) can be formulated as (3.6). For the case $k = 0$, we obtain the optimal control problem

$$\min_{y_0^c, u_0^c} \mathcal{J}_0(y_0^c, u_0^c) := \frac{1}{2} \int_{\Omega} (y_0^c(\mathbf{x}) - y_{d0}^c(\mathbf{x}))^2 d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} (u_0^c(\mathbf{x}))^2 d\mathbf{x} \quad (4.17)$$

subject to the boundary value problem

$$\begin{aligned} -\operatorname{div}(\nu(\mathbf{x}) \nabla y_0^c(\mathbf{x}, t)) &= u_0^c(\mathbf{x}) & \mathbf{x} \in \Omega, \\ y_0^c(\mathbf{x}) &= 0 & \mathbf{x} \in \Gamma, \end{aligned} \quad (4.18)$$

which leads to variational problem (3.7).

Formulating the optimality systems of problems (4.15)-(4.16) and (4.17)-(4.18), and, then, discretizing it by the finite element method leads to the same linear systems of equations (4.11) and (4.12).

4.3 Block-diagonal preconditioned MINRES solver

The resulting linear system (4.11) as well as the system (4.12) are saddle point problems of the form (2.46) with the system matrix (2.47), i.e.,

$$\mathcal{A}u = f, \quad (4.19)$$

where

$$\mathcal{A} := \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}, \quad u := \begin{pmatrix} y \\ p \end{pmatrix} \quad \text{and} \quad f := \begin{pmatrix} y_d \\ 0 \end{pmatrix}.$$

In the case $k = 1, \dots, N$, we have that

$$A := \begin{pmatrix} M_h & 0 \\ 0 & M_h \end{pmatrix}, \quad B := \begin{pmatrix} -K_{h,\nu} & -k\omega M_{h,\sigma} \\ k\omega M_{h,\sigma} & -K_{h,\nu} \end{pmatrix}, \quad C := \lambda^{-1}A$$

and

$$\underline{y}_d := \begin{pmatrix} \underline{y}_{dk}^c \\ \underline{y}_{dk}^s \end{pmatrix}, \quad \underline{y} := \begin{pmatrix} \underline{y}_k^c \\ \underline{y}_k^s \end{pmatrix}, \quad \underline{p} := \begin{pmatrix} \underline{p}_k^c \\ \underline{p}_k^s \end{pmatrix}.$$

These saddle point problems can be solved by a preconditioned MINRES method, see Section 2.7, where a convergence result for this method is stated in Theorem 2.25. Hence, it is crucial to construct preconditioners, which yield robust and fast convergence for the preconditioned MINRES method. This will be done analogously as in Section 3.3. We start with an easier case by assuming that the parameter σ is constant and construct preconditioners following the strategy presented in Zulehner [187], which is based on space interpolation theory. Motivated by the resulting preconditioner, we choose an initial guess for a preconditioner in the more general case of $\sigma(\cdot)$ being piecewise constant. By introducing proper parameter dependent norms, we verify the assumptions of the Babuška-Aziz theorem, which finally yields a parameter robust convergence rate as desired, see [89]. Let us start by assuming that $\sigma(\cdot)$ is constant. Hence, in this case, we have

$$M_{h,\sigma} = \sigma M_h.$$

Then, the system matrix \mathcal{A} in the linear system (4.19) is given by the block matrices

$$A := \begin{pmatrix} M_h & 0 \\ 0 & M_h \end{pmatrix}, \quad B := \begin{pmatrix} -K_{h,\nu} & -k\omega\sigma M_h \\ k\omega\sigma M_h & -K_{h,\nu} \end{pmatrix}, \quad C := \lambda^{-1}A.$$

Due to Theorem 2.25, the convergence rate of the preconditioned MINRES method only depends on the condition number of the preconditioned system. Hence, we are going to construct preconditioners for the preconditioned MINRES method such that the condition number $\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A})$ of the preconditioned system $\mathcal{P}^{-1}\mathcal{A}$ is independent of all “bad” parameters, i.e.,

$$h, \quad N, \quad \omega, \quad \lambda, \quad \nu, \quad \sigma.$$

In order to obtain parameter robust convergence rates, we first construct block-diagonal preconditioners by the operator matrix interpolation technique presented in Section 2.7. From Theorem 2.26 follows that \mathcal{A} can be preconditioned by

$$\mathcal{P}_0 = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \quad \text{and} \quad \mathcal{P}_1 = \begin{pmatrix} R & 0 \\ 0 & C \end{pmatrix},$$

where the negative Schur complements are given by

$$S = \begin{pmatrix} K_{h,\nu}M_h^{-1}K_{h,\nu} + (k^2\omega^2\sigma^2 + \lambda^{-1})M_h & 0 \\ 0 & K_{h,\nu}M_h^{-1}K_{h,\nu} + (k^2\omega^2\sigma^2 + \lambda^{-1})M_h \end{pmatrix}$$

and

$$R = \begin{pmatrix} \lambda K_{h,\nu}M_h^{-1}K_{h,\nu} + (k^2\omega^2\sigma^2\lambda + 1)M_h & 0 \\ 0 & \lambda K_{h,\nu}M_h^{-1}K_{h,\nu} + (k^2\omega^2\sigma^2\lambda + 1)M_h \end{pmatrix}$$

in our model problem. Hence,

$$R = \lambda S.$$

Analogously to Chapter 3, we construct block-diagonal preconditioners \mathcal{P}_θ by interpolating between \mathcal{P}_0 and \mathcal{P}_1 as presented in (2.52), from which we can obtain again parameter independent condition number estimates for all $\theta \in [0, 1]$. We choose $\theta = \frac{1}{2}$ and obtain the block-diagonal matrix

$$\mathcal{P}_{1/2} = \begin{pmatrix} [A, R]_{1/2} & 0 \\ 0 & [S, C]_{1/2} \end{pmatrix}$$

with

$$[A, R]_{1/2} = A^{1/2}(A^{-1/2}RA^{-1/2})^{1/2}A^{1/2} \quad \text{and} \quad [S, C]_{1/2} = S^{1/2}(S^{-1/2}CS^{-1/2})^{1/2}S^{1/2}.$$

Since A and R are block-diagonal, the diagonal entries $[A, R]_{1/2}^{(1,1)} = [A, R]_{1/2}^{(2,2)}$ can be estimated from above and from below as follows

$$\begin{aligned} [A, R]_{1/2}^{(1,1)} &= M_h^{1/2}(\lambda M_h^{-1/2}K_{h,\nu}M_h^{-1}K_{h,\nu}M_h^{-1/2} + (k^2\omega^2\sigma^2\lambda + 1)I)^{1/2}M_h^{1/2} \\ &\leq \sqrt{\lambda}M_h^{1/2}(M_h^{-1/2}K_{h,\nu}M_h^{-1}K_{h,\nu}M_h^{-1/2})^{1/2}M_h^{1/2} + \sqrt{k^2\omega^2\sigma^2\lambda + 1}M_h^{1/2}M_h^{1/2} \\ &= \sqrt{\lambda}K_{h,\nu} + \sqrt{k^2\omega^2\sigma^2\lambda + 1}M_h \\ &\leq \sqrt{\lambda}K_{h,\nu} + (k\omega\sigma\sqrt{\lambda} + 1)M_h =: D, \\ [A, R]_{1/2}^{(1,1)} &= M_h^{1/2}(\lambda M_h^{-1/2}K_{h,\nu}M_h^{-1}K_{h,\nu}M_h^{-1/2} + (k^2\omega^2\sigma^2\lambda + 1)I)^{1/2}M_h^{1/2} \\ &\geq M_h^{1/2}\left(\frac{1}{\sqrt{2}}(\sqrt{\lambda}M_h^{-1/2}K_{h,\nu}M_h^{-1/2} + \sqrt{k^2\omega^2\sigma^2\lambda + 1}I)\right)M_h^{1/2} \\ &= \frac{1}{\sqrt{2}}\left(\sqrt{\lambda}K_{h,\nu} + \sqrt{k^2\omega^2\sigma^2\lambda + 1}M_h\right) \\ &\geq \frac{1}{\sqrt{2}}\left(\sqrt{\lambda}K_{h,\nu} + \frac{1}{\sqrt{2}}(k\omega\sigma\sqrt{\lambda} + 1)M_h\right) \\ &\geq \frac{1}{\sqrt{2}}\min\left\{1, \frac{1}{\sqrt{2}}\right\}\left(\sqrt{\lambda}K_{h,\nu} + (k\omega\sigma\sqrt{\lambda} + 1)M_h\right) \\ &= \frac{1}{2}\left(\sqrt{\lambda}K_{h,\nu} + (k\omega\sigma\sqrt{\lambda} + 1)M_h\right) = \frac{1}{2}D, \end{aligned}$$

where we used the spectral inequality (3.29) again. Analogously, since $S = \lambda^{-1}R$ and $C = \lambda^{-1}A$, we have that

$$\begin{aligned} [S, C]_{1/2}^{(1,1)} &= [S, C]_{1/2}^{(2,2)} \\ &= [\lambda^{-1}R, \lambda^{-1}A]_{1/2}^{(1,1)} \\ &= \lambda^{-1}[A, R]_{1/2}^{(1,1)} \\ &\sim \lambda^{-1}(\sqrt{\lambda}K_{h,\nu} + (k\omega\sigma\sqrt{\lambda} + 1)M_h) = \lambda^{-1}D. \end{aligned}$$

Thus, we have obtained a new block-diagonal preconditioner for the MINRES solver of problem (4.11), which we denote by $\mathcal{P}_{1/2}$, and which is given by

$$\mathcal{P}_{1/2} = \begin{pmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & \lambda^{-1}D & 0 \\ 0 & 0 & 0 & \lambda^{-1}D \end{pmatrix}. \quad (4.20)$$

This block-diagonal preconditioner is much easier to realize in practice than the previous preconditioners \mathcal{P}_0 and \mathcal{P}_1 containing Schur complements. Hence, we obtain the estimates (2.53), i.e.,

$$\underline{c}\|u\|_{\mathcal{P}_{1/2}} \leq \|\mathcal{A}u\|_{\mathcal{P}_{1/2}^{-1}} \leq \bar{c}\|u\|_{\mathcal{P}_{1/2}} \quad \forall u \in \mathbb{R}^{4n}, \quad (4.21)$$

which yield a robust estimate of the condition number

$$\kappa_{\mathcal{P}_{1/2}}(\mathcal{P}_{1/2}^{-1}\mathcal{A}) \leq \bar{c}/\underline{c}$$

with constants $\underline{c} = (\sqrt{5} - 1)/2$ and $\bar{c} = (\sqrt{5} + 1)/2$. Therefore, Theorem 2.25 leads to robust convergence rates of the MINRES method.

Let us now consider the case that $\sigma(\cdot)$ is only piecewise constant, but not constant. Moreover, we allow that σ is zero in some regions of the computational domain Ω . This situation is typical in electromagnetics, where σ is nothing but the conductivity that is zero in non-conducting regions. The system matrix \mathcal{A} of (4.11) is now given by

$$\mathcal{A} := \begin{pmatrix} M_h & 0 & -K_{h,\nu} & k\omega M_{h,\sigma} \\ 0 & M_h & -k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} & -\lambda^{-1}M_h & 0 \\ k\omega M_{h,\sigma} & -K_{h,\nu} & 0 & -\lambda^{-1}M_h \end{pmatrix}.$$

Since now

$$M_{h,\sigma} \neq \sigma M_h,$$

the interpolated matrices $[A, R]_{1/2}$ and $[S, C]_{1/2}$ cannot be computed explicitly. However, we get an inspiration for choosing a suitable block-diagonal preconditioner according to the block-diagonal preconditioner $\mathcal{P}_{1/2}$. Replacing σM_h by $M_{h,\sigma}$ in (4.20), we arrive at the new preconditioner

$$\mathcal{P} = \begin{pmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & \lambda^{-1}D & 0 \\ 0 & 0 & 0 & \lambda^{-1}D \end{pmatrix}, \quad (4.22)$$

and so now, the diagonal block D is given by

$$D = \sqrt{\lambda}K_{h,\nu} + k\omega\sqrt{\lambda}M_{h,\sigma} + M_h.$$

This preconditioner \mathcal{P} is our candidate for a robust preconditioner of the system matrix \mathcal{A} . In order to obtain robust norm estimates for the preconditioned system matrix $\mathcal{P}^{-1}\mathcal{A}$, we look again at the Babuška-Aziz theorem. The clue is once more that the norm estimates which have to be proven are equivalent to the assumptions (inf-sup- and sup-sup-conditions) in the Babuška-Aziz theorem that, at the same time, provides existence, uniqueness, as well as a priori and a posteriori error estimates. The assumptions of the Babuška-Aziz theorem yield discretization error estimates, which we are going to present in Section 4.4.

Let us return to the variational formulation (4.9) of the optimality system for each mode $k = 1, \dots, N$, and let us define the corresponding bilinear form

$$\begin{aligned} \mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k)) &:= \int_{\Omega} (\mathbf{y}_k \cdot \mathbf{v}_k - \nu \nabla \mathbf{p}_k \cdot \nabla \mathbf{v}_k + k\omega \sigma \mathbf{p}_k \cdot \mathbf{v}_k^\perp) \, d\mathbf{x} \\ &+ \int_{\Omega} (\nu \nabla \mathbf{y}_k \cdot \nabla \mathbf{q}_k + k\omega \sigma \mathbf{y}_k \cdot \mathbf{q}_k^\perp + \lambda^{-1} \mathbf{p}_k \cdot \mathbf{q}_k) \, d\mathbf{x}. \end{aligned} \quad (4.23)$$

Hence, the variational problem (4.9) reads now as follows: Find $(\mathbf{y}_k, \mathbf{p}_k) \in \mathbb{V}^2 = (H_0^1(\Omega))^4$ such that

$$\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k)) = \int_{\Omega} \mathbf{y}_k \cdot \mathbf{v}_k \, d\mathbf{x} \quad (4.24)$$

for all test functions $(\mathbf{v}_k, \mathbf{q}_k) \in \mathbb{V}^2$. The initial guess (4.22) for the preconditioner \mathcal{P} yields the following definitions of inner products and associated norms. We first define a non-standard (weighted) inner product in $\mathbb{V} = (H_0^1(\Omega))^2$ by

$$(\mathbf{y}_k, \mathbf{v}_k)_{\mathbb{V}} = \sqrt{\lambda} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)} + k\omega \sqrt{\lambda} (\sigma \mathbf{y}_k, \mathbf{v}_k)_{L^2(\Omega)} + (\mathbf{y}_k, \mathbf{v}_k)_{L^2(\Omega)}.$$

The associated norm is then given by

$$\|\mathbf{y}_k\|_{\mathbb{V}}^2 = \sqrt{\lambda} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + k\omega \sqrt{\lambda} (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} + \|\mathbf{y}_k\|_{L^2(\Omega)}^2, \quad (4.25)$$

which differs from the standard H^1 -norms. Finally, we define an inner product in $\mathbb{V}^2 = (H_0^1(\Omega))^4$ by

$$((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k))_{\mathcal{P}} = (\mathbf{y}_k, \mathbf{v}_k)_{\mathbb{V}} + \lambda^{-1}(\mathbf{p}_k, \mathbf{q}_k)_{\mathbb{V}}.$$

The associated norm is given by

$$\|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2 = \|\mathbf{y}_k\|_{\mathbb{V}}^2 + \lambda^{-1}\|\mathbf{p}_k\|_{\mathbb{V}}^2. \quad (4.26)$$

Next, we verify the assumptions (inf-sup- and sup-sup-conditions) of the theorem of Babuška-Aziz, see [89].

Theorem 4.4. *The following inequalities are valid:*

$$\underline{c} \|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}} \leq \sup_{0 \neq (\mathbf{v}_k, \mathbf{q}_k) \in \mathbb{V}^2} \frac{\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k))}{\|(\mathbf{v}_k, \mathbf{q}_k)\|_{\mathcal{P}}} \leq \bar{c} \|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}} \quad (4.27)$$

for all $(\mathbf{y}_k, \mathbf{p}_k) \in \mathbb{V}^2$ with constants $\underline{c} = 1/\sqrt{3}$ and $\bar{c} = 1$.

Proof. We start with the proof of the inequality from above. Due to the triangle inequality, it follows that

$$\begin{aligned} |\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k))| &\leq \left| \int_{\Omega} \mathbf{y}_k \cdot \mathbf{v}_k \, d\mathbf{x} \right| + \left| \int_{\Omega} \nu \nabla \mathbf{p}_k \cdot \nabla \mathbf{v}_k \, d\mathbf{x} \right| + \left| \int_{\Omega} k\omega \sigma \mathbf{p}_k \cdot \mathbf{v}_k^{\perp} \, d\mathbf{x} \right| \\ &\quad + \left| \int_{\Omega} \nu \nabla \mathbf{y}_k \cdot \nabla \mathbf{q}_k \, d\mathbf{x} \right| + \left| \int_{\Omega} k\omega \sigma \mathbf{y}_k \cdot \mathbf{q}_k^{\perp} \, d\mathbf{x} \right| + \left| \int_{\Omega} \lambda^{-1} \mathbf{p}_k \cdot \mathbf{q}_k \, d\mathbf{x} \right|. \end{aligned}$$

After appropriate scaling with the parameter λ and applying several times the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} |\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k))| &\leq \left| \int_{\Omega} \mathbf{y}_k \cdot \mathbf{v}_k \, d\mathbf{x} \right| + \left| \int_{\Omega} \nu \lambda^{-1/4} \nabla \mathbf{p}_k \cdot \lambda^{1/4} \nabla \mathbf{v}_k \, d\mathbf{x} \right| \\ &\quad + \left| \int_{\Omega} k\omega \sigma \lambda^{-1/4} \mathbf{p}_k \cdot \lambda^{1/4} \mathbf{v}_k^{\perp} \, d\mathbf{x} \right| + \left| \int_{\Omega} \nu \lambda^{-1/4} \nabla \mathbf{y}_k \cdot \lambda^{1/4} \nabla \mathbf{q}_k \, d\mathbf{x} \right| \\ &\quad + \left| \int_{\Omega} k\omega \sigma \lambda^{-1/4} \mathbf{y}_k \cdot \lambda^{1/4} \mathbf{q}_k^{\perp} \, d\mathbf{x} \right| + \left| \int_{\Omega} \lambda^{-1} \mathbf{p}_k \cdot \mathbf{q}_k \, d\mathbf{x} \right| \\ &= |(\mathbf{y}_k, \mathbf{v}_k)_{L^2(\Omega)}| + |(\nu \lambda^{-1/4} \nabla \mathbf{p}_k, \lambda^{1/4} \nabla \mathbf{v}_k)_{L^2(\Omega)}| \\ &\quad + |k\omega (\sigma \lambda^{-1/4} \mathbf{p}_k, \lambda^{1/4} \mathbf{v}_k^{\perp})_{L^2(\Omega)}| + |(\nu \lambda^{1/4} \nabla \mathbf{y}_k, \lambda^{-1/4} \nabla \mathbf{q}_k)_{L^2(\Omega)}| \\ &\quad + |k\omega (\sigma \lambda^{1/4} \mathbf{y}_k, \lambda^{-1/4} \mathbf{q}_k^{\perp})_{L^2(\Omega)}| + |\lambda^{-1}(\mathbf{p}_k, \mathbf{q}_k)_{L^2(\Omega)}| \\ &\leq \|\mathbf{y}_k\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} + \lambda^{-1/4} (\nu \nabla \mathbf{p}_k, \nabla \mathbf{p}_k)_{L^2(\Omega)}^{1/2} \lambda^{1/4} (\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)}^{1/2} \\ &\quad + \sqrt{k\omega} \lambda^{-1/4} (\sigma \mathbf{p}_k, \mathbf{p}_k)_{L^2(\Omega)}^{1/2} \sqrt{k\omega} \lambda^{1/4} (\sigma \mathbf{v}_k^{\perp}, \mathbf{v}_k^{\perp})_{L^2(\Omega)}^{1/2} \\ &\quad + \lambda^{1/4} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)}^{1/2} \lambda^{-1/4} (\nu \nabla \mathbf{q}_k, \nabla \mathbf{q}_k)_{L^2(\Omega)}^{1/2} \\ &\quad + \sqrt{k\omega} \lambda^{1/4} (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)}^{1/2} \sqrt{k\omega} \lambda^{-1/4} (\sigma \mathbf{q}_k^{\perp}, \mathbf{q}_k^{\perp})_{L^2(\Omega)}^{1/2} \\ &\quad + \lambda^{-1/2} \|\mathbf{p}_k\|_{L^2(\Omega)} \lambda^{-1/2} \|\mathbf{q}_k\|_{L^2(\Omega)}. \end{aligned}$$

Applying the Cauchy-Schwarz inequality again several times, we obtain

$$\begin{aligned} |\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k))| &\leq (\|\mathbf{y}_k\|_{L^2(\Omega)}^2 + \lambda^{-1/2} (\nu \nabla \mathbf{p}_k, \nabla \mathbf{p}_k)_{L^2(\Omega)} + k\omega \lambda^{-1/2} (\sigma \mathbf{p}_k, \mathbf{p}_k)_{L^2(\Omega)} \\ &\quad + \lambda^{1/2} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + k\omega \lambda^{1/2} (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} + \lambda^{-1} \|\mathbf{p}_k\|_{L^2(\Omega)}^2)^{1/2} \\ &\quad (\|\mathbf{v}_k\|_{L^2(\Omega)}^2 + \lambda^{1/2} (\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)} + k\omega \lambda^{1/2} (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)} \\ &\quad + \lambda^{-1/2} (\nu \nabla \mathbf{q}_k, \nabla \mathbf{q}_k)_{L^2(\Omega)} + k\omega \lambda^{-1/2} (\sigma \mathbf{q}_k, \mathbf{q}_k)_{L^2(\Omega)} + \lambda^{-1} \|\mathbf{q}_k\|_{L^2(\Omega)}^2)^{1/2} \\ &= (\|\mathbf{y}_k\|_{\mathbb{V}}^2 + \lambda^{-1} \|\mathbf{p}_k\|_{\mathbb{V}}^2)^{1/2} (\|\mathbf{v}_k\|_{\mathbb{V}}^2 + \lambda^{-1} \|\mathbf{q}_k\|_{\mathbb{V}}^2)^{1/2} = \|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}} \|(\mathbf{v}_k, \mathbf{q}_k)\|_{\mathcal{P}}. \end{aligned}$$

Hence, we have proved the upper bound with $\bar{c} = 1$. Now, we want to show the estimate from below. With the choice

$$(\mathbf{v}_k, \mathbf{q}_k) = \left(\mathbf{y}_k - \frac{1}{\sqrt{\lambda}} \mathbf{p}_k - \frac{1}{\sqrt{\lambda}} \mathbf{p}_k^\perp, \mathbf{p}_k + \sqrt{\lambda} \mathbf{y}_k - \sqrt{\lambda} \mathbf{y}_k^\perp \right),$$

we get the relations

$$\begin{aligned} \mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{y}_k, \mathbf{p}_k)) &= \|\mathbf{y}_k\|_{L^2(\Omega)}^2 + \lambda^{-1} \|\mathbf{p}_k\|_{L^2(\Omega)}^2, \\ \mathcal{B}_k\left((\mathbf{y}_k, \mathbf{p}_k), \left(-\frac{1}{\sqrt{\lambda}} \mathbf{p}_k, \sqrt{\lambda} \mathbf{y}_k\right)\right) &= \sqrt{\lambda} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + \frac{1}{\sqrt{\lambda}} (\nu \nabla \mathbf{p}_k, \nabla \mathbf{p}_k)_{L^2(\Omega)}, \\ \mathcal{B}_k\left((\mathbf{y}_k, \mathbf{p}_k), \left(-\frac{1}{\sqrt{\lambda}} \mathbf{p}_k^\perp, -\sqrt{\lambda} \mathbf{y}_k^\perp\right)\right) &= k\omega \sqrt{\lambda} (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} + k\omega \frac{1}{\sqrt{\lambda}} (\sigma \mathbf{p}_k, \mathbf{p}_k)_{L^2(\Omega)}. \end{aligned}$$

Altogether, with this choice, we obtain

$$\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k)) = \|\mathbf{y}_k\|_{\mathbb{V}}^2 + \lambda^{-1} \|\mathbf{p}_k\|_{\mathbb{V}}^2 = \|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2.$$

By using the fact that

$$\|(\mathbf{v}_k, \mathbf{q}_k)\|_{\mathcal{P}}^2 = \left\| \left(\mathbf{y}_k - \frac{1}{\sqrt{\lambda}} \mathbf{p}_k - \frac{1}{\sqrt{\lambda}} \mathbf{p}_k^\perp, \mathbf{p}_k + \sqrt{\lambda} \mathbf{y}_k - \sqrt{\lambda} \mathbf{y}_k^\perp \right) \right\|_{\mathcal{P}}^2 = 3 \|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2,$$

we arrive at the following estimate of the supremum from below:

$$\sup_{0 \neq (\mathbf{v}_k, \mathbf{q}_k) \in \mathbb{V}^2} \frac{\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k))}{\|(\mathbf{v}_k, \mathbf{q}_k)\|_{\mathcal{P}}} \geq \frac{1}{\sqrt{3}} \|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}.$$

Hence, we get $\underline{c} = 1/\sqrt{3}$. This completes the proof of the theorem. \square

Remark 4.5. *The inequalities (4.27) in Theorem 4.4 immediately yield existence and uniqueness of the solution of variational problem (4.9).*

Due to the supremum, the discrete version of the left inequality in Theorem 4.4, i.e., the inf-sup condition, does in general not follow from the continuous version. However, in our case, we can repeat the proof step-by-step, and, finally, we arrive at the same inequalities in the discrete case where \mathbb{V}^2 is replaced by \mathbb{V}_h^2 with the same constants. Therefore, in matrix-vector notation, we have proved similar inequalities as in (3.35), but now for optimal control problems, i.e., the inequalities

$$\underline{c} \|\underline{u}_k\|_{\mathcal{P}} \leq \sup_{\underline{v}_k \in \mathbb{R}^{4n}} \frac{(\mathcal{A} \underline{u}_k, \underline{v}_k)}{\|\underline{v}_k\|_{\mathcal{P}}} \leq \bar{c} \|\underline{u}_k\|_{\mathcal{P}} \quad \forall \underline{u}_k \in \mathbb{R}^{4n} \quad (4.28)$$

implying the condition number estimate

$$\kappa_{\mathcal{P}}(\mathcal{P}^{-1} \mathcal{A}) := \|\mathcal{P}^{-1} \mathcal{A}\|_{\mathcal{P}} \|\mathcal{A}^{-1} \mathcal{P}\|_{\mathcal{P}} \leq \bar{c}/\underline{c} = \sqrt{3}. \quad (4.29)$$

This condition number estimate yields together with Theorem 2.25 a robust convergence rate of the preconditioned MINRES method with

$$q = \frac{\kappa_{\mathcal{P}}(\mathcal{P}^{-1} \mathcal{A}) - 1}{\kappa_{\mathcal{P}}(\mathcal{P}^{-1} \mathcal{A}) + 1} \leq \frac{\sqrt{3} - 1}{\sqrt{3} + 1} \approx 0.267949$$

of the factor q defining the residual reduction $2q^m/(1+q^{2m})$ after $2m$ MINRES iterations.

Finally, we want to determine a preconditioner for the discretized system (4.12) in the case of $k = 0$. This is done in the same way as before. The matrix \mathcal{A} is now given by

$$\mathcal{A} := \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}, \quad (4.30)$$

where $A := M_h$, $B := -K_{h,\nu}$ and $C := \lambda^{-1}M_h$. According to Theorem 2.26, we construct the two block-diagonal preconditioners \mathcal{P}_0 and \mathcal{P}_1 with

$$\mathcal{P}_0 = \begin{pmatrix} A & 0 \\ 0 & C + BA^{-1}B^T \end{pmatrix} \quad \text{and} \quad \mathcal{P}_1 = \begin{pmatrix} A + B^T C^{-1} B & 0 \\ 0 & C \end{pmatrix}.$$

By applying Theorem 2.29, we obtain the new preconditioner \mathcal{P} with

$$\mathcal{P} = \begin{pmatrix} [A, R]_{1/2} & 0 \\ 0 & [S, C]_{1/2} \end{pmatrix} = \begin{pmatrix} [A, R]_{1/2} & 0 \\ 0 & \lambda^{-1}[A, R]_{1/2} \end{pmatrix},$$

where

$$\begin{aligned} [A, R]_{1/2} &= M_h^{1/2}(M_h^{-1/2}(M_h + \lambda K_{h,\nu} M_h^{-1} K_{h,\nu}) M_h^{-1/2})^{1/2} M_h^{1/2} \\ &\sim M_h^{1/2}(M_h^{-1/2} M_h M_h^{-1/2})^{1/2} M_h^{1/2} + \sqrt{\lambda} M_h^{1/2}(M_h^{-1/2} K_{h,\nu} M_h^{-1} K_{h,\nu} M_h^{-1/2})^{1/2} M_h^{1/2} \\ &= M_h^{1/2} M_h^{-1/2} M_h^{1/2} M_h^{1/2} + \sqrt{\lambda} M_h^{1/2} M_h^{-1/2} K_{h,\nu} M_h^{-1/2} M_h^{1/2} \\ &= M_h + \sqrt{\lambda} K_{h,\nu} =: D. \end{aligned}$$

Hence, the preconditioner is given by

$$\mathcal{P} = \begin{pmatrix} D & 0 \\ 0 & \lambda^{-1} D \end{pmatrix}. \quad (4.31)$$

We can again establish similar inequalities as in Theorem 4.4. Indeed, let us define the bilinear form

$$\mathcal{B}_0((y_0^c, p_0^c), (v_0^c, q_0^c)) := \int_{\Omega} (y_0^c v_0^c - \nu \nabla p_0^c \cdot \nabla v_0^c + \nu \nabla y_0^c \cdot \nabla q_0^c + \lambda^{-1} p_0^c q_0^c) \, d\mathbf{x}. \quad (4.32)$$

Then the variational problem (4.10) reads now as follows: Find $(y_0^c, p_0^c) \in \mathbb{V} = V \times V$ with $V = H_0^1(\Omega)$ such that

$$\mathcal{B}_0((y_0^c, p_0^c), (v_0^c, q_0^c)) = \int_{\Omega} y_{d0}^c \cdot v_0^c \, d\mathbf{x} \quad (4.33)$$

for all test functions $(v_0^c, q_0^c) \in \mathbb{V}$. Moreover, defining the inner product

$$((y, p), (v, q))_{\mathcal{P}} = (y, v)_{L^2(\Omega)} + \sqrt{\lambda}(\nu \nabla y, \nabla v)_{L^2(\Omega)} + \lambda^{-1}((p, q)_{L^2(\Omega)} + \sqrt{\lambda}(\nu \nabla p, \nabla q)_{L^2(\Omega)})$$

with associated norm

$$\|(y, p)\|_{\mathcal{P}}^2 = \|y\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu \nabla y, \nabla y)_{L^2(\Omega)} + \lambda^{-1}(\|p\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu \nabla p, \nabla p)_{L^2(\Omega)}),$$

we can again show the following inequalities:

$$\underline{c} \|(y_0^c, p_0^c)\|_{\mathcal{P}} \leq \sup_{0 \neq (v_0^c, q_0^c) \in \mathbb{V}} \frac{\mathcal{B}_0((y_0^c, p_0^c), (v_0^c, q_0^c))}{\|(v_0^c, q_0^c)\|_{\mathcal{P}}} \leq \bar{c} \|(y_0^c, p_0^c)\|_{\mathcal{P}} \quad (4.34)$$

for all $(y_0^c, p_0^c) \in \mathbb{V}$ with constants \underline{c} and \bar{c} independent of all involved parameters. The upper bound of the supremum with the constant $\bar{c} = 1$ is again obtained by applying triangle and Cauchy-Schwarz inequalities. The estimate from below follows by the choice

$$(v_0^c, q_0^c) = \left(y_0^c - \frac{1}{\sqrt{\lambda}} p_0^c, p_0^c + \sqrt{\lambda} y_0^c \right).$$

For this choice, we obtain

$$\|(v_0^c, q_0^c)\|_{\mathcal{P}}^2 = 2\|(y_0^c, p_0^c)\|_{\mathcal{P}}^2 \quad \text{and} \quad \mathcal{B}_0((y_0^c, p_0^c), (v_0^c, q_0^c)) = \|(y_0^c, p_0^c)\|_{\mathcal{P}}^2.$$

Hence, the constant for the lower bound is $\underline{c} = 1/\sqrt{2}$. The same arguments as above lead us to the estimate

$$\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) \leq \sqrt{2}, \quad (4.35)$$

which provides a robust convergence rate of the preconditioned MINRES method by Theorem 2.25 with

$$q \leq \frac{\sqrt{2} - 1}{\sqrt{2} + 1} \approx 0.171573.$$

In a nutshell, we have designed preconditioners for the linear systems (4.11) and (4.12) corresponding to the modes $1 \leq k \leq N$ and $k = 0$, respectively, providing robust convergence rates for solving the preconditioned system by the preconditioned MINRES method.

In practical applications, for a large number of degrees of freedom, it is not very efficient to use sparse direct methods to invert the diagonal blocks appearing in the preconditioners \mathcal{P} in (4.22) of the discretized problem (4.11) for the case $k = 1, \dots, N$ and in (4.31) of the problem (4.12) for the case $k = 0$. Hence, it is important to replace the diagonal blocks $D = \sqrt{\lambda}K_h + k\omega\sqrt{\lambda}M_{h,\sigma} + M_h$ and $D = M_h + \sqrt{\lambda}K_h$ of the preconditioners \mathcal{P} by diagonal blocks \tilde{D} , which are spectrally equivalent to D , robust, symmetric positive definite and more cost efficient. As already mentioned at the end of Section 3.3, such robust and efficient practical preconditioners \tilde{D} for the diagonal blocks D can be constructed by various techniques as by (algebraic) multigrid, multilevel or domain decomposition methods, see, e.g., [101, 142, 168, 173]. In Chapter 5, we consider the algebraic multilevel iteration method for constructing such robust and optimal preconditioners. Moreover, we present numerical results using this algebraic multilevel preconditioner as well as other preconditioners in Chapter 7.

4.4 Discretization error analysis

We start the discretization error analysis by defining some norms in certain function spaces inspired by the \mathcal{P} -norm in the same spirit as it is done in Section 3.4. Let us consider again the function spaces (3.36), i.e.,

$$V_0 := H_0^{1, \frac{1}{2}}(Q_T) \quad \text{and} \quad V_1 := (H^{0, \frac{1}{2}})_0^1(Q_T) \cap H_{per}^{0,1}(Q_T),$$

where

$$(H^{0, \frac{1}{2}})_0^1(Q_T) := \{y \in H^{0, \frac{1}{2}}(Q_T) : \nabla y \in H^{0, \frac{1}{2}}(Q_T), y = 0 \text{ on } \Sigma_T\}.$$

Let us equip these spaces with the norms

$$\begin{aligned} \|y\|_{V_0}^2 &= \|y\|_{L^2(Q_T)}^2 + \sqrt{\lambda}(\nu \nabla y, \nabla y)_{L^2(Q_T)} + \sqrt{\lambda}(\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)}, \\ \|y\|_{V_1}^2 &= \|y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 + \sqrt{\lambda}(\nu \nabla y, \nabla y)_{L^2(Q_T)} \\ &\quad + \sqrt{\lambda}(\nu \partial_t^{1/2} \nabla y, \partial_t^{1/2} \nabla y)_{L^2(Q_T)} + \sqrt{\lambda}(\sigma \partial_t y, \partial_t y)_{L^2(Q_T)}. \end{aligned}$$

These norms are again defined in the Fourier space according to Definition 3.2, i.e.,

$$\begin{aligned} \|y\|_{V_0}^2 &= T(\|y_0^c\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu \nabla y_0^c, \nabla y_0^c)_{L^2(\Omega)}) \\ &\quad + \frac{T}{2} \sum_{k=1}^{\infty} [\|\mathbf{y}_k\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + \sqrt{\lambda}k\omega(\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)}], \\ \|y\|_{V_1}^2 &= T(\|y_0^c\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu \nabla y_0^c, \nabla y_0^c)_{L^2(\Omega)}) \\ &\quad + \frac{T}{2} \sum_{k=1}^{\infty} [(1 + k\omega)\|\mathbf{y}_k\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(1 + k\omega)(\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + \sqrt{\lambda}(k\omega)^2(\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)}]. \end{aligned}$$

Moreover, we introduce the norms

$$\|(y, p)\|_{P_0}^2 = \|y\|_{V_0}^2 + \lambda^{-1} \|p\|_{V_0}^2 \quad \text{and} \quad \|(y, p)\|_{P_1}^2 = \|y\|_{V_1}^2 + \lambda^{-1} \|p\|_{V_1}^2.$$

Remark 4.6. Note that the \mathbb{V} -norm defined in (4.25) as well as the \mathcal{P} -norm defined in (4.26) correspond to the V_0 -norm and the P_0 -norm, respectively, for a single mode k , i.e.,

$$\|y\|_{V_0}^2 = T \|y_0^c\|_{\mathbb{V}}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{y}_k\|_{\mathbb{V}}^2 \quad \text{and} \quad \|(y, p)\|_{P_0}^2 = T \|(y_0^c, p_0^c)\|_{\mathcal{P}}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2.$$

The complete discretization error between the exact solution of the variational problem (4.7) and its multiharmonic finite element approximation is given by

$$\|(y, p) - (y_{Nh}, p_{Nh})\|_{P_0}, \quad (4.36)$$

where the exact solution (y, p) can be represented as Fourier series, i.e.,

$$\begin{aligned} y(\mathbf{x}, t) &= y_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [y_k^c(\mathbf{x}) \cos(k\omega t) + y_k^s(\mathbf{x}) \sin(k\omega t)], \\ p(\mathbf{x}, t) &= p_0^c(\mathbf{x}) + \sum_{k=1}^{\infty} [p_k^c(\mathbf{x}) \cos(k\omega t) + p_k^s(\mathbf{x}) \sin(k\omega t)], \end{aligned}$$

and its multiharmonic finite element approximation (y_{Nh}, p_{Nh}) is given by (4.13) and (4.14). Using the triangle inequality, we can split the discretization error (4.36) again into two parts, i.e.,

$$\|(y, p) - (y_{Nh}, p_{Nh})\|_{P_0} \leq \underbrace{\|(y, p) - (y_N, p_N)\|_{P_0}}_{\text{discretization error in } N} + \underbrace{\|(y_N, p_N) - (y_{Nh}, p_{Nh})\|_{P_0}}_{\text{discretization error in } h}.$$

4.4.1 Discretization error with respect to the truncation index

As in Subsection 3.4.1, we present a theorem which provides an estimate for the discretization error due to truncation of the Fourier series at the mode N under weak regularity assumptions, see [112].

Theorem 4.7. Let us assume that $y, p \in V_1$. Then the discretization error due to truncation of the Fourier series can be estimated by

$$\|(y, p) - (y_N, p_N)\|_{P_0} \leq c_0 N^{-1/2} \|(y, p)\|_{P_1}, \quad (4.37)$$

where c_0 is a constant depending only on the frequency ω .

Proof. Under the assumption that $y \in V_1$, we obtain the following estimate:

$$\begin{aligned} \|y - y_N\|_{V_0}^2 &= \frac{T}{2} \sum_{k=N+1}^{\infty} \left[\|\mathbf{y}_k\|_{L^2(\Omega)}^2 + \sqrt{\lambda} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + \sqrt{\lambda} k\omega (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} \right] \\ &\leq \frac{T}{2} \sum_{k=N+1}^{\infty} \left[\frac{1+k\omega}{k\omega} \|\mathbf{y}_k\|_{L^2(\Omega)}^2 + \sqrt{\lambda} \frac{1+k\omega}{k\omega} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} \right. \\ &\quad \left. + \sqrt{\lambda} \frac{(k\omega)^2}{k\omega} (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} \right] \\ &\leq \frac{1}{(N+1)\omega} \frac{T}{2} \sum_{k=N+1}^{\infty} \left[(1+k\omega) \|\mathbf{y}_k\|_{L^2(\Omega)}^2 + \sqrt{\lambda} (1+k\omega) (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} \right. \\ &\quad \left. + \sqrt{\lambda} (k\omega)^2 (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} \right] \\ &\leq c_0^2 \frac{1}{N} \|y\|_{V_1}^2, \end{aligned}$$

where $c_0 = c_0(\omega) = 1/\sqrt{\omega}$. Since the same estimate is obviously true for the adjoint state p , we finally get the estimate

$$\|(y, p) - (y_N, p_N)\|_{P_0}^2 = \|y - y_N\|_{V_0}^2 + \lambda^{-1} \|p - p_N\|_{V_0}^2 \leq c_0^2 \frac{1}{N} \|(y, p)\|_{P_1}^2,$$

that completes the proof of Theorem 4.7. \square

4.4.2 Discretization error with respect to the finite element discretization parameter

The discretization error between the multiharmonic approximation of the exact solution and its multiharmonic finite element approximation can be deduced from the discretization error between the unknown Fourier coefficients and their finite element approximations due to linearity. More precisely, we have the identity

$$\|(y_N, p_N) - (y_{Nh}, p_{Nh})\|_{P_0}^2 = T \|(y_0^c, p_0^c) - (y_{0h}^c, p_{0h}^c)\|_{\mathcal{P}}^2 + \frac{T}{2} \sum_{k=1}^N \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{y}_{kh}, \mathbf{p}_{kh})\|_{\mathcal{P}}^2.$$

The discretization error analysis with respect to the finite element discretization parameter h starts with proving that the discretization error of the Fourier coefficients can be estimated by the best approximation error. Afterwards we estimate the best approximation error by the interpolation error provided the Fourier coefficients are sufficiently smooth. We mainly consider the case $k = 1, \dots, N$, since the error analysis for the case $k = 0$ can be done analogously, cf. Subsection 3.4.2 as well.

Our decoupled variational problems for $k = 1, \dots, N$ are given by: Find $(\mathbf{y}_k, \mathbf{p}_k) \in \mathbb{V}^2$ such that

$$\begin{aligned} \mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k), (\mathbf{v}_k, \mathbf{q}_k)) &= \int_{\Omega} \mathbf{y}_{d_k} \cdot \mathbf{v}_k \, d\mathbf{x} \\ &= \int_{\Omega} (\mathbf{y}_{d_k}, 0) \cdot (\mathbf{v}_k, \mathbf{q}_k) \, d\mathbf{x} \\ &=: \langle F_k, (\mathbf{v}_k, \mathbf{q}_k) \rangle \end{aligned}$$

for all test functions $(\mathbf{v}_k, \mathbf{q}_k) \in \mathbb{V}^2$. Due to the assumption that the Fourier coefficients of the given desired state y_d can be computed exactly, the corresponding discrete problems are given by: Find $(\mathbf{y}_{kh}, \mathbf{p}_{kh}) \in \mathbb{V}_h^2$ such that

$$\mathcal{B}_k((\mathbf{y}_{kh}, \mathbf{p}_{kh}), (\mathbf{v}_{kh}, \mathbf{q}_{kh})) = \langle F_k, (\mathbf{v}_{kh}, \mathbf{q}_{kh}) \rangle$$

for all test functions $(\mathbf{v}_{kh}, \mathbf{q}_{kh}) \in \mathbb{V}_h^2$, which is equivalent to solving the linear system (4.11). Moreover, from $\mathbb{V}_h \subset \mathbb{V}$ follows the Galerkin orthogonality

$$\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{y}_{kh}, \mathbf{p}_{kh}), (\mathbf{v}_{kh}, \mathbf{q}_{kh})) = 0 \quad \forall (\mathbf{v}_{kh}, \mathbf{q}_{kh}) \in \mathbb{V}_h^2. \quad (4.38)$$

The following theorem provides an estimate for the discretization error between the unknown Fourier coefficients and their finite element approximations, see also [89].

Theorem 4.8. *Under the assumption that $(\mathbf{y}_k, \mathbf{p}_k) \in (H^2(\Omega))^4$ the discretization error for the Fourier coefficients can be estimated by*

$$\|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{y}_{kh}, \mathbf{p}_{kh})\|_{\mathcal{P}} \leq c_1 c_{par}(\lambda, k, \omega, \bar{\nu}, \bar{\sigma}, h) h |(\mathbf{y}_k, \mathbf{p}_k)|_{H^2(\Omega)}, \quad (4.39)$$

where $c_{par}^2(\lambda, k, \omega, \bar{\nu}, \bar{\sigma}, h) = \sqrt{\lambda \bar{\nu}} c_{1,2}^2 + (1 + k\omega\sqrt{\lambda \bar{\sigma}}) c_{0,2}^2 h^2$ with constants $c_{0,2}$ and $c_{1,2}$ from the approximation theorem and c_1 is a positive constant. The weighted H^2 -seminorm $|\cdot|_{H^2(\Omega)}$ is defined by the relation

$$|(\mathbf{y}_k, \mathbf{p}_k)|_{H^2(\Omega)}^2 = |\mathbf{y}_k|_{H^2(\Omega)}^2 + \lambda^{-1} |\mathbf{p}_k|_{H^2(\Omega)}^2. \quad (4.40)$$

Proof. Inserting an arbitrary $(\mathbf{v}_{kh}, \mathbf{q}_{kh}) \in \mathbb{V}_h^2$ and using triangle inequality, the discrete inf-sup condition as well as the sup-sup condition of the Babuška-Aziz theorem together with (4.38), we obtain the following estimate:

$$\begin{aligned}
\|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{y}_{kh}, \mathbf{p}_{kh})\|_{\mathcal{P}} &\leq \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} + \|(\mathbf{y}_{kh}, \mathbf{p}_{kh}) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} \\
&\leq \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} + \sqrt{3} \sup_{0 \neq (\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh}) \in \mathbb{V}_h^2} \frac{\mathcal{B}_k((\mathbf{y}_{kh}, \mathbf{p}_{kh}) - (\mathbf{v}_{kh}, \mathbf{q}_{kh}), (\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh}))}{\|(\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh})\|_{\mathcal{P}}} \\
&\leq \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} + \sqrt{3} \underbrace{\sup_{0 \neq (\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh}) \in \mathbb{V}_h^2} \frac{\mathcal{B}_k((\mathbf{y}_{kh}, \mathbf{p}_{kh}) - (\mathbf{y}_k, \mathbf{p}_k), (\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh}))}{\|(\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh})\|_{\mathcal{P}}}}_{=0} \\
&\quad + \sqrt{3} \sup_{0 \neq (\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh}) \in \mathbb{V}_h^2} \frac{\mathcal{B}_k((\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh}), (\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh}))}{\|(\tilde{\mathbf{v}}_{kh}, \tilde{\mathbf{q}}_{kh})\|_{\mathcal{P}}} \\
&\leq \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} + \sqrt{3} \cdot 1 \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} \\
&\leq \underbrace{(1 + \sqrt{3})}_{=: c_1} \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}}.
\end{aligned}$$

We can estimate the discretization error by the best approximation error, i.e.,

$$\|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{y}_{kh}, \mathbf{p}_{kh})\|_{\mathcal{P}} \leq c \inf_{(\mathbf{v}_{kh}, \mathbf{q}_{kh}) \in \mathbb{V}_h^2} \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}}, \quad (4.41)$$

and then, estimate the best approximation error by the interpolation error, i.e.

$$\inf_{(\mathbf{v}_{kh}, \mathbf{q}_{kh}) \in \mathbb{V}_h^2} \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} \leq \|(\mathbf{y}_k, \mathbf{p}_k) - I_h^2(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}, \quad (4.42)$$

where $I_h^2 : \mathbb{V}^2 \rightarrow \mathbb{V}_h^2$ (respectively $I_h : \mathbb{V} \rightarrow \mathbb{V}_h$) is some interpolation operator. The \mathcal{P} -norm

$$\|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2 = \|\mathbf{y}_k\|_{\mathbb{V}}^2 + \lambda^{-1} \|\mathbf{p}_k\|_{\mathbb{V}}^2$$

with

$$\begin{aligned}
\|\mathbf{y}_k\|_{\mathbb{V}}^2 &= \sqrt{\lambda} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + k\omega \sqrt{\lambda} (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} + \|\mathbf{y}_k\|_{L^2(\Omega)}^2 \\
&\leq \sqrt{\lambda \bar{\nu}} |\mathbf{y}_k|_{H^1(\Omega)}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) \|\mathbf{y}_k\|_{L^2(\Omega)}^2
\end{aligned}$$

is bounded by

$$\begin{aligned}
\|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2 &\leq \sqrt{\lambda \bar{\nu}} |\mathbf{y}_k|_{H^1(\Omega)}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) \|\mathbf{y}_k\|_{L^2(\Omega)}^2 \\
&\quad + \lambda^{-1} (\sqrt{\lambda \bar{\nu}} |\mathbf{p}_k|_{H^1(\Omega)}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) \|\mathbf{p}_k\|_{L^2(\Omega)}^2).
\end{aligned}$$

Under the assumption that the Fourier coefficients are from $H^2(\Omega)$, the interpolation error can be estimated by

$$\begin{aligned}
\|(\mathbf{y}_k, \mathbf{p}_k) - I_h^2(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2 &= \|(I - I_h)^2(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2 = \|(I - I_h)\mathbf{y}_k\|_{\mathbb{V}}^2 + \lambda^{-1} \|(I - I_h)\mathbf{p}_k\|_{\mathbb{V}}^2 \\
&\leq \sqrt{\lambda \bar{\nu}} |(I - I_h)\mathbf{y}_k|_{H^1(\Omega)}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) \|(I - I_h)\mathbf{y}_k\|_{L^2(\Omega)}^2 \\
&\quad + \lambda^{-1} (\sqrt{\lambda \bar{\nu}} |(I - I_h)\mathbf{p}_k|_{H^1(\Omega)}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) \|(I - I_h)\mathbf{p}_k\|_{L^2(\Omega)}^2) \\
&\leq \sqrt{\lambda \bar{\nu}} c_{1,2}^2 h^2 |\mathbf{y}_k|_{H^2(\Omega)}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) c_{0,2}^2 h^4 |\mathbf{y}_k|_{H^2(\Omega)}^2 \\
&\quad + \lambda^{-1} (\sqrt{\lambda \bar{\nu}} c_{1,2}^2 h^2 |\mathbf{p}_k|_{H^2(\Omega)}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) c_{0,2}^2 h^4 |\mathbf{p}_k|_{H^2(\Omega)}^2) \\
&= ((\sqrt{\lambda \bar{\nu}} c_{1,2}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) c_{0,2}^2 h^2) h^2) |\mathbf{y}_k|_{H^2(\Omega)}^2 \\
&\quad + \lambda^{-1} \underbrace{(\sqrt{\lambda \bar{\nu}} c_{1,2}^2 + (1 + k\omega \sqrt{\lambda \bar{\sigma}}) c_{0,2}^2 h^2)}_{=: c_{par}^2(\lambda, k, \omega, \bar{\nu}, \bar{\sigma}, h)} h^2 |\mathbf{p}_k|_{H^2(\Omega)}^2 \\
&= c_{par}^2(\lambda, k, \omega, \bar{\nu}, \bar{\sigma}, h) h^2 (|\mathbf{y}_k|_{H^2(\Omega)}^2 + \lambda^{-1} |\mathbf{p}_k|_{H^2(\Omega)}^2),
\end{aligned}$$

where $c_{0,2}$ and $c_{1,2}$ are constants coming from applying the approximation theorem from finite element discretization theory, see, e.g., [42, 46]. Hence, it follows that

$$\|(\mathbf{y}_k, \mathbf{p}_k) - I_h^2(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}} \leq c_{par}(\lambda, k, \omega, \bar{\nu}, \bar{\sigma}, h) h \underbrace{(|\mathbf{y}_k|_{H^2(\Omega)}^2 + \lambda^{-1} |\mathbf{p}_k|_{H^2(\Omega)}^2)^{1/2}}_{=:(\mathbf{y}_k, \mathbf{p}_k)|_{H^2(\Omega)}},$$

where $|(\mathbf{y}_k, \mathbf{p}_k)|_{H^2(\Omega)}$ is a weighted H^2 -seminorm defined by (4.40). Altogether the discretization error for the Fourier coefficients can be estimated by

$$\begin{aligned} \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{y}_{kh}, \mathbf{p}_{kh})\|_{\mathcal{P}} &\leq c_1 \inf_{(\mathbf{v}_{kh}, \mathbf{q}_{kh}) \in \mathbb{V}_h^2} \|(\mathbf{y}_k, \mathbf{p}_k) - (\mathbf{v}_{kh}, \mathbf{q}_{kh})\|_{\mathcal{P}} \\ &\leq c_1 \|(\mathbf{y}_k, \mathbf{p}_k) - I_h^2(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}} \\ &\leq c_1 c_{par}(\lambda, k, \omega, \bar{\nu}, \bar{\sigma}, h) h |(\mathbf{y}_k, \mathbf{p}_k)|_{H^2(\Omega)} \end{aligned} \quad (4.43)$$

with the constant $c_1 = 1 + \sqrt{3}$. \square

Under the assumption that $(y_0^c, p_0^c) \in (H^2(\Omega))^2$ we obtain the following estimate for the discretization error in the case of $k = 0$, which can analogously be proven as Theorem 4.8, by using (4.34):

$$\|(y_0^c, p_0^c) - (y_{0h}^c, p_{0h}^c)\|_{\mathcal{P}} \leq (1 + \sqrt{2}) c_{par}(\lambda, \bar{\nu}, h) h |(y_0^c, p_0^c)|_{H^2(\Omega)}, \quad (4.44)$$

where $c_{par}^2(\lambda, \bar{\nu}, h) = \sqrt{\lambda \bar{\nu}} c_{1,2}^2 + c_{0,2}^2 h^2$ with constants $c_{0,2}$ and $c_{1,2}$ coming from the approximation theorem.

According to (3.47), we define again a $H^{2,0}(Q_T)$ -seminorm in the Fourier space, i.e.,

$$|(y, p)|_{H^2} = \left(T |(y_0^c, p_0^c)|_{H^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} |(\mathbf{y}_k, \mathbf{p}_k)|_{H^2(\Omega)}^2 \right)^{1/2}, \quad (4.45)$$

where the $H^2(\Omega)$ -seminorm for the Fourier coefficients is defined in (4.40). The following theorem provides the estimate for the complete discretization error with respect to the spatial discretization parameter h .

Theorem 4.9. *Under the assumptions that $(y_0^c, p_0^c) \in (H^2(\Omega))^2$ and $(\mathbf{y}_k, \mathbf{p}_k) \in (H^2(\Omega))^4$ for all $k = 1, \dots, N$, the error with respect to the discretization parameter of the finite element discretization can be estimated as follows*

$$\|(y_N, p_N) - (y_{Nh}, p_{Nh})\|_{P_0} \leq c_1 c_{par}(\lambda, N, \omega, \bar{\nu}, \bar{\sigma}, h) h |(y_N, p_N)|_{H^2}, \quad (4.46)$$

where $c_{par}^2(\lambda, N, \omega, \bar{\nu}, \bar{\sigma}, h) = \sqrt{\lambda \bar{\nu}} c_{1,2}^2 + (1 + N\omega\sqrt{\lambda \bar{\sigma}}) c_{0,2}^2 h^2$ with constants $c_{0,2}$ and $c_{1,2}$ from the approximation theorem and $c_1 = 1 + \sqrt{3}$. The H^2 -seminorm is given by

$$|(y_N, p_N)|_{H^2}^2 = T |(y_0^c, p_0^c)|_{H^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N |(\mathbf{y}_k, \mathbf{p}_k)|_{H^2(\Omega)}^2.$$

Proof. The proof immediately follows from Theorem 4.8, more precisely, by using (4.39), and from (4.44). \square

4.4.3 Complete discretization error

Finally, the following theorem presents the result for the complete discretization error (4.36).

Theorem 4.10. *Let us assume that $y, p \in V_1 \cap H^{2,0}(Q_T)$. Then the complete discretization error arising from the multiharmonic finite element discretization can be estimated as follows*

$$\|(y, p) - (y_{Nh}, p_{Nh})\|_{P_0} \leq c_0 N^{-1/2} \|(y, p)\|_{P_1} + c_1 c_{par}(\lambda, N, \omega, \bar{\nu}, \bar{\sigma}, h) h |(y, p)|_{H^2},$$

where c_0 and c_1 come from Theorem 4.7 and Theorem 4.9, respectively, and $c_{par}^2(\lambda, N, \omega, \bar{\nu}, \bar{\sigma}, h) = \sqrt{\lambda \bar{\nu}} c_{1,2}^2 + (1 + N\omega\sqrt{\lambda \bar{\sigma}}) c_{0,2}^2 h^2$ with constants $c_{0,2}$ and $c_{1,2}$ from the approximation theorem.

Proof. Applying the triangle inequality and using Theorems 4.7 and 4.9 yield the estimates

$$\begin{aligned} \|(y, p) - (y_{Nh}, p_{Nh})\|_{P_0} &\leq \|(y, p) - (y_N, p_N)\|_{P_0} + \|(y_N, p_N) - (y_{Nh}, p_{Nh})\|_{P_0} \\ &\leq c_0 N^{-1/2} \|(y, p)\|_{P_1} + c_1 c_{par}(\lambda, N, \omega, \bar{\nu}, \bar{\sigma}, h) h |(y_N, p_N)|_{H^2}, \end{aligned}$$

where the seminorm $|(y_N, p_N)|_{H^2}$ can trivially be estimated by (4.45). \square

Remark 4.11. *The same statements for optimal control problems can be made as in Remark 3.20, i.e., the convergence rate with respect to h reduces from h to h^s with some $s \in (0, 1)$, if*

$$y, p \in V_1 \cap H^{1+s, 0}(Q_T),$$

whereas higher order elements are needed in order to obtain h^s with $s > 1$. Moreover, the convergence with respect to N will improve, if y and p are smoother with respect to the time variable, i.e., the factor $N^{-1/2}$ improves to $N^{-\ell/2}$ provided that

$$y, p \in V_\ell \cap H^{1+s, 0}(Q_T),$$

with $V_\ell := (H^{0, \ell/2})_0^1(Q_T) \cap H_{per}^{0, (\ell+1)/2}(Q_T)$ and with some $\ell > 1$. This is confirmed by our numerical experiments presented in Chapter 7. In particular, we observe very fast convergence with respect to N for time-analytic solutions.

Chapter 5

Robustness and optimality of algebraic multilevel preconditioners for reaction-diffusion type problems

The examination of the block-diagonal preconditioners in Chapters 3 and 4, whose diagonal blocks are weighted sums of stiffness and mass matrices, strongly motivates to construct robust and optimal preconditioners for these sums of stiffness and mass matrices. Moreover, we do not only want to construct efficient preconditioners but to provide a rigorous proof of their robustness and optimal complexity. Hence, this chapter is devoted to the analysis of preconditioners for reaction-diffusion type problems that are both, uniform with respect to the reaction and diffusion coefficients, and optimal in terms of computational complexity. The considered preconditioners belong to the class of algebraic multilevel iteration (AMLI) methods, which are based on multilevel block factorization and polynomial stabilization. In Section 2.8, we have already presented some of the fundamental results regarding the AMLI method.

The main focus of this chapter is on the construction and on the analysis of a hierarchical splitting of the conforming finite element space of piecewise linear functions that allows to meet the optimality conditions for the related AMLI preconditioner in case of second-order elliptic problems with non-vanishing zero-order term. The finite element method then leads to a system of linear equations with a system matrix that is a weighted sum of stiffness and mass matrices. We compute bounds for the constant γ in the strengthened Cauchy-Bunyakowski-Schwarz inequality (2.59) for both, mass and stiffness, matrices in case of a general m -refinement, including a new estimate for the mass matrix. Moreover, we present an additive preconditioner for the pivot blocks with (2.70) that arise in the course of the multilevel block factorization and prove its optimality for the case $m = 3$. Together with the estimates for γ this shows that the construction of a uniformly convergent AMLI method with optimal complexity is possible (for all $m \geq 3$). In other words, the derived uniform condition number estimates together with the verification of the optimality conditions guarantee the existence of optimal linear AMLI methods for linear systems with weighted sums of stiffness and mass matrices. We discuss the practical application of this preconditioning technique in the context of parabolic time-periodic problems at the end of this chapter. First numerical results using the linear AMLI preconditioned CG algorithm in case of a 3-refinement together with the additive preconditioner for the pivot block of the two-by-two splitting can be found in Chapter 7. Altogether, this linear AMLI preconditioned CG algorithm leads to a robust solver of optimal complexity.

5.1 A reaction-diffusion type problem

Let $\Omega \subset \mathbb{R}^2$ be a two-dimensional bounded Lipschitz domain with boundary $\Gamma := \partial\Omega$. For simplicity, we assume that Ω is a polygonal domain. We consider the following heterogeneous reaction-diffusion model problem:

$$\begin{aligned} -\operatorname{div}(\nu(\mathbf{x})\nabla u(\mathbf{x})) + \mu(\mathbf{x})u(\mathbf{x}) &= f(\mathbf{x}) & \mathbf{x} \in \Omega, \\ u(\mathbf{x}) &= 0 & \mathbf{x} \in \Gamma, \end{aligned} \quad (5.1)$$

where the coefficients ν and μ are assumed to be measurable, uniformly bounded, and positive and non-negative, respectively, i.e.,

$$0 < \underline{\nu} \leq \nu(\mathbf{x}) \leq \bar{\nu} \quad \text{and} \quad 0 \leq \underline{\mu} \leq \mu(\mathbf{x}) \leq \bar{\mu}, \quad \mathbf{x} \in \Omega.$$

Usually, these coefficients are piecewise constant, e.g., due to different material parameters in different subdomains.

5.1.1 The variational problem

In order to formulate the variational problem corresponding to (5.1), one multiplies the first equation of (5.1) by a test function $v \in V$, where V is the Hilbert space

$$V := H_0^1(\Omega) = \{u \in L^2(\Omega) : \nabla u \in L^2(\Omega), u = 0 \text{ on } \Gamma\}$$

equipped with the norm

$$\|u\|_{H^1(\Omega)} = \left(\|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 \right)^{1/2},$$

and integrates over Ω . Integration by parts finally yields the following variational problem: Given $f \in L^2(\Omega)$, find $u \in V$ such that

$$\int_{\Omega} (\nu(\mathbf{x})\nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) + \mu(\mathbf{x})u(\mathbf{x})v(\mathbf{x})) \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x} \quad (5.2)$$

for all test functions $v \in V$.

Theorem 5.1. *The variational problem (5.2) has a unique solution.*

Proof. Existence and uniqueness of the solution of problem (5.2) follows immediately from the Lax-Milgram theorem. \square

5.1.2 The finite element discretization

In order to solve the reaction-diffusion problem (5.1), we discretize problem (5.2) by a conforming finite element method (FEM), see, e.g., [41, 46, 84, 161] and Section 2.4. Hence, we approximate the solution $u \in V$ by a finite element function $u_h \in V_h \subset V$. Let us consider the space V_h to be the largest of a sequence of nested spaces, i.e.,

$$V^{(0)} \subset V^{(1)} \subset \dots \subset V^{(\ell)} \subset \dots \subset V^{(L)} = V_h,$$

which correspond to a sequence of nested meshes $\mathcal{T}^{(\ell)}$ for $\ell = 0, \dots, L$ and $\mathcal{T}^{(L)} = \mathcal{T}_h$ is the finest mesh. The spaces

$$V^{(\ell)} = \operatorname{span}\{\varphi_1^{(\ell)}, \dots, \varphi_{n^{(\ell)}}^{(\ell)}\}$$

are finite element spaces spanned by the standard nodal basis functions

$$\{\varphi_i^{(\ell)} : i = 1, \dots, n^{(\ell)}\},$$

where $n^{(\ell)} = \dim V^{(\ell)}$. For the finest triangulation \mathcal{T}_h with $n = n^{(L)} = n_h = \dim V_h$, the following linear system arises from the variational formulation (5.2):

$$\underbrace{(K_{\nu,h} + M_{\mu,h})}_{=: A_h} \underline{u}_h = \underline{f}_h, \quad (5.3)$$

where $K_{\nu,h}$ and $M_{\mu,h}$ correspond to the weighted stiffness and weighted mass matrix, respectively, and \underline{f}_h denotes the load vector. Their entries are computed by the formulas

$$K_{\nu,h}^{ij} = \int_{\Omega} \nu \nabla \varphi_i^{(L)} \cdot \nabla \varphi_j^{(L)} \, d\mathbf{x}, \quad M_{\mu,h}^{ij} = \int_{\Omega} \mu \varphi_i^{(L)} \varphi_j^{(L)} \, d\mathbf{x}$$

with $i, j = 1, \dots, n$ and

$$\underline{f}_h = \left[\int_{\Omega} f \varphi_j^{(L)} \, d\mathbf{x} \right]_{j=1, \dots, n}.$$

The system (5.3) has to be solved for the vector

$$\underline{u}_h = (u_i)_{i=1, \dots, n} \in \mathbb{R}^n$$

of nodal unknowns of the finite element approximation

$$u_h(x) = \sum_{i=1}^n u_i \varphi_i^{(L)}(x).$$

In order to solve problem (5.3) efficiently one needs a robust optimal preconditioner. Such a preconditioner can be implemented by various methods such as algebraic multigrid (AMG), domain decomposition (DD) or the AMLI method. In the following, we will construct AMLI preconditioners as presented in Subsection 2.8.1, which have been introduced in [14, 15], see also [101, 173]. We will present a rigorous proof of their robustness and optimal complexity when used for solving the linear system (5.3).

Let the symmetric and positive definite matrix $A_h = A^{(L)}$ in (5.3) be obtained in the course of a regular refinement procedure, which defines a sequence of symmetric positive definite matrices starting from a coarsest level system matrix $A^{(0)}$, i.e.,

$$\{A^{(\ell)}\}, \quad A^{(\ell)} \in \mathbb{R}^{n^{(\ell)} \times n^{(\ell)}},$$

where $\ell = 0, \dots, L$, and $n^{(\ell)} > n^{(\ell-1)}$, for $\ell = 1, \dots, L$, see [17]. These matrices are constructed for the sequence of nested spaces $V^{(\ell)}$, i.e., (2.56), corresponding to nested meshes $\mathcal{T}^{(\ell)}$ for $\ell = 0, \dots, L$. We partition the matrix $A^{(\ell)}$ on each level ℓ in a two-by-two block form, i.e., (2.57), where its standard FEM assembling can be written as in (2.63). Moreover, let us assume that the parameters ν and μ of problem (5.3) are constant on the coarsest mesh partitioning $\mathcal{T}^{(0)}$, and let us denote by $e^{(0)} \in \mathcal{T}^{(0)}$ an arbitrary element at the coarsest level. Then, the system matrix corresponding to the coarsest mesh can be written as

$$\begin{aligned} A^{(0)} &= \sum_{e^{(0)} \in \mathcal{T}^{(0)}} R_{e^{(0)}}^T A_{e^{(0)}} R_{e^{(0)}} \\ &= \sum_{e^{(0)} \in \mathcal{T}^{(0)}} R_{e^{(0)}}^T (\nu_{e^{(0)}} K_{e^{(0)}} + \mu_{e^{(0)}} M_{e^{(0)}}) R_{e^{(0)}} \\ &= \sum_{e^{(0)} \in \mathcal{T}^{(0)}} \nu_{e^{(0)}} R_{e^{(0)}}^T (K_{e^{(0)}} + \tilde{\mu}_{e^{(0)}} M_{e^{(0)}}) R_{e^{(0)}}, \end{aligned} \quad (5.4)$$

where $\tilde{\mu}_{e^{(0)}} = \mu_{e^{(0)}} / \nu_{e^{(0)}} \geq 0$ and $\nu_{e^{(0)}} > 0$. We denote by $R_{e^{(0)}}$ the restriction mapping of a global vector of unknowns at the coarsest level to the local vector corresponding to $e^{(0)} \in \mathcal{T}^{(0)}$.

5.2 Robust AMLI algorithms for conforming linear finite elements

Since the system matrices on the coarsest level satisfy the assembling property (5.4), we consider the element system matrix

$$A_{e^{(0)}} = K_{e^{(0)}} + \tilde{\mu}_{e^{(0)}} M_{e^{(0)}}$$

for an arbitrary element $e^{(0)} \in \mathcal{T}^{(0)}$, which is a weighted sum of stiffness and mass matrices. The analysis of uniform local bounds has to be carried out for an element matrix corresponding to an arbitrary triangle denoted by $e \in \mathcal{T}^{(\ell)}$.

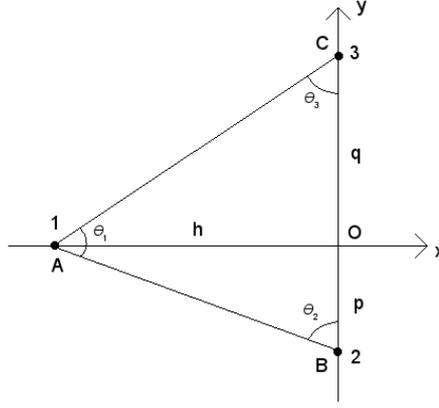


Figure 5.1: An arbitrary non-degenerate triangle $e \in \mathcal{T}^{(\ell)}$.

The following lemma regarding the element stiffness matrix for the Laplace operator can be found in, e.g., [9, 101, 122].

Lemma 5.2. *The element stiffness matrix K_e for the Laplace operator can be written in the general form*

$$K_e = \frac{1}{2} \begin{pmatrix} b+c & -c & -b \\ -c & a+c & -a \\ -b & -a & a+b \end{pmatrix}, \quad (5.5)$$

where a, b and c are equal to the cotangent of the angles in the triangle e , i.e.,

$$a = \cot \theta_1, \quad b = \cot \theta_2, \quad c = \cot \theta_3.$$

Proof. See, e.g., [101]. □

Lemma 5.3. *The element mass matrix M_e can be written in the general form*

$$M_e = \frac{h^2(b+c)}{24} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad (5.6)$$

where a, b and c are equal to the cotangent of the angles in the triangle e , i.e.,

$$a = \cot \theta_1, \quad b = \cot \theta_2, \quad c = \cot \theta_3$$

and h is the triangle height measured perpendicular to the side BC , where B and C are the vertices with the angles θ_2 and θ_3 , respectively.

Proof. The element mass matrix for a given arbitrary non-degenerate triangle e is given by

$$M_e(u, v) = \int_e uv \, de.$$

We introduce the notations $h = |OA|$, $p = |OB|$ and $q = |OC|$, where O is the origin, see Figure 5.1. Then we have the following relations given:

$$b = \frac{p}{h}, \quad c = \frac{q}{h}, \quad a = \cot(\pi - (\theta_2 + \theta_3)) = \frac{h^2 - pq}{h(p+q)}.$$

The element basis functions are given by

$$\varphi_1 = -\frac{x}{h}, \quad \varphi_2 = \frac{qx + h(q-y)}{h(p+q)}, \quad \varphi_3 = \frac{px + h(p+y)}{h(p+q)}.$$

Moreover,

$$|e| = \int de = \frac{h(p+q)}{2} = J_e \cdot \frac{1}{2},$$

where $J_e = h^2(b+c)$ is the Jacobi determinant. We obtain the following first two entries of the element mass matrix:

$$M_{e11} = \int_{-h}^0 \int_{-p/hx-p}^{q/hx+q} (\varphi_1)^2 \, dy \, dx = \frac{h^2(b+c)}{12}$$

and

$$M_{e12} = \int_{-h}^0 \int_{-p/hx-p}^{q/hx+q} \varphi_1 \varphi_2 \, dy \, dx = \frac{h^2(b+c)}{24}.$$

Analogously, we obtain all other entries and we finally derive the element mass matrix (5.6). \square

We assume without loss of generality that

$$|a| \leq b \leq c.$$

Moreover, we define $\alpha = a/c$ and $\beta = b/c$ and obtain the following representations for the element stiffness and mass matrices:

$$K_e = \frac{c}{2} \begin{pmatrix} \beta+1 & -1 & -\beta \\ -1 & \alpha+1 & -\alpha \\ -\beta & -\alpha & \alpha+\beta \end{pmatrix}$$

and

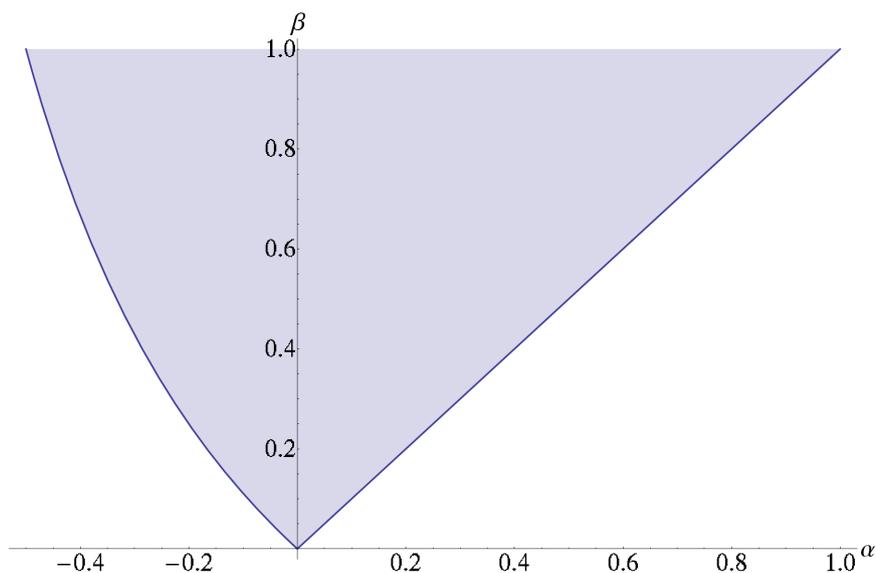
$$M_e = \frac{h^2 c (\beta+1)}{24} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

where $(\alpha, \beta) \in D$ with

$$D = \left\{ (\alpha, \beta) \in \mathbb{R}^2 : -\frac{1}{2} < \alpha \leq 1, \max\left\{-\frac{\alpha}{\alpha+1}, |\alpha|\right\} \leq \beta \leq 1 \right\}, \quad (5.7)$$

see [101] and the references therein. The domain D is illustrated in Figure 5.2.

In case of discretizing diffusion problems by conforming linear finite elements, the standard choice is a (uniform) 2-refinement, which means that each coarse element is subdivided into four congruent elements in every refinement step. We will consider the general case of an m -refinement, where each element is subdivided into m^2 elements in every refinement step. In the next subsection, we will give the reason why a 2-refinement in general is not sufficient for problems of the form (5.3).

Figure 5.2: Domain D of the parameters (α, β) .

5.2.1 The 2-refinement

Let us consider the two consecutive levels $\ell - 1$ and ℓ . We consider the element stiffness matrix (5.5) for an arbitrary element $e \in \mathcal{T}^{(\ell-1)}$. On the related macro element $E \subset \mathcal{T}^{(\ell)}$, which is obtained by

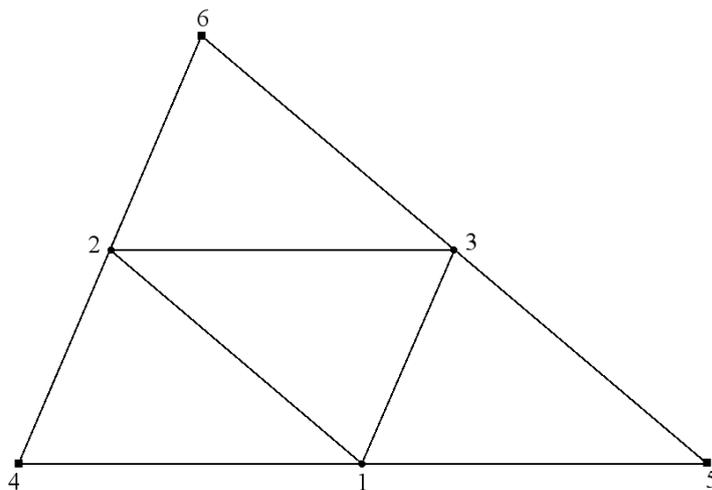


Figure 5.3: Subdivision of one triangle into four congruent ones.

subdivision of the coarse triangle into four congruent triangles, see Figure 5.3, we obtain the following stiffness matrix:

$$K_E = \begin{pmatrix} a+b+c & -a & -b & -\frac{c}{2} & -\frac{c}{2} & 0 \\ -a & a+b+c & -c & -\frac{b}{2} & 0 & -\frac{b}{2} \\ -b & -c & a+b+c & 0 & -\frac{a}{2} & -\frac{a}{2} \\ -\frac{c}{2} & -\frac{b}{2} & 0 & \frac{b+c}{2} & 0 & 0 \\ -\frac{c}{2} & 0 & -\frac{a}{2} & 0 & \frac{a+c}{2} & 0 \\ 0 & -\frac{b}{2} & -\frac{a}{2} & 0 & 0 & \frac{a+b}{2} \end{pmatrix}.$$

The hierarchical stiffness matrix is given by

$$\tilde{K}_E = J^T K_E J = \begin{pmatrix} \tilde{K}_{E:11} & \tilde{K}_{E:12} \\ \tilde{K}_{E:21} & \tilde{K}_{E:22} \end{pmatrix}$$

with

$$J = \begin{pmatrix} I & J_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$\begin{aligned} \tilde{K}_{E:11} &= \begin{pmatrix} a+b+c & -a & -b \\ -a & a+b+c & -c \\ -b & -c & a+b+c \end{pmatrix}, & \tilde{K}_{E:12} &= -\frac{1}{2} \begin{pmatrix} -b & -a & a+b \\ -c & a+c & -a \\ b+c & -c & b \end{pmatrix}, \\ \tilde{K}_{E:21} &= (\tilde{K}_{E:12})^T, & \tilde{K}_{E:22} &= K_e. \end{aligned}$$

In [122], the authors proved the following explicit formula for the local CBS constant $\gamma_{K,E}$ in case of a 2-refinement:

$$\gamma_{K,E}^2 = \frac{3}{8} + \frac{1}{4} \sqrt{\sum_{i=1}^3 \cos^2 \theta_i - \frac{3}{4}}. \quad (5.8)$$

One way to derive this result is to solve the generalized eigenvalue problem (2.62) and then to compute $\gamma_{K,E}$ via the rule (2.61). From (5.8) follows that

$$\gamma_{K,E}^2 = 3/4$$

in the worst case.

Analogously, we can compute the CBS constant for the mass matrix. The mass matrix on a macro element $E \subset \mathcal{T}^{(\ell)}$ is given by

$$M_E = \frac{h^2(b+c)}{24} \begin{pmatrix} 6 & 2 & 2 & 1 & 1 & 0 \\ 2 & 6 & 2 & 1 & 0 & 1 \\ 2 & 2 & 6 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 & 2 \end{pmatrix}$$

and the hierarchical mass matrix by

$$\tilde{M}_E = J^T M_E J = \begin{pmatrix} \tilde{M}_{E:11} & \tilde{M}_{E:12} \\ \tilde{M}_{E:21} & \tilde{M}_{E:22} \end{pmatrix},$$

where

$$\tilde{M}_{E:11} = \frac{h^2(b+c)}{24} \begin{pmatrix} 6 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 6 \end{pmatrix}, \quad \tilde{M}_{E:12} = (\tilde{M}_{E:21})^T = \frac{h^2(b+c)}{24} \begin{pmatrix} 5 & 5 & 2 \\ 5 & 2 & 5 \\ 2 & 5 & 5 \end{pmatrix}$$

and

$$\tilde{M}_{E:22} = 4 M_e.$$

We solve the generalized eigenvalue problem

$$S_M \mathbf{v}_{E:2} = \lambda \tilde{M}_{E:22} \mathbf{v}_{E:2}, \quad (5.9)$$

where

$$S_M = M_{E:22} - M_{E:21} M_{E:11}^{-1} M_{E:12} = \tilde{S}_M$$

is the Schur complement of the mass matrix, and obtain the eigenvalue $7/16$ twice and the eigenvalue $1/10$ once (independent of all parameters a, b, c !). The local CBS constant for the mass matrix therefore is given by

$$\gamma_{M,E} = \sqrt{1 - \min \left\{ \frac{7}{16}, \frac{1}{10} \right\}} = \sqrt{1 - \frac{1}{10}} = \sqrt{\frac{9}{10}}.$$

Hence, we obtain the following parameter-robust estimate for the CBS constant of the weighted sum $A_E = K_E + \tilde{\mu}_E M_E$:

$$\gamma_{A,E}^2 \leq \max \{ \gamma_{K,E}^2, \gamma_{M,E}^2 \} = \max \left\{ \frac{3}{4}, \frac{9}{10} \right\} = \frac{9}{10}, \quad (5.10)$$

see also [39]. We observe that the estimate (5.10) does not imply the optimality conditions (2.71) of the linear AMLI method because there is no integer degree ν of the stabilization polynomial satisfying

$$\frac{1}{\sqrt{1 - 9/10}} = \sqrt{10} \approx 3.16228 < \nu < \varrho = m^2 = 4. \quad (5.11)$$

This is the reason to consider m -refinements for $m > 2$. In the next subsection, we analyze the case $m = 3$.

5.2.2 The 3-refinement

In case of a 3-refinement, see Figure 5.4, one macro element $E \subset \mathcal{T}^{(\ell)}$ is subdivided into nine congruent triangles.

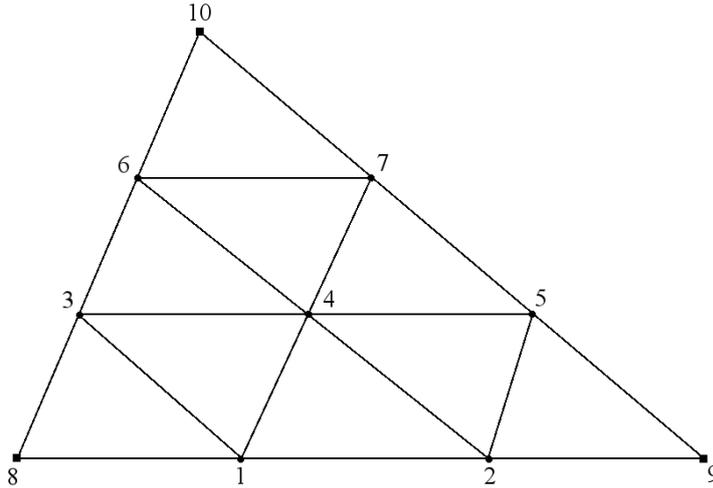


Figure 5.4: Subdivision of one triangle into nine congruent ones.

Denoting

$$s = a + b + c,$$

the corresponding stiffness matrix K_E and its hierarchical stiffness matrix \tilde{K}_E on the macro element are then given by

$$K_E = \begin{pmatrix} K_{E:11} & K_{E:12} \\ K_{E:21} & K_{E:22} \end{pmatrix} \quad \text{and} \quad \tilde{K}_E = J^T K_E J,$$

where

$$K_{E:11} = \begin{pmatrix} s & -\frac{c}{2} & -a & -b & 0 & 0 & 0 \\ -\frac{c}{2} & s & 0 & -a & -b & 0 & 0 \\ -a & 0 & s & -c & 0 & -\frac{b}{2} & 0 \\ -b & -a & -c & 2s & -c & -a & -b \\ 0 & -b & 0 & -c & s & 0 & -\frac{a}{2} \\ 0 & 0 & -\frac{b}{2} & -a & 0 & s & -c \\ 0 & 0 & 0 & -b & -\frac{a}{2} & -c & s \end{pmatrix}, \quad K_{E:12} = \begin{pmatrix} -\frac{c}{2} & 0 & 0 \\ 0 & -\frac{c}{2} & 0 \\ -\frac{b}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -\frac{a}{2} & 0 \\ 0 & 0 & -\frac{b}{2} \\ 0 & 0 & -\frac{a}{2} \end{pmatrix},$$

$$K_{E:21} = (K_{E:12})^T, \quad K_{E:22} = \begin{pmatrix} \frac{b+c}{2} & 0 & 0 \\ 0 & \frac{a+c}{2} & 0 \\ 0 & 0 & \frac{a+b}{2} \end{pmatrix}.$$

Here, the transformation matrix J is given by

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix},$$

and the blocks of the hierarchical stiffness matrix \tilde{K}_E are as follows

$$\tilde{K}_{E:11} = K_{E:11}, \quad \tilde{K}_{E:12} = (\tilde{K}_{E:21})^T = -\frac{1}{3} \begin{pmatrix} -b & -a & a+b \\ -b & -a & a+b \\ -c & a+c & -a \\ 0 & 0 & 0 \\ b+c & -c & -b \\ -c & a+c & -a \\ b+c & -c & -b \end{pmatrix}, \quad \tilde{K}_{E:22} = K_e.$$

In [10], the authors proved the following estimate of the CBS constant of a macro element stiffness matrix arising from a uniform m -refinement:

$$\gamma_{K,E} \leq \sqrt{\frac{m^2 - 1}{m^2}}. \quad (5.12)$$

Hence, for $m = 3$, we have the estimate

$$\gamma_{K,E} \leq \sqrt{\frac{8}{9}}. \quad (5.13)$$

Now, we estimate the CBS constant for the corresponding macro element mass matrix, where we use the node numbering as shown in Figure 5.4. First, we find

$$M_E = \frac{h^2(b+c)}{24} \begin{pmatrix} 6 & 1 & 2 & 2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 6 & 0 & 2 & 2 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 6 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\ 2 & 2 & 2 & 12 & 2 & 2 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & 6 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 6 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 & 2 & 6 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 2 \end{pmatrix}$$

and the hierarchical macro element mass matrix

$$\tilde{M}_E = J^T M_E J = \begin{pmatrix} \tilde{M}_{E:11} & \tilde{M}_{E:12} \\ \tilde{M}_{E:21} & \tilde{M}_{E:22} \end{pmatrix},$$

where $\tilde{M}_{E:11} = M_{E:11}$, $\tilde{M}_{E:22} = 9 M_e$ and

$$\tilde{M}_{E:12} = (\tilde{M}_{E:21})^T = \frac{h^2(b+c)}{24} \begin{pmatrix} \frac{22}{3} & \frac{10}{3} & \frac{4}{3} \\ \frac{10}{3} & \frac{22}{3} & \frac{4}{3} \\ \frac{22}{3} & \frac{4}{3} & \frac{10}{3} \\ \frac{4}{3} & \frac{22}{3} & \frac{10}{3} \\ \frac{10}{3} & \frac{4}{3} & \frac{22}{3} \\ \frac{4}{3} & \frac{10}{3} & \frac{22}{3} \\ \frac{22}{3} & \frac{4}{3} & \frac{10}{3} \\ \frac{4}{3} & \frac{22}{3} & \frac{10}{3} \\ \frac{10}{3} & \frac{4}{3} & \frac{22}{3} \\ \frac{4}{3} & \frac{10}{3} & \frac{22}{3} \end{pmatrix}.$$

We solve the eigenvalue problem (5.9) corresponding to the 3-refinement and obtain the two-fold eigenvalue $19/99$ and the eigenvalue $1/21$ which are again independent of the parameters a , b , and c . Hence, the local CBS constant for the mass matrix can be computed via the rule (2.61) and is given by

$$\gamma_{M,E} = \sqrt{1 - \min \left\{ \frac{19}{99}, \frac{1}{21} \right\}} = \sqrt{\frac{20}{21}},$$

and finally, in view of (5.13), the CBS constant of the weighted sum $A_E = K_E + \tilde{\mu}_E M_E$ can be estimated by

$$\gamma_{A,E} \leq \sqrt{\max \left\{ \gamma_{K,E}^2, \gamma_{M,E}^2 \right\}} = \sqrt{\max \left\{ \frac{8}{9}, \frac{20}{21} \right\}} = \sqrt{\frac{20}{21}}. \quad (5.14)$$

From (5.14) we conclude that the optimality condition (2.71) is fulfilled since

$$\frac{1}{\sqrt{1 - 20/21}} = \sqrt{21} \approx 4.58258 < v < \varrho = m^2 = 9 \quad (5.15)$$

holds for polynomial degree $v \in \{5, 6, 7, 8\}$.

In the following subsection, we discuss the estimation of the CBS constant for uniform m -refinements for all $m > 2$. In particular, we present a uniform estimate of the CBS constant for $A_E = K_E + \tilde{\mu}_E M_E$ on a macro element $E \subset \mathcal{T}^{(\ell)}$.

5.2.3 The uniform m -refinement

The CBS constant for the mass matrix for a refinement factor $\rho = m^2$ can be computed in the same way as it was presented in Subsection 5.2.1 and Subsection 5.2.2. In Theorem 5.4, we present a general result for the estimation of the CBS constant of the mass matrix, see [103].

Theorem 5.4 (Kraus and Wolfmayr [103]). *Consider a uniform m -refinement for conforming linear finite elements where $m > 2$. The CBS constant of the mass matrix can be estimated as follows*

$$\gamma_M \leq \sqrt{\frac{12m^2 - 5}{12m^2}}. \quad (5.16)$$

Proof. The global CBS constant can be estimated by the maximum of the local CBS constants on the macro elements (2.60), which can be again computed via the rule (2.61), i.e.,

$$\gamma_{M,E}^2 = 1 - \lambda_E^{\min},$$

where λ_E^{\min} is the minimal eigenvalue of the eigenvalue problem (5.9), which we write in the form

$$\mathbf{v}_{E:2}^T \left(S_M - \lambda \tilde{M}_{E:22} \right) \mathbf{v}_{E:2} = 0,$$

We try to find a lower bound $\underline{\lambda}_E^{\min}$ for the minimal eigenvalue λ_E^{\min} such that

$$\mathbf{v}_{E:2}^T \left(S_M - \underline{\lambda}_E^{\min} \tilde{M}_{E:22} \right) \mathbf{v}_{E:2} \geq 0.$$

For that reason, we estimate the Schur complement S_M from below by \underline{S}_M , and, after that we solve the problem

$$\mathbf{v}_{E:2}^T \left(\underline{S}_M - \lambda \tilde{M}_{E:22} \right) \mathbf{v}_{E:2} = 0.$$

For every m -refinement, we systematically use a bottom-up lexicographical ordering for the fine nodes and number the three coarse nodes last. Hence, the lower right block $M_{E:22}$ of the macro element mass matrix is always

$$M_{E:22} = 2m_c I,$$

where I is the identity matrix and

$$m_c = \frac{h^2(b+c)}{24}.$$

Let n_E denote the number of nodes on a macro element subdivided into m^2 elements. The Schur complement S_M can be estimated from below in the following way:

$$\begin{aligned} S_M &= M_{E:22} - M_{E:21} M_{E:11}^{-1} M_{E:12} \\ &= 2m_c I - M_{E:21} M_{E:11}^{-1} M_{E:12} \\ &\geq 2m_c I - \frac{1}{6} m_c^{-1} M_{E:21} I M_{E:12}, \end{aligned}$$

where we use the estimate

$$M_{E:11} \geq \min_{i \in \{1, \dots, n_E\}} (M_{E:11})_{ii} I = 6m_c I$$

because the weakly diagonally dominant matrix $M_{E:11}$ has only the diagonal entries $12m_c$ and $6m_c$. Moreover, since the matrices $M_{E:12}$ and $M_{E:21} = (M_{E:12})^T$ have exactly two entries equal to one in each of the three columns and rows, respectively, we obtain

$$M_{E:21} M_{E:12} = (M_{E:12})^T M_{E:12} = 2m_c^2 I.$$

Finally, the Schur complement S_M can be estimated from below by

$$\begin{aligned} S_M &\geq 2m_c I - \frac{1}{6}m_c^{-1}2m_c^2 I \\ &= 2m_c I - \frac{1}{3}m_c I \\ &= \frac{5}{3}m_c I. \end{aligned}$$

Next, we have that

$$\tilde{M}_{E:22} = M_{E:22} + J_{E:12}^T M_{E:12} + M_{E:21} J_{E:12} + J_{E:12}^T M_{E:12} J_{E:12} = m^2 M_e$$

with a transformation matrix $J_{E:12}$ of the form (2.64). Then

$$J_{E:12}^T M_{E:12} = m_c \begin{pmatrix} \frac{2m-2}{m} & \frac{1}{m} & \frac{1}{m} \\ \frac{1}{m} & \frac{2m-2}{m} & \frac{1}{m} \\ \frac{1}{m} & \frac{1}{m} & \frac{2m-2}{m} \end{pmatrix}$$

because $M_{E:12}$ has exactly two entries with value one in each row and $J_{E:12}$ has the value $(m-1)/m$ in exactly the same positions. Hence, we solve the problem

$$\mathbf{v}_{E:2}^T \left(\frac{5}{3}m_c I - \lambda m^2 M_e \right) \mathbf{v}_{E:2} = 0$$

and obtain the two-fold eigenvalue $5/(3m^2)$ and the eigenvalue $5/(12m^2)$, which yields the lower bound for λ_E^{\min} . According to (2.61), we obtain the estimate

$$\gamma_{M,E}^2 = 1 - \lambda_E^{\min} \leq 1 - \frac{5}{12m^2} = \frac{12m^2 - 5}{12m^2},$$

which together with (2.60) gives the upper bound (5.16). \square

Remark 5.5. If the parameter $\tilde{\mu}_E = 0$, i.e., $A_E = K_E$, then the CBS constant can be estimated by the formula (5.12), see [10]. Moreover, the estimate (5.16) at the same time provides a bound for the local CBS constant $\gamma_{A,E}$ corresponding to the weighted sum of mass and stiffness matrix, since together with (5.12), we obtain

$$\gamma_{A,E} \leq \max \{ \gamma_{K,E}, \gamma_{M,E} \} \leq \max \left\{ \sqrt{\frac{m^2 - 1}{m^2}}, \sqrt{\frac{12m^2 - 5}{12m^2}} \right\} = \sqrt{\frac{12m^2 - 5}{12m^2}}. \quad (5.17)$$

By applying estimate (5.16) of Theorem 5.4 and using (5.12), we obtain the following estimates for the CBS constants in case of m -refinements for $m = 3, 4, 5$:

$$\gamma_{A,E}^2 \leq \max \{ \gamma_{K,E}^2, \gamma_{M,E}^2 \} \leq \begin{cases} \max \left\{ \frac{8}{9}, \frac{103}{108} \right\} \approx 0.953704 & \text{for } m = 3, \\ \max \left\{ \frac{15}{16}, \frac{187}{192} \right\} \approx 0.973958 & \text{for } m = 4, \\ \max \left\{ \frac{24}{25}, \frac{59}{60} \right\} \approx 0.983333 & \text{for } m = 5. \end{cases} \quad (5.18)$$

For comparison, we compute also the sharp bounds for the CBS constants $\gamma_{K,E}$ and $\gamma_{M,E}$ up to a 5-refinement. The cases $m = 2$ and $m = 3$ have already been worked out in Subsections 5.2.1 and 5.2.2. In an analogous manner, one computes the estimates for the cases $m = 4$ and $m = 5$ using again a bottom-up lexicographical ordering for the fine nodes and numbering the three coarse nodes last.

For $m = 4$, we obtain the following block J_{12} of the transformation matrix J :

$$J_{12} = \begin{pmatrix} \frac{3}{4} & \frac{1}{2} & \frac{1}{4} & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{3}{4} & 0 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} & 0 & \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} & \frac{3}{4} \end{pmatrix}^T.$$

The resulting hierarchical macro element mass matrix is given by

$$\tilde{M}_E = J^T M_E J = \begin{pmatrix} \tilde{M}_{E:11} & \tilde{M}_{E:12} \\ \tilde{M}_{E:21} & \tilde{M}_{E:22} \end{pmatrix},$$

where $\tilde{M}_{E:22} = 16 M_e$,

$$\tilde{M}_{E:12} = (\tilde{M}_{E:21})^T = \frac{h^2(b+c)}{24} \begin{pmatrix} \frac{17}{2} & \frac{11}{2} & \frac{5}{2} & \frac{17}{2} & 12 & 6 & 1 & \frac{11}{2} & 6 & 1 & \frac{5}{2} & 1 \\ \frac{5}{2} & \frac{11}{2} & \frac{17}{2} & \frac{1}{2} & 6 & 12 & \frac{17}{2} & \frac{1}{2} & 6 & \frac{11}{2} & \frac{1}{2} & \frac{5}{2} \\ \frac{2}{2} & \frac{2}{2} & \frac{2}{2} & \frac{5}{2} & 6 & 6 & \frac{2}{2} & \frac{11}{2} & 12 & \frac{11}{2} & \frac{17}{2} & \frac{17}{2} \\ 1 & 1 & 1 & \frac{5}{2} & 6 & 6 & \frac{2}{2} & \frac{11}{2} & 12 & \frac{11}{2} & \frac{17}{2} & \frac{17}{2} \end{pmatrix}^T$$

and

$$\tilde{M}_{E:11} = \frac{h^2(b+c)}{24} \begin{pmatrix} 6 & 1 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 6 & 1 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 6 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 2 & 12 & 2 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 2 & 12 & 2 & 0 & 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 2 & 6 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 6 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 2 & 0 & 2 & 12 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 2 & 6 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 6 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 6 & 6 \end{pmatrix}.$$

Analogously, one determines the transformation matrix J , the macro element mass matrix M_E and the hierarchical mass matrix \tilde{M}_E for the 5-refinement. For $m = 5$, we obtain the following block J_{12} of the transformation matrix J :

$$J_{12} = \begin{pmatrix} \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} & \frac{4}{5} & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{3}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} & 0 \\ \frac{1}{5} & \frac{3}{5} & \frac{2}{5} & \frac{4}{5} & 0 & \frac{1}{5} & \frac{3}{5} & \frac{2}{5} & \frac{4}{5} & 0 & \frac{1}{5} & \frac{3}{5} & \frac{2}{5} & 0 & \frac{1}{5} & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{4}{5} & \frac{1}{5} & 0 \end{pmatrix}^T.$$

The resulting hierarchical macro element mass matrix is given by

$$\tilde{M}_E = J^T M_E J = \begin{pmatrix} \tilde{M}_{E:11} & \tilde{M}_{E:12} \\ \tilde{M}_{E:21} & \tilde{M}_{E:22} \end{pmatrix},$$

where $\tilde{M}_{E:22} = 25 M_e$, $\tilde{M}_{E:12} = (\tilde{M}_{E:21})^T = \frac{h^2(b+c)}{24} \tilde{N}$ with

$$\tilde{N} = \begin{pmatrix} \frac{46}{5} & \frac{34}{5} & \frac{22}{5} & 2 & \frac{46}{5} & \frac{72}{5} & \frac{48}{5} & \frac{24}{5} & \frac{4}{5} & \frac{34}{5} & \frac{48}{5} & \frac{24}{5} & \frac{4}{5} & \frac{22}{5} & \frac{24}{5} & \frac{4}{5} & 2 & \frac{4}{5} \\ \frac{2}{5} & \frac{22}{5} & \frac{34}{5} & \frac{46}{5} & \frac{2}{5} & \frac{24}{5} & \frac{48}{5} & \frac{72}{5} & \frac{46}{5} & \frac{4}{5} & \frac{24}{5} & \frac{48}{5} & \frac{34}{5} & \frac{22}{5} & \frac{24}{5} & \frac{4}{5} & \frac{4}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{4}{5} & \frac{4}{5} & \frac{4}{5} & 2 & \frac{24}{5} & \frac{24}{5} & \frac{24}{5} & 2 & \frac{22}{5} & \frac{48}{5} & \frac{48}{5} & \frac{22}{5} & \frac{34}{5} & \frac{34}{5} & \frac{4}{5} & \frac{4}{5} & \frac{46}{5} \\ \frac{4}{5} & \frac{4}{5} & \frac{4}{5} & \frac{4}{5} & 2 & \frac{24}{5} & \frac{24}{5} & \frac{24}{5} & 2 & \frac{22}{5} & \frac{48}{5} & \frac{48}{5} & \frac{22}{5} & \frac{34}{5} & \frac{34}{5} & \frac{4}{5} & \frac{4}{5} & \frac{46}{5} \end{pmatrix}^T,$$

and

$$\tilde{M}_{E:11} = \frac{h^2(b+c)}{24} \begin{pmatrix} 6 & 1 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 6 & 1 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 6 & 1 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 6 & 0 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 6 & 2 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 2 & 12 & 2 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 2 & 12 & 2 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 2 & 12 & 2 & 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 6 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 6 & 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 & 0 & 0 & 2 & 12 & 2 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 0 & 0 & 2 & 12 & 2 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 & 2 & 6 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 6 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 0 & 2 & 12 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 2 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 2 & 6 \end{pmatrix}.$$

Finally, one solves the corresponding eigenvalue problem (5.9) for $m = 4, 5$, and uses the minimal eigenvalues to compute the CBS constants via the rule (2.61). The resulting sharp estimates of the CBS constants for problems of the form (5.3) corresponding to m -refinements for $m \in \{2, 3, 4, 5\}$ (under the assumption (5.4)) are as follows

$$\gamma_{A,E}^2 \leq \max \{ \gamma_{K,E}^2, \gamma_{M,E}^2 \} = \begin{cases} \max \{ \frac{3}{4}, \frac{9}{10} \} = 0.9 & \text{for } m = 2, \\ \max \{ \frac{8}{9}, \frac{20}{21} \} \approx 0.952381 & \text{for } m = 3, \\ \max \{ \frac{15}{16}, \frac{36}{37} \} \approx 0.972973 & \text{for } m = 4, \\ \max \{ \frac{24}{25}, \frac{11916}{12125} \} \approx 0.982763 & \text{for } m = 5. \end{cases} \quad (5.19)$$

Comparing now (5.18) and (5.19) shows that formula (5.16) provides a very good estimate for the CBS constant of the macro element mass matrix.

Corollary 5.6. *Summarizing our findings, the smallest value for m that guarantees that the optimality conditions can be satisfied with an m -refinement is $m = 3$. Moreover, since*

$$\frac{1}{\sqrt{1 - \frac{12m^2 - 5}{12m^2}}} = 2m \sqrt{\frac{3}{5}} < 2m < m^2 \quad \text{for all } m > 2, \quad (5.20)$$

any m -refinement for $m > 2$ allows to meet the optimality conditions (2.71), see also [103].

In the next section, we will present the construction and the analysis of an additive preconditioner for the pivot block A_{11} arising in the 3-refinement.

5.3 Additive preconditioning of the pivot block

Applying the AMLI method requires the action of (an approximation of) the inverse of the pivot blocks $A_{11}^{(\ell)}$ on a vector. It is well known (see, e.g., [15]) that the (linear) AMLI preconditioner

with approximate pivot block $C_{11}^{(\ell)}$ is optimal if apart from the optimality conditions (2.71) the preconditioners $C_{11}^{(\ell)}$ are spectrally equivalent to $A_{11}^{(\ell)}$ on all levels ℓ , i.e.,

$$C_{11}^{(\ell)} \approx A_{11}^{(\ell)},$$

in the sense of (2.70) and their action on a vector has linear complexity, i.e., requires $\mathcal{O}(n^{(\ell)})$ arithmetic operations.

Here, we generalize the additive preconditioner $C_{11}^{(\ell)}$, which was proposed in [13] for the 2-refinement, for the 3-refinement and derive the corresponding condition number bounds. The construction as well as the analysis of $C_{11}^{(\ell)}$ relies on a macro-element-by-macro-element assembling procedure, i.e.,

$$A_{11}^{(\ell)} = \sum_{E \in \mathcal{T}^{(\ell)}} R_E^T A_{E:11} R_E \quad (5.21)$$

and

$$C_{11}^{(\ell)} = \sum_{E \in \mathcal{T}^{(\ell)}} R_E^T C_{E:11} R_E. \quad (5.22)$$

The pivot block of the macro element matrices is given by

$$A_{E:11} = K_{E:11} + \tilde{\mu}_E M_{E:11}.$$

The idea is to construct an additive preconditioner $C_{E:11}$ having the form

$$C_{E:11} = C_{E:11}^K + \tilde{\mu}_E C_{E:11}^M$$

with the same weighting as the pivot block $A_{E:11}$ and where the matrices $C_{E:11}^K$ and $C_{E:11}^M$ have the same structure, i.e., the same non-zero pattern, in order to implement the preconditioner $C_{E:11}$. We obtain the preconditioners $C_{E:11}^K$ and $C_{E:11}^M$ by preserving the largest (in magnitude) off-diagonal entries of $K_{E:11}$ and $M_{E:11}$, respectively. *Note that the same nonzero pattern is chosen for the preconditioner of the stiffness matrix pivot block and for the one of the mass matrix pivot block!* The couplings corresponding to the largest (in magnitude) off-diagonal entries are shown in Figure 5.5.

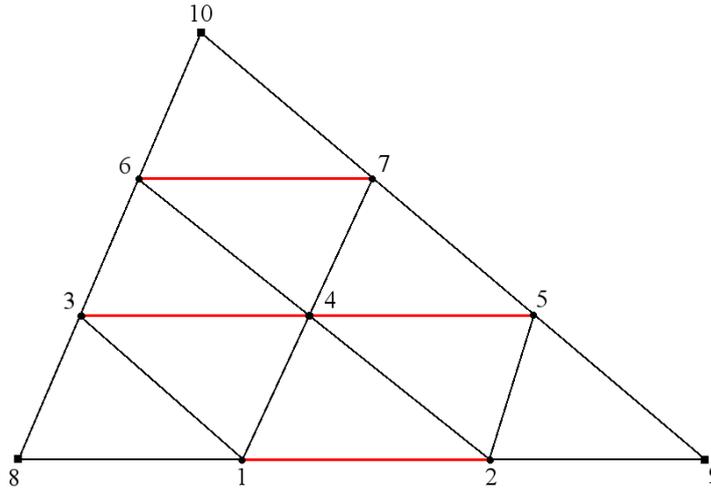


Figure 5.5: Couplings corresponding to the largest entries in the macro element pivot block (red) for the 3-refinement.

We start with the computation and analysis of the additive preconditioner $C_{E:11}^M$ for the pivot block of the macro element mass matrix $M_{E:11}$.

5.3.1 Additive preconditioning for the pivot block of the mass matrix

Let us consider the element mass matrix (5.6) and let $\alpha = a/c$ and $\beta = b/c$ with $|a| \leq b \leq c$, where $(\alpha, \beta) \in D$ as illustrated in Figure 5.2. Then the pivot block of the macro element mass matrix corresponding to the node numbering presented in Figure 5.4 is given by

$$M_{E:11} = \frac{h^2 c (\beta + 1)}{24} \begin{pmatrix} 6 & 1 & 2 & 2 & 0 & 0 & 0 \\ 1 & 6 & 0 & 2 & 2 & 0 & 0 \\ 2 & 0 & 6 & 2 & 0 & 1 & 0 \\ 2 & 2 & 2 & 12 & 2 & 2 & 2 \\ 0 & 2 & 0 & 2 & 6 & 0 & 1 \\ 0 & 0 & 1 & 2 & 0 & 6 & 2 \\ 0 & 0 & 0 & 2 & 1 & 2 & 6 \end{pmatrix}.$$

We choose the following additive preconditioner by preserving only the largest (in magnitude) off-diagonal entries as illustrated in Figure 5.5:

$$C_{E:11}^M = \frac{h^2 c (\beta + 1)}{24} \begin{pmatrix} 6 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 12 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 6 \end{pmatrix}. \quad (5.23)$$

Hence, the additive preconditioner of M_{11} is defined via the macro-element-by-macro-element assembling by

$$C_{11}^M = \sum_{E \in \mathcal{T}^{(\ell)}} R_E^T C_{E:11}^M R_E. \quad (5.24)$$

Theorem 5.7 (Kraus and Wolfmayr [103]). *The additive preconditioner (5.24) of M_{11} with (5.23) yields a relative condition number uniformly bounded by*

$$\kappa((C_{11}^M)^{-1} M_{11}) \leq \frac{1 + \sqrt{\frac{205 + 3\sqrt{3873}}{1792}}}{1 - \sqrt{\frac{205 + 3\sqrt{3873}}{1792}}} \approx 2.75607, \quad (5.25)$$

which holds independently of the shape, the size of each element and of the problem parameters ν and μ , cf. (5.3).

Proof. In order to obtain the relative condition number of the preconditioned system $\kappa((C_{11}^M)^{-1} M_{11})$, we have to solve the local eigenproblem

$$M_{E:11} v_E = \lambda_E C_{E:11}^M v_E.$$

In the characteristic equation

$$\det(M_{E:11} - \lambda_E C_{E:11}^M) = 0,$$

we substitute $\lambda_E = 1 - \mu_E$ and obtain the equation

$$(\beta + 1) c h \mu_E (40 \mu_E^2 - 7) (896 \mu_E^4 - 205 \mu_E^2 + 2) = 0.$$

Since $c \neq 0$, $h \neq 0$ and $\beta \neq -1$, we obtain the following solutions of this equation:

$$\begin{aligned}\mu_E^{(1)} &= 0, \\ \mu_E^{(2/3)} &= \pm \frac{1}{2} \sqrt{\frac{7}{10}} \approx \pm 0.41833, \\ \mu_E^{(4/5)} &= \pm \frac{1}{16} \sqrt{\frac{1}{7} (205 - 3\sqrt{3873})} \approx \pm 0.101054, \\ \mu_E^{(6/7)} &= \pm \sqrt{\frac{205}{1792} + \frac{3\sqrt{3873}}{1792}} \approx \pm 0.467528.\end{aligned}$$

Hence, the largest and smallest eigenvalues λ_E^{\max} and λ_E^{\min} are given by

$$\lambda_E^{\max} = 1 + \sqrt{\frac{205}{1792} + \frac{3\sqrt{3873}}{1792}} \approx 1.46753$$

and

$$\lambda_E^{\min} = 1 - \sqrt{\frac{205}{1792} + \frac{3\sqrt{3873}}{1792}} \approx 0.532472,$$

respectively. Thus, we obtain

$$1 - \sqrt{\frac{205 + 3\sqrt{3873}}{1792}} \leq \lambda_E \leq 1 + \sqrt{\frac{205 + 3\sqrt{3873}}{1792}}. \quad (5.26)$$

Finally, using (5.26) together with (5.24) it follows that the relative condition number estimate (5.25) holds. \square

In the next subsection, we compute and analyze the additive preconditioner for the pivot block of the stiffness matrix.

5.3.2 Additive preconditioner for the pivot block of the stiffness matrix

Let us consider now the element stiffness matrix (5.5) and let again $\alpha = a/c$ and $\beta = b/c$ with $|a| \leq b \leq c$, where $(\alpha, \beta) \in D$ as illustrated in Figure 5.2. Then the pivot block of the macro element stiffness matrix corresponding to the node numbering presented in Figure 5.4 is given by

$$K_{E:11} = c \begin{pmatrix} \sigma & -\frac{1}{2} & -\alpha & -\beta & 0 & 0 & 0 \\ -\frac{1}{2} & \sigma & 0 & -\alpha & -\beta & 0 & 0 \\ -\alpha & 0 & \sigma & -1 & 0 & -\frac{\beta}{2} & 0 \\ -\beta & -\alpha & -1 & 2\sigma & -1 & -\alpha & -\beta \\ 0 & -\beta & 0 & -1 & \sigma & 0 & -\frac{\alpha}{2} \\ 0 & 0 & -\frac{\beta}{2} & -\alpha & 0 & \sigma & -1 \\ 0 & 0 & 0 & -\beta & -\frac{\alpha}{2} & -1 & \sigma \end{pmatrix},$$

where $\sigma = \alpha + \beta + 1$. We define the additive preconditioner of K_{11} by

$$C_{11}^K = \sum_{E \in \mathcal{T}^{(\ell)}} R_E^T C_{E:11}^K R_E \quad (5.27)$$

with

$$C_{E:11}^K = c \begin{pmatrix} \sigma & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & \sigma & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2\sigma & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & \sigma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & \sigma \end{pmatrix}, \quad (5.28)$$

where we have preserved only the largest (in magnitude) off-diagonal entries of the macro element stiffness matrix pivot block $K_{E:11}$ in order to get $C_{E:11}^K$ as illustrated in Figure 5.5. Note that this preconditioner has the same nonzero pattern as the one for the pivot block of the mass matrix.

Theorem 5.8 (Kraus and Wolfmayr [103]). *The additive preconditioner (5.27) of K_{11} with (5.28) yields a relative condition number uniformly bounded by*

$$\kappa((C_{11}^K)^{-1}K_{11}) \leq \frac{1 + \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}}{1 - \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}} \approx 10.7185. \quad (5.29)$$

which holds independently of the shape, the size of each element and of the problem parameters ν and μ , cf. (5.3).

Proof. In order to estimate the condition number of the preconditioned pivot block K_{11} we consider the local generalized eigenproblem

$$K_{E:11} v_{E:1} = \lambda_E C_{E:11}^K v_{E:1}. \quad (5.30)$$

We rewrite (5.30) in the form

$$v_{E:1}^T (K_{E:11} - \lambda_E C_{E:11}^K) v_{E:1} = 0,$$

substitute $\lambda_E = 1 - \mu_E$ and define

$$cP(\mu_E, \alpha, \beta) = K_{E:11} - (1 - \mu_E) C_{E:11}^K$$

with

$$P(\mu_E, \alpha, \beta) := P_0(\mu_E) + \alpha P_\alpha(\mu_E) + \beta P_\beta(\mu_E),$$

where

$$P_0(\mu_E) := \mu_E \begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix},$$

$$P_\alpha(\mu_E) := \begin{pmatrix} \mu_E & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & \mu_E & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & \mu_E & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 2\mu_E & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & \mu_E & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & -1 & 0 & \mu_E & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & \mu_E \end{pmatrix}$$

and

$$P_\beta(\mu_E) := \begin{pmatrix} \mu_E & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & \mu_E & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & \mu_E & 0 & 0 & -\frac{1}{2} & 0 \\ -1 & 0 & 0 & 2\mu_E & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & \mu_E & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & \mu_E & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & \mu_E \end{pmatrix}$$

do not depend on α and β . Since $P(\mu_E, \alpha, \beta)$ depends linearly on α and β , this matrix-valued function can be only maximal or minimal (in an symmetric positive semidefinite sense) on the boundary of the domain D defined in (5.7), see also Figure 5.2. Hence, either for $-\frac{1}{2} < \alpha \leq 0$ and $\beta = -\frac{\alpha}{\alpha+1}$ or for $\alpha = \beta = 1$. It remains to determine the corresponding value μ_E . Let us firstly consider the simpler case $\alpha = \beta = 1$.

In the case $\alpha = \beta = 1$, we solve the characteristic equation corresponding to problem (5.30) which yields the equation

$$\mu_E (107520\mu_E^6 - 87408\mu_E^4 + 9369\mu_E^2 - 78) = 0$$

and has the following solutions:

$$\begin{aligned} \mu_E^{(1)} &= 0, \\ \mu_E^{(2/3)} &= \pm \frac{1}{4} \sqrt{\frac{13}{7}} \approx \pm 0.340693, \\ \mu_E^{(4/5)} &= \pm \frac{1}{8} \sqrt{\frac{1}{10} (223 - 3\sqrt{5241})} \approx \pm 0.0953263, \\ \mu_E^{(6/7)} &= \pm \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}} \approx \pm 0.82933. \end{aligned}$$

Hence, the local largest and smallest eigenvalues corresponding to the case $\alpha = \beta = 1$ are given by

$$\begin{aligned} \lambda_E^{\max} &= 1 + \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}} \approx 1.82933, \\ \lambda_E^{\min} &= 1 - \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}} \approx 0.17067, \end{aligned}$$

respectively.

Now, we consider the second case, i.e., $-\frac{1}{2} < \alpha \leq 0$ and $\beta = -\frac{\alpha}{\alpha+1}$. The characteristic equation corresponding to (5.30) together with $c \neq 0$ and with the substitution $\lambda_E = 1 - \mu_E$ is given by

$$\begin{aligned} 0 = \frac{\alpha^4 \mu_E}{32(\alpha+1)^7} & [16(\alpha^2 + \alpha + 1)(2\alpha^2 + \alpha + 1)(\alpha(\alpha+2) + 2)^2(\alpha(2\alpha+3) + 3)\mu_E^6 \\ & - 2\alpha^2(\alpha(\alpha(\alpha(\alpha(\alpha(8\alpha+7) + 191) + 421) + 589) + 457) + 157) \\ & + 4) + 1) - 6(\alpha(\alpha(\alpha(\alpha(\alpha(2\alpha(\alpha(12\alpha+7) + 293) + 647) + 2001) \\ & + 2229) + 1807) + 1026) + 399) + 95) + 19)\mu_E^4 \\ & + 3(\alpha(\alpha(\alpha(\alpha(\alpha(2\alpha(\alpha(16\alpha+7) + 375) + 781) + 2203) + 2119) \\ & + 1397) + 598) + 157) + 5) + 1)\mu_E^2], \end{aligned}$$

where $-\frac{1}{2} < \alpha \leq 0$. The first solution is $\mu_E^{(1)} = 0$ and $\lambda_E^{(1)} = 1$. Moreover, the equation is fulfilled for $\alpha = 0$. So, we consider the case $\alpha \in (-\frac{1}{2}, 0)$. We now substitute $\nu_E = \mu_E^2$ and solve the cubic

equation by Cardano's formula, see, e.g., [34]. The equation has the form

$$A(\alpha)\nu_E^3 + B(\alpha)\nu_E^2 + C(\alpha)\nu_E + D(\alpha) = 0, \quad (5.31)$$

where

$$A(\alpha) = 16(\alpha^2 + \alpha + 1)(2\alpha^2 + \alpha + 1)(\alpha(\alpha + 2) + 2)^2(\alpha(2\alpha + 3) + 3).$$

We divide by $A(\alpha) \neq 0$ and obtain an equation of the form

$$\nu_E^3 + a(\alpha)\nu_E^2 + b(\alpha)\nu_E + c(\alpha) = 0$$

with

$$\begin{aligned} a(\alpha) &= \frac{B(\alpha)}{A(\alpha)} \\ &= \frac{3}{8} \left(-\frac{6(\alpha+1)}{\alpha^2 + \alpha + 1} + \frac{6(\alpha+1)}{2\alpha^2 + \alpha + 1} + \frac{-10\alpha - 7}{\alpha(\alpha+2) + 2} + \frac{26(\alpha+1)}{\alpha(2\alpha+3) + 3} - \frac{3}{(\alpha(\alpha+2) + 2)^2} - 6 \right). \end{aligned}$$

The symmetry of all the involved matrices implies that the cubic equation (5.31) has three real solutions. Using the substitution

$$\nu_E = z - \frac{a(\alpha)}{3},$$

we obtain the equation

$$z^3 + p(\alpha)z + q(\alpha) = 0,$$

where

$$p(\alpha) = b(\alpha) - \frac{a(\alpha)^2}{3} \quad \text{and} \quad q(\alpha) = \frac{2a(\alpha)^3}{27} - \frac{a(\alpha)b(\alpha)}{3} + c(\alpha).$$

The solutions of this equation are given by

$$\begin{aligned} z^{(1)} &= \sqrt{-\frac{4}{3}p(\alpha)} \cos\left(\frac{1}{3}\arccos(r(\alpha))\right), \\ z^{(2)} &= -\sqrt{-\frac{4}{3}p(\alpha)} \cos\left(\frac{1}{3}\arccos(r(\alpha)) + \frac{\pi}{3}\right), \\ z^{(3)} &= -\sqrt{-\frac{4}{3}p(\alpha)} \cos\left(\frac{1}{3}\arccos(r(\alpha)) - \frac{\pi}{3}\right), \end{aligned}$$

where

$$r(\alpha) = -\frac{q(\alpha)}{2} \sqrt{-\frac{27}{p(\alpha)^3}}.$$

Finally, the solutions $\nu_E = \mu_E^2$ are given by

$$\nu_E^{(i)}(\alpha) = z^{(i)} - \frac{a(\alpha)}{3}, \quad i = 1, 2, 3.$$

The three solutions (as functions of α) are illustrated in Figure 5.6. We see that $\nu^{(1)}(\alpha)$ corresponds to the largest value for μ_E and has its maximum for $\alpha = -1/2$. This can be proven by a standard tool from symbolic computation called *Cylindrical Algebraic Decomposition (CAD)*, see [47, 48, 86],

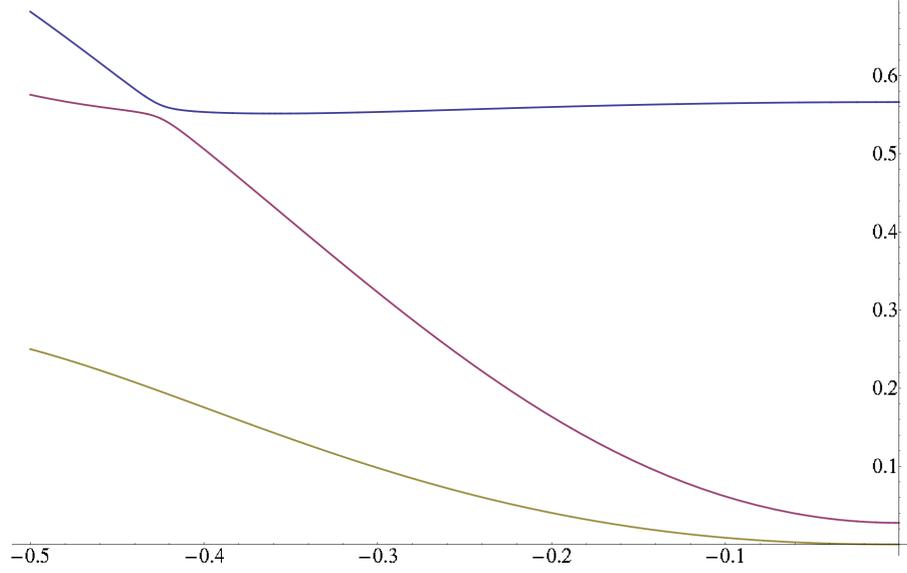


Figure 5.6: The three solutions $\nu^{(1)}(\alpha)$, $\nu^{(2)}(\alpha)$ and $\nu^{(3)}(\alpha)$ for $\alpha \in (-1/2, 0)$.

of which several implementations are available. We used the *Mathematica* commands *CylindricalDecomposition* and *Resolve* for the proof. For $\alpha = -1/2$,

$$\nu^{(1)}(-1/2) = \frac{1}{800} \left(503 + 3\sqrt{201} \right) \approx 0.681915,$$

and the corresponding maximal and minimal eigenvalues are given by

$$\lambda_E^{\max} = 1 + \sqrt{\nu^{(1)}(-1/2)} \approx 1.82578 \quad \text{and} \quad \lambda_E^{\min} = 1 - \sqrt{\nu^{(1)}(-1/2)} \approx 0.174218,$$

respectively. The outcome of comparing these eigenvalues to the ones for the case $\alpha = \beta = 1$ is that the maximal and minimal eigenvalues are both attained for $\alpha = \beta = 1$, which leads to the following local eigenvalue estimate:

$$0.17067 \approx 1 - s < \lambda_E < 1 + s \approx 1.82933, \quad (5.32)$$

where

$$s = \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}.$$

Together with the macro-element-by-macro-element assembling procedure (5.27), we finally arrive at the condition number estimate

$$\kappa((C_{11}^K)^{-1}K_{11}) \leq \frac{1 + \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}}{1 - \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}} \approx 10.7185.$$

□

Instead of using CAD for proving that

$$\max_{1 \leq i \leq 3} \max_{\alpha \in D} \nu^{(i)}(\alpha) = \nu^{(1)}(-1/2),$$

we can, alternatively, prove a less sharp bound by estimating the cosine in all formulas of $z^{(i)}$, $i = 1, 2, 3$, by 1. This yields the following bound for all three solutions:

$$\nu_E^{(i)}(\alpha) \leq \sqrt{-\frac{4}{3}p(\alpha)} \cdot 1 - \frac{a(\alpha)}{3} =: \nu_{\max}(\alpha).$$

This bound can be written as

$$\begin{aligned} \nu_{\max}(\alpha) &= -\frac{a(\alpha)}{3} + \sqrt{-\frac{4}{3}p(\alpha)} \\ &= -\frac{1}{8}f(\alpha) + \frac{1}{4}\sqrt{f(\alpha)^2 - 4g(\alpha)} \end{aligned}$$

where

$$f(\alpha) = \frac{8}{3}a(\alpha) \quad \text{and} \quad g(\alpha) = \frac{16}{3}b(\alpha).$$

As can be seen from Figure 5.7, the function $\nu_{\max}(\alpha)$ has its maximum at $\alpha = -1/2$. For the sake of

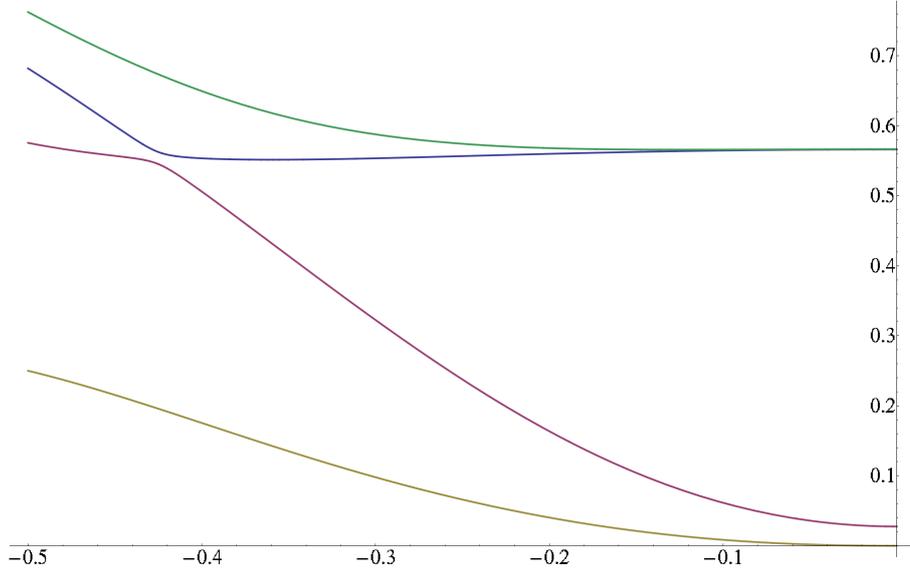


Figure 5.7: $\nu^{(1)}(\alpha)$, $\nu^{(2)}(\alpha)$, $\nu^{(3)}(\alpha)$ and their bound $\nu_{\max}(\alpha)$ for $\alpha \in (-1/2, 0)$.

a simple proof we determine functions $\bar{f}(\alpha)$ and $\underline{g}(\alpha)$ such that we can bound ν_{\max} from above, i.e.,

$$\begin{aligned} \nu_{\max}(\alpha) &= -\frac{1}{8}f(\alpha) + \frac{1}{4}\sqrt{f(\alpha)^2 - 4g(\alpha)} \\ &\leq \frac{1}{8}\bar{f}(\alpha) + \frac{1}{4}\sqrt{\bar{f}(\alpha)^2 - 4\underline{g}(\alpha)} =: \bar{\nu}_{\max}(\alpha). \end{aligned}$$

This can be done by choosing quadratic functions as upper bound for $-f$ and as lower bound for g . We compute three points on the curves $-f$ and g and then fit quadratic polynomials through these points. For instance, we may choose the following two quadratic functions bounding $-f$ and g from above and from below, respectively:

$$\bar{f}(\alpha) = \frac{119}{27}\alpha^2 - \frac{1802}{675}\alpha + \frac{19}{12},$$

$$\underline{g}(\alpha) = \frac{232}{15}\alpha^2 + \frac{9}{25}\alpha + \frac{1}{12}.$$

For this choice of \bar{f} and \underline{g} , the function $\bar{\nu}_{\max}(\alpha)$ does not have a local extremum and, hence, its maximum is attained for $\alpha = -1/2$. In this case, we have that

$$\bar{\nu}_{\max}(-1/2) = \nu_{\max}(-1/2) = \frac{201}{400} + \frac{\sqrt{2701}}{200} \approx 0.762356,$$

and $P(\mu_E, \alpha, \beta)$ is maximal and minimal for $\alpha = -1/2$ and $\beta = 1$. Together with the macro-element-by-macro-element assembling (5.27), we obtain the following condition number estimate:

$$\kappa((C_{11}^K)^{-1}K_{11}) \leq \frac{1 + \sqrt{\bar{\nu}_{\max}(-1/2)}}{1 - \sqrt{\bar{\nu}_{\max}(-1/2)}} \approx 14.7641.$$

5.3.3 Additive preconditioning for the pivot block of the whole system

Theorem 5.9 (Kraus and Wolfmayr [103]). *The additive preconditioner (5.22) of A_{11} with the additive preconditioners for the pivot block of the mass matrix (5.24) and for the stiffness matrix (5.27), where we have used the preconditioners (5.23) and (5.28) corresponding to the macro elements, is uniform. More precisely, the relative condition number bound*

$$\kappa((C_{11})^{-1}A_{11}) \leq \frac{1 + \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}}{1 - \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}} \approx 10.7185 \quad (5.33)$$

holds independently of the shape, the size of each element and of the problem parameters ν and μ , cf. (5.3).

Proof. It follows from macro-element-by-macro-element assembling that

$$\begin{aligned} v_1^T A_{11} v_1 &= \sum_{E \in \mathcal{T}^{(\ell)}} v_{E:1}^T R_E^T A_{E:11} R_E v_{E:1} \\ &= \sum_{E \in \mathcal{T}^{(\ell)}} v_{E:1}^T R_E^T (K_{E:11} + \tilde{\mu}_E M_{E:11}) R_E v_{E:1}. \end{aligned}$$

Using the local eigenvalue estimates (5.26) and (5.32) corresponding to the local generalized eigenproblems with $M_{E:11}$ and $C_{E:11}^M$ and with $K_{E:11}$ and $C_{E:11}^K$, respectively, yields the following upper bound:

$$\begin{aligned} v_1^T A_{11} v_1 &= \sum_{E \in \mathcal{T}^{(\ell)}} v_{E:1}^T R_E^T K_{E:11} R_E v_{E:1} + \sum_{E \in \mathcal{T}^{(\ell)}} v_{E:1}^T R_E^T \tilde{\mu}_E M_{E:11} R_E v_{E:1} \\ &\leq \sum_{E \in \mathcal{T}^{(\ell)}} \lambda_{K,E}^{\max} v_{E:1}^T R_E^T C_{E:11}^K R_E v_{E:1} + \sum_{E \in \mathcal{T}^{(\ell)}} \lambda_{M,E}^{\max} v_{E:1}^T R_E^T \tilde{\mu}_E C_{E:11}^M R_E v_{E:1}. \end{aligned}$$

Taking the maximum of the two eigenvalues $\lambda_{K,E}^{\max}$ and $\lambda_{M,E}^{\max}$, i.e., $\lambda_{A,E}^{\max} = \max\{\lambda_{K,E}^{\max}, \lambda_{M,E}^{\max}\}$, leads to

$$\begin{aligned} v_1^T A_{11} v_1 &= \sum_{E \in \mathcal{T}^{(\ell)}} v_{E:1}^T R_E^T A_{E:11} R_E v_{E:1} \\ &\leq \sum_{E \in \mathcal{T}^{(\ell)}} \lambda_{A,E}^{\max} v_{E:1}^T R_E^T (C_{E:11}^K + \tilde{\mu}_E C_{E:11}^M) R_E v_{E:1} \\ &\leq \lambda_A^{\max} \sum_{E \in \mathcal{T}^{(\ell)}} v_{E:1}^T R_E^T (C_{E:11}^K + \tilde{\mu}_E C_{E:11}^M) R_E v_{E:1} \\ &= \lambda_A^{\max} \sum_{E \in \mathcal{T}^{(\ell)}} v_{E:1}^T R_E^T C_{E:11} R_E v_{E:1} \\ &= \lambda_A^{\max} v_1^T C_{11} v_1, \end{aligned}$$

where $\lambda_A^{\max} = 1 + \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}$. Similarly, one proves

$$v_1^T A_{11} v_1 \geq \lambda_A^{\min} v_1^T C_{11} v_1$$

with $\lambda_A^{\min} = 1 - \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}$. Altogether, the two inequalities yield the condition number estimate (5.33). \square

5.4 Stabilization polynomials of higher degree

As briefly discussed in Subsection 2.8.1, we combine hierarchical basis preconditioners with stabilization techniques in order to obtain a linear AMLI method of optimal order. In this section, we discuss the construction of the matrix polynomials $P^{(\ell)}(t) = P_{v_\ell}(t)$, which occur in the approximations (2.68), i.e.,

$$(Z^{(\ell-1)})^{-1} := \left(I - P^{(\ell)} \left((B^{(\ell-1)})^{-1} A^{(\ell-1)} \right) \right) (A^{(\ell-1)})^{-1}$$

of the Schur complements (2.58), i.e.,

$$S^{(\ell)} = A_{22}^{(\ell)} - A_{21}^{(\ell)} (A_{11}^{(\ell)})^{-1} A_{12}^{(\ell)}$$

at levels $\ell = 1, \dots, L$, not only of degree 1 and 2 but also of higher degree since the considered 3-refinement requires and allows for higher-degree polynomial stabilization in order to fulfill the optimality conditions (2.71), i.e.,

$$\frac{1}{\sqrt{1-\gamma^2}} < v < \varrho.$$

Proper choices of matrix polynomials $P^{(\ell)}(t) = P_{v_\ell}(t)$ approximating matrix inverses are based on *Chebyshev polynomials*, see [14, 15] as well as [8, 101]. This work goes back to Chebyshev (1821-1894), see [45]. Polynomials of best approximation are discussed in, e.g., [102, 106]. The following matrix polynomials based on Chebyshev polynomials can be used on the interval $[\alpha, 1]$ with $0 < \alpha < 1$:

$$P_{v_\ell}(t) = \frac{T_{v_\ell} \left(\frac{1+\alpha-2t}{1-\alpha} \right) + 1}{T_{v_\ell} \left(\frac{1+\alpha}{1-\alpha} \right) + 1}. \quad (5.34)$$

Here, $T_{v_\ell}(t)$ is the Chebyshev polynomial of the first kind defined via the recursion

$$\begin{aligned} T_{v_\ell}(t) &= 2t T_{v_\ell-1}(t) - T_{v_\ell-2}(t), \quad v_\ell = 2, 3, \dots, \\ T_0(t) &= 1, \quad T_1(t) = t. \end{aligned}$$

Next, we define the polynomial $Q^{(\ell)}(t) = Q_{v_\ell-1}(t)$ by

$$\begin{aligned} Q_{v_\ell-1}(t) &= \frac{1 - P_{v_\ell}(t)}{t} \\ &= q_0^{(\ell)} + q_1^{(\ell)} t + \dots + q_{v_\ell-1}^{(\ell)} t^{v_\ell-1}. \end{aligned} \quad (5.35)$$

In Table 5.1, Table 5.2, Table 5.3, Table 5.4 Table 5.5 and Table 5.6, we list the coefficients $q_i^{(\ell)}$, $i \in \{0, 1, \dots, 7\}$, of all polynomials $Q_{v_\ell-1}(t)$ for $v_\ell \in \{1, 2, \dots, 8\}$ on the interval $[\alpha, 1]$ with $0 < \alpha < 1$.

Table 5.1: Coefficients of the polynomials $Q_0(t)$ and $Q_1(t)$.

$q_i^{(\ell)}$	$v_\ell = 1$	$q_i^{(\ell)}$	$v_\ell = 2$
$i = 0$	$q_0^{(1)} = 1$	$i = 0$	$q_0^{(2)} = \frac{4}{1+\alpha}$
		$i = 1$	$q_1^{(2)} = -\frac{4}{(1+\alpha)^2}$

Table 5.2: Coefficients of the polynomials $Q_2(t)$ and $Q_3(t)$.

$q_i^{(\ell)}$	$v_\ell = 3$	$q_i^{(\ell)}$	$v_\ell = 4$
$i = 0$	$q_0^{(3)} = \frac{9+3\alpha}{1+3\alpha}$	$i = 0$	$q_0^{(4)} = \frac{16(1+\alpha)}{1+\alpha(6+\alpha)}$
$i = 1$	$q_1^{(3)} = -\frac{24(1+\alpha)}{(1+3\alpha)^2}$	$i = 1$	$q_1^{(4)} = -\frac{16(5+\alpha(14+5\alpha))}{(1+\alpha(6+\alpha))^2}$
$i = 2$	$q_2^{(3)} = \frac{16}{(1+3\alpha)^2}$	$i = 2$	$q_2^{(4)} = \frac{128(1+\alpha)}{(1+\alpha(6+\alpha))^2}$
		$i = 3$	$q_3^{(4)} = -\frac{64}{(1+\alpha(6+\alpha))^2}$

Table 5.3: Coefficients of the polynomial $Q_4(t)$.

$q_i^{(\ell)}$	$v_\ell = 5$
$i = 0$	$q_0^{(5)} = \frac{5(5+\alpha(10+\alpha))}{1+5\alpha(2+\alpha)}$
$i = 1$	$q_1^{(5)} = -\frac{40(1+\alpha)(5+\alpha(22+5\alpha))}{(1+5\alpha(2+\alpha))^2}$
$i = 2$	$q_2^{(5)} = \frac{80(7+\alpha(18+7\alpha))}{(1+5\alpha(2+\alpha))^2}$
$i = 3$	$q_3^{(5)} = -\frac{640(1+\alpha)}{(1+5\alpha(2+\alpha))^2}$
$i = 4$	$q_4^{(5)} = \frac{256}{(1+5\alpha(2+\alpha))^2}$

Table 5.4: Coefficients of the polynomial $Q_5(t)$.

$q_i^{(\ell)}$	$v_\ell = 6$
$i = 0$	$q_0^{(6)} = \frac{12(3+\alpha)(3\alpha+1)}{(1+\alpha)(1+\alpha(\alpha+14))}$
$i = 1$	$q_1^{(6)} = -\frac{12(35+\alpha(308+\alpha(594+7\alpha(44+5\alpha))))}{(1+\alpha)^2(1+\alpha(\alpha+14))^2}$
$i = 2$	$q_2^{(6)} = \frac{256(\alpha(7\alpha+26)+7)}{(1+\alpha)(1+\alpha(\alpha+14))^2}$
$i = 3$	$q_3^{(6)} = -\frac{384(\alpha(9\alpha+22)+9)}{(1+\alpha)^2(1+\alpha(\alpha+14))^2}$
$i = 4$	$q_4^{(6)} = \frac{3072}{(1+\alpha)(1+\alpha(\alpha+14))^2}$
$i = 5$	$q_5^{(6)} = -\frac{1024}{(1+\alpha)^2(1+\alpha(\alpha+14))^2}$

Table 5.5: Coefficients of the polynomial $Q_6(t)$.

$q_i^{(\ell)}$	$v_\ell = 7$
$i = 0$	$q_0^{(7)} = \frac{7(\alpha(\alpha+21)+35)+7}{1+7\alpha(3+\alpha(5+\alpha))}$
$i = 1$	$q_1^{(7)} = -\frac{112(\alpha+1)(\alpha(\alpha(7\alpha(\alpha+12)+202)+84)+7)}{(1+7\alpha(3+\alpha(5+\alpha)))^2}$
$i = 2$	$q_2^{(7)} = \frac{224(\alpha(\alpha(3\alpha(7\alpha+52)+286)+156)+21)}{(1+7\alpha(3+\alpha(5+\alpha)))^2}$
$i = 3$	$q_3^{(7)} = -\frac{4480(\alpha+1)(\alpha+3)(3\alpha+1)}{(1+7\alpha(3+\alpha(5+\alpha)))^2}$
$i = 4$	$q_4^{(7)} = \frac{1792(\alpha(11\alpha+26)+11)}{(1+7\alpha(3+\alpha(5+\alpha)))^2}$
$i = 5$	$q_5^{(7)} = -\frac{14336(\alpha+1)}{(1+7\alpha(3+\alpha(5+\alpha)))^2}$
$i = 6$	$q_6^{(7)} = \frac{4096}{(1+7\alpha(3+\alpha(5+\alpha)))^2}$

Table 5.6: Coefficients of the polynomial $Q_7(t)$.

$q_i^{(\ell)}$	$v_\ell = 8$
$i = 0$	$q_0^{(8)} = \frac{64(\alpha+1)(\alpha(\alpha+6)+1)}{1+\alpha(28+\alpha(70+\alpha(28+\alpha)))}$
$i = 1$	$q_1^{(8)} = -\frac{64(\alpha(\alpha(\alpha(3\alpha(7\alpha+18)+585)+2860)+1755)+378)+21}{(1+\alpha(28+\alpha(70+\alpha(28+\alpha))))^2}$
$i = 2$	$q_2^{(8)} = \frac{512(\alpha+1)(\alpha(\alpha(3\alpha(7\alpha+68)+446)+204)+21)}{(1+\alpha(28+\alpha(70+\alpha(28+\alpha))))^2}$
$i = 3$	$q_3^{(8)} = -\frac{1280(\alpha(\alpha(11\alpha(3\alpha+20)+390)+220)+33)}{(1+\alpha(28+\alpha(70+\alpha(28+\alpha))))^2}$
$i = 4$	$q_4^{(8)} = \frac{8192(\alpha+1)(\alpha(11\alpha+34)+11)}{(1+\alpha(28+\alpha(70+\alpha(28+\alpha))))^2}$
$i = 5$	$q_5^{(8)} = -\frac{8192(\alpha(13\alpha+30)+13)}{(1+\alpha(28+\alpha(70+\alpha(28+\alpha))))^2}$
$i = 6$	$q_6^{(8)} = \frac{65536(\alpha+1)}{(1+\alpha(28+\alpha(70+\alpha(28+\alpha))))^2}$
$i = 7$	$q_7^{(8)} = -\frac{16384}{(1+\alpha(28+\alpha(70+\alpha(28+\alpha))))^2}$

Remark 5.10. *In order to ensure that the linear AMLI preconditioner is uniform and of optimal order of computational complexity, the stabilization polynomial $P_{v_\ell}(t)$ for the 3-refinement has to be at least of degree $v_\ell = 5$, which corresponds to a polynomial $Q_{v_\ell-1}(t)$ of degree 4. More precisely, the optimality condition (5.15) in the 3-refinement case, i.e.,*

$$\frac{1}{\sqrt{1-20/21}} = \sqrt{21} \approx 4.58258 < v < \varrho = 9 \quad \text{with } v = \max_{0 \leq \ell \leq L} v_\ell$$

is fulfilled for polynomial degree $v \in \{5, 6, 7, 8\}$.

5.5 The relative condition number of the AMLI preconditioned system

We want to consider some bounds for the spectral condition number

$$\kappa((B^{(\ell)})^{-1}A^{(\ell)})$$

of the AMLI preconditioned linear systems, which consist of weighted sums of stiffness and mass matrices, in the spirit of [11, 14, 15]. The preconditioner $B^{(\ell)}$ at level ℓ for (2.66) and replacing the pivot block $A_{11}^{(\ell)}$ by an approximation $C_{11}^{(\ell)}$ satisfying (2.70), i.e.,

$$\underline{c} v_1^T C_{11}^{(\ell)} v_1 \leq v_1^T A_{11}^{(\ell)} v_1 \leq \bar{c} v_1^T C_{11}^{(\ell)} v_1,$$

yields the linear AMLI preconditioner (2.73), which is given by

$$B^{(\ell)} := \begin{pmatrix} C_{11}^{(\ell)} & 0 \\ \tilde{A}_{21}^{(\ell)} & Z^{(\ell-1)} \end{pmatrix} \begin{pmatrix} I & (C_{11}^{(\ell)})^{-1} \tilde{A}_{12}^{(\ell)} \\ 0 & I \end{pmatrix}.$$

Here, we denote by $Z^{(\ell-1)}$ the approximation of the inverse of the Schur complements (2.68), i.e.,

$$(Z^{(\ell-1)})^{-1} := \left(I - P^{(\ell)} \left((B^{(\ell-1)})^{-1} A^{(\ell-1)} \right) \right) (A^{(\ell-1)})^{-1}$$

Since we have discussed the choice of the stabilization polynomials in Section 5.4, we now want to present bounds for the relative condition number of $A^{(\ell)}$ with respect to $B^{(\ell)}$, which are depending on the CBS constant γ , the degree of the stabilization polynomial v , the interval on which the stabilization polynomial is defined and on the constants \underline{c} and \bar{c} for approximating the pivot blocks according to (2.70). The following estimates are proven for stabilization polynomials defined on the interval $[\alpha, 1]$ with $0 < \alpha < 1$, see [15].

Theorem 5.11. *Let $C_{11}^{(\ell)}$ be a spectral equivalent preconditioner for the pivot block $A_{11}^{(\ell)}$ satisfying (2.70) and let $B^{(\ell)}$ be the linear AMLI preconditioner defined in (2.73). Then the relative condition number of $A^{(\ell)}$ with respect to $B^{(\ell)}$ is bounded by*

$$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq \frac{1}{1-\gamma^2} \left(c\gamma^2 + \left(\frac{1+\vartheta(\alpha)}{1-\vartheta(\alpha)} \right)^2 \right) + \frac{c^2}{c + \gamma^2 + \frac{4(1-\alpha)^v}{(1-\vartheta(\alpha))^2}}, \quad (5.36)$$

where $c = \bar{c}/\underline{c} - 1$ and $\vartheta(\alpha)$ denotes the so-called (asymptotic) reduction rate of the Chebyshev method given by

$$\vartheta(\alpha) = \left(\frac{1-\sqrt{\alpha}}{1+\sqrt{\alpha}} \right)^v. \quad (5.37)$$

Proof. See [15]. □

The parameter α can be chosen by solving the equation

$$1 = \frac{1}{1-\gamma^2} \left(\alpha c \gamma^2 + \left(\frac{1+\vartheta(\alpha)}{2 \sum_{s=1}^v (1+\sqrt{\alpha})^{v-s} (1-\sqrt{\alpha})^{s-1}} \right)^2 \right) + \frac{\alpha c^2}{c + \gamma^2 + \frac{4(1-\alpha)^v}{(1-\vartheta(\alpha))^2}},$$

see [15], which is equivalent to solving the equation

$$1 = \gamma^2 + \alpha \left(c \gamma^2 + \left(\frac{1+\vartheta(\alpha)}{1-\vartheta(\alpha)} \right)^2 + \frac{c^2 (1-\gamma^2)}{c + \gamma^2 + \frac{4(1-\alpha)^v}{(1-\vartheta(\alpha))^2}} \right). \quad (5.38)$$

Remark 5.12. We can simplify the estimate (5.36) by obtaining the (only slightly worse) upper bound

$$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq \frac{1}{1-\gamma^2} \left(c + \left(\frac{1+\vartheta(\alpha)}{1-\vartheta(\alpha)} \right)^2 \right), \quad (5.39)$$

see [15]. Then, we can choose the parameter α by solving the equation

$$1 - \gamma^2 = \alpha c + \left(\frac{1+\vartheta(\alpha)}{2 \sum_{s=1}^v (1+\sqrt{\alpha})^{v-s} (1-\sqrt{\alpha})^{s-1}} \right)^2,$$

or, equivalently,

$$1 - \gamma^2 = \alpha \left(c + \left(\frac{1+\vartheta(\alpha)}{1-\vartheta(\alpha)} \right)^2 \right).$$

Using the linear AMLI preconditioner (2.67) with the exact pivot block $A_{11}^{(\ell)}$ reduces the estimate (5.36) to

$$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq \frac{1}{1-\gamma^2} \left(\frac{1+\vartheta(\alpha)}{1-\vartheta(\alpha)} \right)^2, \quad (5.40)$$

since we have $\bar{c}/\underline{c} = 1$, hence, $c = 0$, in this case, see [14, 15]. Here, the parameter α can be chosen by solving the equation

$$1 - \gamma^2 = \alpha \left(\frac{1+\vartheta(\alpha)}{1-\vartheta(\alpha)} \right)^2. \quad (5.41)$$

Let us consider the 3-refinement case and the discretized problem (5.3), i.e.,

$$\underbrace{(K_{\nu,h} + M_{\mu,h})}_{=:A_h} \underline{u}_h = \underline{f}_h.$$

As first example, we consider the case without mass matrix. Then, the linear system is given by

$$K_{\nu,h} \underline{u}_h = \underline{f}_h,$$

the CBS constant is estimated by $\gamma \leq \sqrt{8/9}$, and the optimality conditions

$$\frac{1}{\sqrt{1-8/9}} = 3 < \nu < \varrho = m^2 = 9$$

are fulfilled for polynomials of degree $\nu \in \{4, 5, 6, 7, 8\}$. By using the exact pivot block $A_{11}^{(\ell)}$, we can compute α from solving equation (5.41). This yields the following choice for α :

$$\alpha = \frac{1}{3} \left(2\sqrt{13} - 7 \right) \approx 0.0704.$$

Table 5.7: Condition number estimates for the AMLI method in the 3-refinement case using additive preconditioners for approximating the pivot blocks and stabilization polynomials of degree $v_\ell = 5$.

α	$\kappa((B^{(\ell)})^{-1}A^{(\ell)})$
$\alpha = 0.001$	$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq 1049.72$
$\alpha = 0.01$	$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq 298.741$
$\alpha = 0.1$	$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq 227.521$
$\alpha = 0.5$	$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq 224.233$
$\alpha = 0.9$	$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq 224.221$

We can compute an estimate for the condition number by (5.40) and obtain

$$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq 7 + 2\sqrt{13} \approx 14.2111.$$

Now, let us consider the problem (5.3) with mass matrix. In this case, the CBS constant is estimated by $\gamma \leq \sqrt{20/21}$, see (5.14), and the optimality conditions (5.15)

$$\frac{1}{\sqrt{1-20/21}} = \sqrt{21} \approx 4.58258 < v < \varrho = m^2 = 9$$

are fulfilled for polynomials of degree $v \in \{5, 6, 7, 8\}$. By using the exact pivot block $A_{11}^{(\ell)}$, we can compute α from solving equation (5.41). This yields the choice

$$\alpha = \frac{1}{131} \left(-130 + 5\sqrt{21} + \sqrt{2(7075 - 257\sqrt{21})} \right) \approx 0.0116$$

and an estimate for the condition number by (5.40), i.e.,

$$\kappa((B^{(\ell)})^{-1}A^{(\ell)}) \leq 5 \left(4 + \sqrt{21} \right) + \sqrt{950 + 206\sqrt{21}} \approx 86.4331.$$

Finally, let us consider solving the finite element problem (5.3) by the AMLI preconditioned CG method together with approximating the pivot block $A_{11}^{(\ell)}$ by the additive preconditioner $C_{11}^{(\ell)}$ presented in Section 5.3. In this case,

$$c = \frac{1 + \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}}{1 - \sqrt{\frac{223}{640} + \frac{3\sqrt{5241}}{640}}} - 1 \approx 9.7185.$$

In Table 5.7, we present some estimates for the condition numbers by computing (5.36) for different choices of α .

In conclusion of this section, the parameter α should be chosen carefully, especially for different choices of polynomial degrees v . For further information regarding the choices of α and condition number estimates for the AMLI preconditioned CG method, we refer the reader to [8, 11, 12, 14, 15] as well as [101] and the references therein.

5.6 AMLI preconditioned MINRES solver for parabolic time-periodic problems

In this section, we briefly present the model problems from Chapter 3 and Chapter 4 for the case $\Omega \subset \mathbb{R}^2$ being a (two-dimensional) bounded (polygonal) Lipschitz domain and discuss the use of

the AMLI preconditioner in order to solve the problems by the preconditioned MINRES method. Moreover, we assume that the coefficients σ and ν are piecewise constant on the coarsest mesh.

Let us start with the parabolic time-periodic partial differential equation (3.1)-(3.3). After the multiharmonic finite element discretization of a corresponding variational problem, we finally arrived at saddle point systems (3.26) corresponding to every mode $k = 1, \dots, N$, i.e.,

$$\begin{pmatrix} k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} \end{pmatrix} \begin{pmatrix} \underline{u}_k^s \\ \underline{u}_k^c \end{pmatrix} = \begin{pmatrix} -\underline{f}_k^c \\ -\underline{f}_k^s \end{pmatrix},$$

for which we constructed the block-diagonal preconditioner (3.30)

$$\mathcal{P} = \begin{pmatrix} k\omega M_{h,\sigma} + K_{h,\nu} & 0 \\ 0 & k\omega M_{h,\sigma} + K_{h,\nu} \end{pmatrix},$$

for a MINRES solver yielding the robust condition number estimate $\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) \leq \sqrt{2}$. For $k = 0$, we had to solve the linear system

$$K_{h,\nu} \underline{u}_0^c = \underline{f}_0^c.$$

Now, we are able to precondition the diagonal blocks $D = (k\omega M_{h,\sigma} + K_{h,\nu})$ of the preconditioner \mathcal{P} in (3.30) of the discretized problem (3.26) for $k = 1, 2, \dots, N$ by the AMLI method. More precisely, we can replace these diagonal blocks D by spectral equivalent ones \tilde{D} , i.e., $\underline{c}_D \tilde{D} \leq D \leq \bar{c}_D \tilde{D}$, with a condition number estimate

$$\kappa_{\tilde{\mathcal{P}}}(\tilde{\mathcal{P}}^{-1}\mathcal{A}) \leq \kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) (\bar{c}_D/\underline{c}_D).$$

The corresponding optimal control problem of Chapter 4, i.e.,

$$\min_{y,u} \mathcal{J}(y,u) := \frac{1}{2} \int_0^T \int_{\Omega} [y(\mathbf{x},t) - y_d(\mathbf{x},t)]^2 d\mathbf{x} dt + \frac{\lambda}{2} \int_0^T \int_{\Omega} [u(\mathbf{x},t)]^2 d\mathbf{x} dt$$

subject to the partial differential equation (3.1)-(3.3), is treated analogously: The multiharmonic finite element discretization of its variational problem leads to the linear saddle point systems (4.11) and (4.12), i.e.,

$$\begin{pmatrix} M_h & 0 & -K_{h,\nu} & k\omega M_{h,\sigma} \\ 0 & M_h & -k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} & -\lambda^{-1}M_h & 0 \\ k\omega M_{h,\sigma} & -K_{h,\nu} & 0 & -\lambda^{-1}M_h \end{pmatrix} \begin{pmatrix} \underline{y}_k^c \\ \underline{y}_k^s \\ \underline{p}_k^c \\ \underline{p}_k^s \end{pmatrix} = \begin{pmatrix} \underline{y}_{dk}^c \\ \underline{y}_{dk}^s \\ 0 \\ 0 \end{pmatrix}$$

for $k = 1, 2, \dots, N$, and

$$\begin{pmatrix} M_h & -K_{h,\nu} \\ -K_{h,\nu} & -\lambda^{-1}M_h \end{pmatrix} \begin{pmatrix} \underline{y}_0^c \\ \underline{p}_0^c \end{pmatrix} = \begin{pmatrix} \underline{y}_{d0}^c \\ 0 \end{pmatrix}$$

for $k = 0$, respectively. For the cases $k = 1, 2, \dots, N$, we constructed the preconditioners (4.22), i.e.,

$$\mathcal{P} = \begin{pmatrix} D & 0 & 0 & 0 \\ 0 & D & 0 & 0 \\ 0 & 0 & \lambda^{-1}D & 0 \\ 0 & 0 & 0 & \lambda^{-1}D \end{pmatrix}$$

with $D = \sqrt{\lambda}K_{h,\nu} + k\omega\sqrt{\lambda}M_{h,\sigma} + M_h$ and proved the condition number estimate

$$\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) := \|\mathcal{P}^{-1}\mathcal{A}\|_{\mathcal{P}} \|\mathcal{A}^{-1}\mathcal{P}\|_{\mathcal{P}} \leq \sqrt{3}.$$

In the case $k = 0$, we constructed the preconditioner (4.31), i.e.,

$$\mathcal{P} = \begin{pmatrix} D & 0 \\ 0 & \lambda^{-1}D \end{pmatrix}$$

with $D = M_h + \sqrt{\lambda}K_{h,\nu}$ implying the condition number estimate

$$\kappa_{\mathcal{P}}(\mathcal{P}^{-1}\mathcal{A}) := \|\mathcal{P}^{-1}\mathcal{A}\|_{\mathcal{P}} \|\mathcal{A}^{-1}\mathcal{P}\|_{\mathcal{P}} \leq \sqrt{2}.$$

If all parameters are piecewise constant on the coarsest mesh, the AMLI preconditioner discussed in this chapter yields a robust and optimal preconditioner for D and, hence, an optimal MINRES solver either for the partial differential equation (3.1)-(3.3) or for the optimal control problem (4.1)-(4.2). Finally, we will present numerical results confirming our theoretical findings and demonstrating the robustness and optimal complexity of the AMLI preconditioned MINRES solver for parabolic time-periodic problems in Chapter 7.

Chapter 6

A posteriori error analysis of parabolic time-periodic problems

This chapter is devoted to the a posteriori error analysis of parabolic time-periodic boundary value and optimal control problems that are presented in Chapters 3 and 4, respectively. The functional a posteriori error estimation techniques, which we use, are based on the works by Repin, see, e.g., [152, 61, 59, 60, 154, 153, 123] and the references therein. In particular, our a posteriori error analysis is based on the method presented in [152], but the analysis contains proper changes regarding the space $H^{1, \frac{1}{2}}$ and the special features of the approximation via truncated Fourier series.

6.1 Functional a posteriori error estimates for parabolic time-periodic boundary value problems

As starting point let us consider the variational problem (3.16): Given $f \in L^2(Q_T)$, find $u \in H_0^{1, \frac{1}{2}}(Q_T)$ such that

$$\int_0^T \int_{\Omega} \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} v^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla v \right) d\mathbf{x} dt = \int_0^T \int_{\Omega} f v d\mathbf{x} dt \quad (6.1)$$

for all test functions $v \in H_0^{1, \frac{1}{2}}(Q_T)$, where all functions are given in their Fourier series expansion in time, i.e., everything has to be understood in the sense of Definition 3.2. Moreover, we have defined the space-time bilinear form $a(\cdot, \cdot)$ in (3.17) as follows

$$a(u, v) = \int_0^T \int_{\Omega} \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} v^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla v \right) d\mathbf{x} dt.$$

Let η be an approximation to the exact solution u , e.g., the multiharmonic finite element approximation u_{Nh} constructed in Chapter 3.

A first a posteriori error result

First, we assume that η is a bit more regular than u , i.e., $\eta \in H_{0, per}^{1,1}(Q_T)$, that is of course true for the multiharmonic finite element approximation u_{Nh} . Our goal is to deduce a computable upper bound of the error

$$e := u - \eta$$

in $H_0^{1, \frac{1}{2}}(Q_T)$.

Relation (6.1) immediately implies that the integral identity

$$\begin{aligned} & \int_0^T \int_{\Omega} \left(\sigma(\mathbf{x}) \partial_t^{1/2} (u - \eta) \partial_t^{1/2} v^\perp + \nu(\mathbf{x}) \nabla (u - \eta) \cdot \nabla v \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(f v - \sigma(\mathbf{x}) \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp - \nu(\mathbf{x}) \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \end{aligned} \quad (6.2)$$

is valid for all $v \in H_0^{1, \frac{1}{2}}(Q_T)$. The left hand side of (6.2) is nothing but

$$a(u - \eta, v),$$

and we know from Lemma 3.5 that

$$\sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u - \eta, v)}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}} \geq \mu_1 \|u - \eta\|_{H^{1, \frac{1}{2}}(Q_T)} \quad (6.3)$$

with the positive constant $\mu_1 = \min\{\frac{\underline{\nu}}{C_F^2 + 1}, \underline{\sigma}\}$. Moreover, we can similarly prove inf-sup and sup-sup conditions with the $H^{1, \frac{1}{2}}$ -seminorm. In fact, the $H^{1, \frac{1}{2}}$ -seminorm is, due to the Friedrichs inequality, an equivalent norm.

Lemma 6.1. *The space-time bilinear form $a(\cdot, \cdot)$ defined by (3.17) fulfills the following inf-sup and sup-sup conditions:*

$$\mu_1 |u|_{H^{1, \frac{1}{2}}(Q_T)} \leq \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u, v)}{|v|_{H^{1, \frac{1}{2}}(Q_T)}} \leq \mu_2 |u|_{H^{1, \frac{1}{2}}(Q_T)} \quad (6.4)$$

for all $u \in H_0^{1, \frac{1}{2}}(Q_T)$ with positive constants $\mu_1 = \min\{\underline{\nu}, \underline{\sigma}\}$ and $\mu_2 = \max\{\bar{\sigma}, \bar{\nu}\}$.

Proof. The inf-sup and sup-sup conditions are analogously proven as in Lemma 3.5. \square

We denote the right-hand side of (6.2) by $\mathcal{F}_\eta(v)$. Indeed,

$$\mathcal{F}_\eta(v) = \int_0^T \int_{\Omega} \left(f v - \sigma(\mathbf{x}) \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp - \nu(\mathbf{x}) \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt$$

is a linear functional defined on $v \in H_0^{1, \frac{1}{2}}(Q_T)$. We need to find an upper bound of

$$\sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}} \quad \text{or} \quad \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{|v|_{H^{1, \frac{1}{2}}(Q_T)}}. \quad (6.5)$$

For getting such a bound we need to reconstruct $\mathcal{F}_\eta(v)$. First, we note that the σ -weighted identity

$$\left(\sigma \partial_t^{1/2} \eta, \partial_t^{1/2} v^\perp \right)_{L^2(Q_T)} = \left(\sigma \partial_t \eta, v \right)_{L^2(Q_T)} \quad (6.6)$$

is valid since $\eta \in H_{0, per}^{1, 1}(Q_T)$ and $v \in H_0^{1, \frac{1}{2}}(Q_T)$, see also Lemma 3.3 and Remark 3.4. Second, we introduce a vector-valued function

$$\boldsymbol{\tau} \in H(\operatorname{div}, Q_T) := \{ \boldsymbol{\tau} \in [L^2(Q_T)]^d : \operatorname{div} \boldsymbol{\tau}(\cdot, t) \in L^2(\Omega) \text{ for a.e. } t \in (0, T) \}.$$

Here, $\operatorname{div} = \operatorname{div}_{\mathbf{x}}$ denotes the weak spatial divergence defined by the identity

$$\int_{\Omega} \operatorname{div} \boldsymbol{\tau} v d\mathbf{x} = - \int_{\Omega} \boldsymbol{\tau} \cdot \nabla v d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega).$$

Due to the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathcal{F}_\eta(v) &= \int_0^T \int_\Omega \left(f v - \sigma(\mathbf{x}) \partial_t \eta v + \operatorname{div} \boldsymbol{\tau} v + (\boldsymbol{\tau} - \nu(\mathbf{x}) \nabla \eta) \cdot \nabla v \right) d\mathbf{x} dt \\ &\leq \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \end{aligned} \quad (6.7)$$

with

$$\mathcal{R}_1(\eta, \boldsymbol{\tau}) = \sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau} - f \quad \text{and} \quad \mathcal{R}_2(\eta, \boldsymbol{\tau}) = \boldsymbol{\tau} - \nu \nabla \eta.$$

Applying the Friedrichs inequality (2.17), or rather the Friedrichs inequality (3.19) in the space-time cylinder Q_T , yields

$$\begin{aligned} \mathcal{F}_\eta(v) &\leq \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\leq \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} C_F \|\nabla v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\leq (C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}) \|\nabla v\|_{L^2(Q_T)}, \end{aligned}$$

where again $\nabla = \nabla_{\mathbf{x}}$ denotes the weak spatial gradient. Hence,

$$\begin{aligned} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{|v|_{H^{1, \frac{1}{2}}(Q_T)}} &\leq \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{(C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}) \|\nabla v\|_{L^2(Q_T)}}{|v|_{H^{1, \frac{1}{2}}(Q_T)}} \\ &= \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{(C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}) \|\nabla v\|_{L^2(Q_T)}}{(\|\nabla v\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2)^{1/2}} \\ &\leq C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}. \end{aligned} \quad (6.8)$$

We arrive at the following result:

Theorem 6.2. *Let $\eta \in H_{0, \text{per}}^{1,1}(Q_T)$ and the bilinear form $a(\cdot, \cdot)$ satisfy (6.4). Then,*

$$|u - \eta|_{H^{1, \frac{1}{2}}(Q_T)} \leq \frac{1}{\mu_1} (C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}) =: \mathcal{M}_{|\cdot|}^\oplus(\eta, \boldsymbol{\tau}), \quad (6.9)$$

where $\mu_1 = \min\{\underline{\nu}, \underline{\sigma}\}$ and $\boldsymbol{\tau} \in H(\operatorname{div}, Q_T)$.

Proof. Using the left inequality of (6.4) we get the estimate

$$|u - \eta|_{H^{1, \frac{1}{2}}(Q_T)} \leq \frac{1}{\mu_1} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u - \eta, v)}{|v|_{H^{1, \frac{1}{2}}(Q_T)}} = \frac{1}{\mu_1} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{|v|_{H^{1, \frac{1}{2}}(Q_T)}}$$

that together with (6.8) immediately lead to (6.9). \square

A similar estimate as (6.9) for the seminorm can be proven for the full norm using the inf-sup condition (6.3).

Theorem 6.3. *Let $\eta \in H_{0, \text{per}}^{1,1}(Q_T)$ and the bilinear form $a(\cdot, \cdot)$ satisfy (6.3). Then,*

$$\|u - \eta\|_{H^{1, \frac{1}{2}}(Q_T)} \leq \frac{1}{\mu_1} \left(\|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right)^{1/2} =: \mathcal{M}_{\|\cdot\|}^\oplus(\eta, \boldsymbol{\tau}), \quad (6.10)$$

where $\boldsymbol{\tau} \in H(\operatorname{div}, Q_T)$ and now $\mu_1 = \min\{\frac{\underline{\nu}}{C_F^2 + 1}, \underline{\sigma}\}$.

Proof. Applying the Cauchy-Schwarz inequality again on (6.7), we obtain

$$\begin{aligned} \mathcal{F}_\eta(v) &\leq \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\leq \left(\|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right)^{1/2} \left(\|v\|_{L^2(Q_T)}^2 + \|\nabla v\|_{L^2(Q_T)}^2 \right)^{1/2}. \end{aligned}$$

Hence, estimate (6.10) now follows from (6.3) and the estimate

$$\begin{aligned} &\sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}} \\ &\leq \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\left(\|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right)^{1/2} \left(\|v\|_{L^2(Q_T)}^2 + \|\nabla v\|_{L^2(Q_T)}^2 \right)^{1/2}}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}} \\ &\leq \left(\|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right)^{1/2}. \end{aligned}$$

□

We call the functions $\mathcal{M}_{|\cdot|}^\oplus(\eta, \boldsymbol{\tau})$ and $\mathcal{M}_{\|\cdot\|}^\oplus(\eta, \boldsymbol{\tau})$ error majorants. They denote the upper bounds for the error in the $H^{1, \frac{1}{2}}$ -seminorm and full norm, respectively.

Remark 6.4. If $\mathcal{R}_1(\eta, \boldsymbol{\tau}) = 0$ and $\mathcal{R}_2(\eta, \boldsymbol{\tau}) = 0$, then

$$\begin{aligned} \sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau} &= f, \\ \boldsymbol{\tau} &= \nu \nabla \eta. \end{aligned}$$

Since $\eta \in H_{0, \text{per}}^{1,1}(Q_T)$ is a periodic function and satisfies the Dirichlet condition on Σ_T , it is the solution. In other words, the majorants vanish if and only if η is the exact solution and $\boldsymbol{\tau}$ is the exact flux.

The multiharmonic approximation

Since we assume that the function f is from $L^2(Q_T)$, we expand it into a Fourier series. Moreover, we choose our approximation η of the solution u as well as the vector-valued function $\boldsymbol{\tau}$ to be some truncated Fourier series, i.e.,

$$\begin{aligned} \eta(\mathbf{x}, t) &= \eta_0^c(\mathbf{x}) + \sum_{k=1}^N [\eta_k^c(\mathbf{x}) \cos(k\omega t) + \eta_k^s(\mathbf{x}) \sin(k\omega t)], \\ \boldsymbol{\tau}(\mathbf{x}, t) &= \boldsymbol{\tau}_0^c(\mathbf{x}) + \sum_{k=1}^N [\boldsymbol{\tau}_k^c(\mathbf{x}) \cos(k\omega t) + \boldsymbol{\tau}_k^s(\mathbf{x}) \sin(k\omega t)], \end{aligned} \tag{6.11}$$

where all Fourier coefficients are from the space $L^2(\Omega)$ and are defined by the relations

$$\begin{aligned} \eta_0^c(\mathbf{x}) &= \frac{1}{T} \int_0^T \eta(\mathbf{x}, t) dt, & \boldsymbol{\tau}_0^c(\mathbf{x}) &= \frac{1}{T} \int_0^T \boldsymbol{\tau}(\mathbf{x}, t) dt, \\ \eta_k^c(\mathbf{x}) &= \frac{2}{T} \int_0^T \eta(\mathbf{x}, t) \cos(k\omega t) dt, & \boldsymbol{\tau}_k^c(\mathbf{x}) &= \frac{2}{T} \int_0^T \boldsymbol{\tau}(\mathbf{x}, t) \cos(k\omega t) dt, \\ \eta_k^s(\mathbf{x}) &= \frac{2}{T} \int_0^T \eta(\mathbf{x}, t) \sin(k\omega t) dt, & \boldsymbol{\tau}_k^s(\mathbf{x}) &= \frac{2}{T} \int_0^T \boldsymbol{\tau}(\mathbf{x}, t) \sin(k\omega t) dt. \end{aligned}$$

Hence, we can compute the $L^2(Q_T)$ -norms of the functions

$$\mathcal{R}_1(\eta, \boldsymbol{\tau}) = \sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau} - f \quad \text{and} \quad \mathcal{R}_2(\eta, \boldsymbol{\tau}) = \boldsymbol{\tau} - \nu \nabla \eta,$$

where

$$\begin{aligned} \partial_t \eta(\mathbf{x}, t) &= \sum_{k=1}^N [k\omega \eta_k^s(\mathbf{x}) \cos(k\omega t) - k\omega \eta_k^c(\mathbf{x}) \sin(k\omega t)], \\ \nabla \eta(\mathbf{x}, t) &= \nabla \eta_0^c(\mathbf{x}) + \sum_{k=1}^N [\nabla \eta_k^c(\mathbf{x}) \cos(k\omega t) + \nabla \eta_k^s(\mathbf{x}) \sin(k\omega t)], \\ \operatorname{div} \boldsymbol{\tau}(\mathbf{x}, t) &= \operatorname{div} \boldsymbol{\tau}_0^c(\mathbf{x}) + \sum_{k=1}^N [\operatorname{div} \boldsymbol{\tau}_k^c(\mathbf{x}) \cos(k\omega t) + \operatorname{div} \boldsymbol{\tau}_k^s(\mathbf{x}) \sin(k\omega t)]. \end{aligned}$$

We have that

$$\begin{aligned} \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 &= \int_0^T \int_{\Omega} (\sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau} - f)^2 \, d\mathbf{x} \, dt \\ &= \int_0^T \int_{\Omega} \left((\sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau})^2 - 2 (\sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau}) f + f^2 \right) \, d\mathbf{x} \, dt. \end{aligned}$$

We want to mention that we split only this first integral into three parts, since f does not have a multiharmonic representation, and we want to compute the first integral more in detail. Due to the orthogonalities of the cosine and sine functions with respect to the $L^2(0, T)$ -inner product, i.e., orthogonalities (2.9), the integrals in time can be computed easily. For instance, we have that

$$\begin{aligned} &\int_0^T \int_{\Omega} (\sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau})^2 \, d\mathbf{x} \, dt \\ &= \int_0^T \int_{\Omega} \left(\sum_{k=1}^N [k\omega \sigma(\mathbf{x}) \eta_k^s(\mathbf{x}) \cos(k\omega t) - k\omega \sigma(\mathbf{x}) \eta_k^c(\mathbf{x}) \sin(k\omega t)] \right. \\ &\quad \left. - \operatorname{div} \boldsymbol{\tau}_0^c(\mathbf{x}) - \sum_{k=1}^N [\operatorname{div} \boldsymbol{\tau}_k^c(\mathbf{x}) \cos(k\omega t) + \operatorname{div} \boldsymbol{\tau}_k^s(\mathbf{x}) \sin(k\omega t)] \right)^2 \, d\mathbf{x} \, dt \\ &= \int_0^T \int_{\Omega} \left(-\operatorname{div} \boldsymbol{\tau}_0^c(\mathbf{x}) + \sum_{k=1}^N [(k\omega \sigma(\mathbf{x}) \eta_k^s(\mathbf{x}) - \operatorname{div} \boldsymbol{\tau}_k^c(\mathbf{x})) \cos(k\omega t) \right. \\ &\quad \left. + (-k\omega \sigma(\mathbf{x}) \eta_k^c(\mathbf{x}) - \operatorname{div} \boldsymbol{\tau}_k^s(\mathbf{x})) \sin(k\omega t)] \right)^2 \, d\mathbf{x} \, dt \\ &= T \int_{\Omega} (\operatorname{div} \boldsymbol{\tau}_0^c(\mathbf{x}))^2 \, d\mathbf{x} + \frac{T}{2} \int_{\Omega} \sum_{k=1}^N [(k\omega \sigma(\mathbf{x}) \eta_k^s(\mathbf{x}) - \operatorname{div} \boldsymbol{\tau}_k^c(\mathbf{x}))^2 \\ &\quad + (-k\omega \sigma(\mathbf{x}) \eta_k^c(\mathbf{x}) - \operatorname{div} \boldsymbol{\tau}_k^s(\mathbf{x}))^2] \, d\mathbf{x} \\ &= T \|\operatorname{div} \boldsymbol{\tau}_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N [\|k\omega \sigma \eta_k^s - \operatorname{div} \boldsymbol{\tau}_k^c\|_{L^2(\Omega)}^2 + \|-k\omega \sigma \eta_k^c - \operatorname{div} \boldsymbol{\tau}_k^s\|_{L^2(\Omega)}^2] \\ &= T \|\operatorname{div} \boldsymbol{\tau}_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N [\|-k\omega \sigma \eta_k^s + \operatorname{div} \boldsymbol{\tau}_k^c\|_{L^2(\Omega)}^2 + \|k\omega \sigma \eta_k^c + \operatorname{div} \boldsymbol{\tau}_k^s\|_{L^2(\Omega)}^2] \\ &= T \|\operatorname{div} \boldsymbol{\tau}_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|k\omega \sigma \eta_k^\perp + \operatorname{div} \boldsymbol{\tau}_k\|_{L^2(\Omega)}^2, \end{aligned}$$

where $\boldsymbol{\eta}_k^\perp = (-\eta_k^s, \eta_k^c)^T$ and $\mathbf{div} \boldsymbol{\tau}_k = (\mathbf{div} \boldsymbol{\tau}_k^c, \mathbf{div} \boldsymbol{\tau}_k^s)^T$. We remind the reader that

$$\|\mathbf{u}_k\|_{L^2(\Omega)}^2 = \|u_k^c\|_{L^2(\Omega)}^2 + \|u_k^s\|_{L^2(\Omega)}^2.$$

Analogously, we compute the following time-integrals:

$$\begin{aligned} -2 \int_0^T \int_{\Omega} (\sigma \partial_t \eta - \mathbf{div} \boldsymbol{\tau}) f \, d\mathbf{x} \, dt &= -2 \left(T(-\mathbf{div} \boldsymbol{\tau}_0^c, f_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^N (-k\omega \sigma \boldsymbol{\eta}_k^\perp - \mathbf{div} \boldsymbol{\tau}_k, \mathbf{f}_k)_{L^2(\Omega)} \right) \\ &= 2 \left(T(\mathbf{div} \boldsymbol{\tau}_0^c, f_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^N (k\omega \sigma \boldsymbol{\eta}_k^\perp + \mathbf{div} \boldsymbol{\tau}_k, \mathbf{f}_k)_{L^2(\Omega)} \right) \end{aligned}$$

and

$$\int_0^T \int_{\Omega} f^2 \, d\mathbf{x} \, dt = T \|f_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{f}_k\|_{L^2(\Omega)}^2.$$

Altogether, we obtain the L^2 -norm of \mathcal{R}_1 , i.e.,

$$\begin{aligned} \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 &= \|\sigma \partial_t \eta - \mathbf{div} \boldsymbol{\tau}\|_{L^2(Q_T)}^2 - 2(\sigma \partial_t \eta - \mathbf{div} \boldsymbol{\tau}, f)_{L^2(Q_T)} + \|f\|_{L^2(Q_T)}^2 \\ &= \int_0^T \int_{\Omega} (\sigma \partial_t \eta - \mathbf{div} \boldsymbol{\tau})^2 \, d\mathbf{x} \, dt - 2 \int_0^T \int_{\Omega} (\sigma \partial_t \eta - \mathbf{div} \boldsymbol{\tau}) f \, d\mathbf{x} \, dt + \int_0^T \int_{\Omega} f^2 \, d\mathbf{x} \, dt \\ &= T \|\mathbf{div} \boldsymbol{\tau}_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|k\omega \sigma \boldsymbol{\eta}_k^\perp + \mathbf{div} \boldsymbol{\tau}_k\|_{L^2(\Omega)}^2 \\ &\quad + 2 \left(T(\mathbf{div} \boldsymbol{\tau}_0^c, f_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^N (k\omega \sigma \boldsymbol{\eta}_k^\perp + \mathbf{div} \boldsymbol{\tau}_k, \mathbf{f}_k)_{L^2(\Omega)} \right) \\ &\quad + T \|f_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^{\infty} \|\mathbf{f}_k\|_{L^2(\Omega)}^2 \\ &= T \|\mathbf{div} \boldsymbol{\tau}_0^c + f_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|k\omega \sigma \boldsymbol{\eta}_k^\perp + \mathbf{div} \boldsymbol{\tau}_k + \mathbf{f}_k\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=N+1}^{\infty} \|\mathbf{f}_k\|_{L^2(\Omega)}^2. \end{aligned}$$

Analogously, we can compute the L^2 -norm of \mathcal{R}_2 and obtain

$$\|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 = \int_0^T \int_{\Omega} |\boldsymbol{\tau} - \nu \nabla \eta|^2 \, d\mathbf{x} \, dt = T \|\boldsymbol{\tau}_0^c - \nu \nabla \eta_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|\boldsymbol{\tau}_k - \nu \nabla \boldsymbol{\eta}_k\|_{L^2(\Omega)}^2,$$

where $\boldsymbol{\tau}_k = ((\boldsymbol{\tau}_k^c)^T, (\boldsymbol{\tau}_k^s)^T)^T$. In fact, the L^2 -norms of \mathcal{R}_1 and \mathcal{R}_2 corresponding to every single mode k are decoupled. In the following, we denote by $\|\cdot\|_{L^2}$ the norm $\|\cdot\|_{L^2(\Omega)}$. Altogether, we have the following L^2 -norms of \mathcal{R}_1 and \mathcal{R}_2 :

$$\begin{aligned} \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 &= T \|\mathbf{div} \boldsymbol{\tau}_0^c + f_0^c\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N \|k\omega \sigma \boldsymbol{\eta}_k^\perp + \mathbf{div} \boldsymbol{\tau}_k + \mathbf{f}_k\|_{L^2}^2 + \frac{T}{2} \sum_{k=N+1}^{\infty} \|\mathbf{f}_k\|_{L^2}^2 \\ &= T \|\mathbf{div} \boldsymbol{\tau}_0^c + f_0^c\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\| -k\omega \sigma \boldsymbol{\eta}_k^s + \mathbf{div} \boldsymbol{\tau}_k^c + f_k^c \|_{L^2}^2 + \|k\omega \sigma \boldsymbol{\eta}_k^c + \mathbf{div} \boldsymbol{\tau}_k^s + f_k^s \|_{L^2}^2] \\ &\quad + \frac{T}{2} \sum_{k=N+1}^{\infty} [\|f_k^c\|_{L^2}^2 + \|f_k^s\|_{L^2}^2] \\ &= T \|\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c)\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{1k}^c(\boldsymbol{\eta}_k^s, \boldsymbol{\tau}_k^c)\|_{L^2}^2 + \|\mathcal{R}_{1k}^s(\boldsymbol{\eta}_k^c, \boldsymbol{\tau}_k^s)\|_{L^2}^2] + \frac{T}{2} \sum_{k=N+1}^{\infty} [\|f_k^c\|_{L^2}^2 + \|f_k^s\|_{L^2}^2], \end{aligned}$$

where

$$\begin{aligned}
 \mathcal{R}_{1_0^c}(\boldsymbol{\tau}_0^c) &= \operatorname{div} \boldsymbol{\tau}_0^c + f_0^c, \\
 \mathcal{R}_{1_k^c}(\eta_k^s, \boldsymbol{\tau}_k^c) &= -k\omega \sigma \eta_k^s + \operatorname{div} \boldsymbol{\tau}_k^c + f_k^c, \quad \forall k = 1, \dots, N, \\
 \mathcal{R}_{1_k^s}(\eta_k^c, \boldsymbol{\tau}_k^s) &= k\omega \sigma \eta_k^c + \operatorname{div} \boldsymbol{\tau}_k^s + f_k^s, \quad \forall k = 1, \dots, N,
 \end{aligned} \tag{6.12}$$

and

$$\begin{aligned}
 \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 &= T \|\boldsymbol{\tau}_0^c - \nu \nabla \eta_0^c\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N \|\boldsymbol{\tau}_k - \nu \nabla \eta_k\|_{L^2}^2 \\
 &= T \|\boldsymbol{\tau}_0^c - \nu \nabla \eta_0^c\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|\boldsymbol{\tau}_k^c - \nu \nabla \eta_k^c\|_{L^2}^2 + \|\boldsymbol{\tau}_k^s - \nu \nabla \eta_k^s\|_{L^2}^2] \\
 &= T \|\mathcal{R}_{2_0^c}(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{2_k^c}(\eta_k^c, \boldsymbol{\tau}_k^c)\|_{L^2}^2 + \|\mathcal{R}_{2_k^s}(\eta_k^s, \boldsymbol{\tau}_k^s)\|_{L^2}^2],
 \end{aligned}$$

where

$$\begin{aligned}
 \mathcal{R}_{2_0^c}(\eta_0^c, \boldsymbol{\tau}_0^c) &= \boldsymbol{\tau}_0^c - \nu \nabla \eta_0^c, \\
 \mathcal{R}_{2_k^j}(\eta_k^j, \boldsymbol{\tau}_k^j) &= \boldsymbol{\tau}_k^j - \nu \nabla \eta_k^j, \quad \forall k = 1, \dots, N, \quad j \in \{c, s\}.
 \end{aligned} \tag{6.13}$$

Corollary 6.5. *The error majorants $\mathcal{M}_{|\cdot|}^\oplus(\eta, \boldsymbol{\tau})$ and $\mathcal{M}_{\|\cdot\|}^\oplus(\eta, \boldsymbol{\tau})$ are given by*

$$\begin{aligned}
 \mathcal{M}_{|\cdot|}^\oplus(\eta, \boldsymbol{\tau}) &= \frac{1}{\mu_{1,|\cdot|}} \left(C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} \right) \\
 &= \frac{1}{\mu_{1,|\cdot|}} \left(C_F (T \|\mathcal{R}_{1_0^c}(\boldsymbol{\tau}_0^c)\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{1_k^c}(\eta_k^s, \boldsymbol{\tau}_k^c)\|_{L^2}^2 + \|\mathcal{R}_{1_k^s}(\eta_k^c, \boldsymbol{\tau}_k^s)\|_{L^2}^2] \right. \\
 &\quad \left. + \frac{T}{2} \sum_{k=N+1}^{\infty} [\|f_k^c\|_{L^2}^2 + \|f_k^s\|_{L^2}^2] \right)^{1/2} \\
 &\quad + (T \|\mathcal{R}_{2_0^c}(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{2_k^c}(\eta_k^c, \boldsymbol{\tau}_k^c)\|_{L^2}^2 + \|\mathcal{R}_{2_k^s}(\eta_k^s, \boldsymbol{\tau}_k^s)\|_{L^2}^2])^{1/2},
 \end{aligned}$$

where $\mu_{1,|\cdot|} = \min\{\underline{\nu}, \underline{\sigma}\}$, and

$$\begin{aligned}
 \mathcal{M}_{\|\cdot\|}^\oplus(\eta, \boldsymbol{\tau}) &= \frac{1}{\mu_{1,\|\cdot\|}} \left(\|\mathcal{R}_1(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right)^{1/2} \\
 &= \frac{1}{\mu_{1,\|\cdot\|}} \left(T (\|\mathcal{R}_{1_0^c}(\boldsymbol{\tau}_0^c)\|_{L^2}^2 + \|\mathcal{R}_{2_0^c}(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2}^2) \right. \\
 &\quad \left. + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{1_k^c}(\eta_k^s, \boldsymbol{\tau}_k^c)\|_{L^2}^2 + \|\mathcal{R}_{1_k^s}(\eta_k^c, \boldsymbol{\tau}_k^s)\|_{L^2}^2] \right. \\
 &\quad \left. + \|\mathcal{R}_{2_k^c}(\eta_k^c, \boldsymbol{\tau}_k^c)\|_{L^2}^2 + \|\mathcal{R}_{2_k^s}(\eta_k^s, \boldsymbol{\tau}_k^s)\|_{L^2}^2] \right. \\
 &\quad \left. + \frac{T}{2} \sum_{k=N+1}^{\infty} [\|f_k^c\|_{L^2}^2 + \|f_k^s\|_{L^2}^2] \right)^{1/2}
 \end{aligned}$$

where $\mu_{1,\|\cdot\|} = \min\{\frac{\underline{\nu}}{C_F^2+1}, \underline{\sigma}\}$, respectively.

The term

$$\frac{T}{2} \sum_{k=N+1}^{\infty} \|\mathbf{f}_k\|_{L^2(\Omega)}^2 = \frac{T}{2} \sum_{k=N+1}^{\infty} [\|f_k^c\|_{L^2}^2 + \|f_k^s\|_{L^2}^2]$$

corresponds to the high oscillatory part of the right-hand side f . This term can be controlled due to the knowledge on the data f .

Remark 6.6. *It is obvious that η is the exact solution of problem (6.1) and $\boldsymbol{\tau}$ is the exact flux if and only if the error majorants vanish, i.e.,*

$$\begin{aligned} \mathcal{R}_{1k}^c = 0 \quad \text{and} \quad \mathcal{R}_{2k}^c = 0 \quad \forall k = 0, 1, \dots, N, \\ \mathcal{R}_{1k}^s = 0 \quad \text{and} \quad \mathcal{R}_{2k}^s = 0 \quad \forall k = 1, 2, \dots, N, \end{aligned} \quad (6.14)$$

i.e.,

$$\begin{aligned} -\operatorname{div} \boldsymbol{\tau}_0^c = f_0^c, \quad \boldsymbol{\tau}_0^c = \nu \nabla \eta_0^c, \\ k\omega \sigma \eta_k^s - \operatorname{div} \boldsymbol{\tau}_k^c = f_k^c, \quad -k\omega \sigma \eta_k^c - \operatorname{div} \boldsymbol{\tau}_k^s = f_k^s, \quad \boldsymbol{\tau}_k^c = \nu \nabla \eta_k^c, \quad \boldsymbol{\tau}_k^s = \nu \nabla \eta_k^s, \quad \forall k = 1, \dots, N, \end{aligned}$$

and the data f has a multiharmonic representation, i.e.,

$$f(\mathbf{x}, t) = f_0^c(\mathbf{x}) + \sum_{k=1}^N [f_k^c(\mathbf{x}) \cos(k\omega t) + f_k^s(\mathbf{x}) \sin(k\omega t)].$$

Moreover, η and $\boldsymbol{\tau}$ converge to the exact solution and flux, respectively, if and only if $\eta = \eta_N$ and $\boldsymbol{\tau} = \boldsymbol{\tau}_N$ in (6.11) with N going to infinity and the error majorants corresponding to the modes $k = 0, 1, \dots$ vanish as in (6.14).

Another approach to derive some kind of Fourier series representation of the majorant is to insert the Fourier series ansatz immediately into the bilinear form $a(u - \eta, v)$ and into the functional $\mathcal{F}_\eta(v)$ as defined in (6.2), but now we consider the variational problem (3.4) since $\eta \in H_{0,per}^{1,1}(Q_T)$. Due to the orthogonalities of the cosine and sine functions (2.9), we obtain the following integral identities corresponding to every single mode $k = 1, \dots, N$:

$$\begin{aligned} \int_{\Omega} (\nu(\mathbf{x}) \nabla(\mathbf{u}_k(\mathbf{x}) - \boldsymbol{\eta}_k(\mathbf{x})) \cdot \nabla \mathbf{v}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x})(\mathbf{u}_k(\mathbf{x}) - \boldsymbol{\eta}_k(\mathbf{x})) \cdot \mathbf{v}_k^\perp(\mathbf{x})) \, d\mathbf{x} \\ = \int_{\Omega} (\mathbf{f}_k(\mathbf{x}) \cdot \mathbf{v}_k(\mathbf{x}) - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) - k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k(\mathbf{x}) \cdot \mathbf{v}_k^\perp(\mathbf{x})) \, d\mathbf{x}, \end{aligned} \quad (6.15)$$

which are valid for all $\mathbf{v}_k \in (H_0^1(\Omega))^2$, and, in the case $k = 0$, we obtain the integral identity

$$\int_{\Omega} \nu(\mathbf{x}) \nabla(u_0^c(\mathbf{x}) - \eta_0^c(\mathbf{x})) \cdot \nabla v_0^c(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} (f_0^c(\mathbf{x}) v_0^c(\mathbf{x}) - \nu(\mathbf{x}) \nabla \eta_0^c(\mathbf{x}) \cdot \nabla v_0^c(\mathbf{x})) \, d\mathbf{x}, \quad (6.16)$$

which is valid for all $v_0^c \in H_0^1(\Omega)$. We define the left hand sides of (6.15) and (6.16) by

$$a_k(\mathbf{u}_k - \boldsymbol{\eta}_k, \mathbf{v}_k) \quad \text{and} \quad a_0(u_0^c - \eta_0^c, v_0^c),$$

respectively. We start with the case $k = 1, \dots, N$. Let us compute an upper bound for the errors

$$\mathbf{e}_k := \mathbf{u}_k - \boldsymbol{\eta}_k, \quad \forall k = 1, \dots, N,$$

in $(H_0^1(\Omega))^2$. First, we define the bilinear form $a_k(\cdot, \cdot)$ as in (3.8), i.e.,

$$a_k(\mathbf{u}_k, \mathbf{v}_k) = \int_{\Omega} (\nu(\mathbf{x}) \nabla \mathbf{u}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) + k\omega \sigma(\mathbf{x}) \mathbf{u}_k(\mathbf{x}) \cdot \mathbf{v}_k^\perp(\mathbf{x})) \, d\mathbf{x}. \quad (6.17)$$

Following the proof of Theorem 3.1, we analogously deduce the inf-sup condition

$$\sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{a_k(\mathbf{u}_k - \boldsymbol{\eta}_k, \mathbf{v}_k)}{\|\mathbf{v}_k\|_{H^1(\Omega)}} \geq \underline{c}_k \|\mathbf{u}_k - \boldsymbol{\eta}_k\|_{H^1(\Omega)} \quad (6.18)$$

with the inf-sup constant $\underline{c}_k = \min\{\underline{\nu}, k\omega \underline{\sigma}\}$. We denote the right-hand side of (6.15) by $\mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k)$, i.e.,

$$\mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k) = \int_{\Omega} (\mathbf{f}_k(\mathbf{x}) \cdot \mathbf{v}_k(\mathbf{x}) - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k(\mathbf{x}) \cdot \nabla \mathbf{v}_k(\mathbf{x}) - k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k(\mathbf{x}) \cdot \mathbf{v}_k^\perp(\mathbf{x})) \, d\mathbf{x},$$

and need to find an upper bound of

$$\sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{\mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k)}{\|\mathbf{v}_k\|_{H^1(\Omega)}}.$$

Moreover, we introduce a function of vector-valued functions

$$\boldsymbol{\tau}_k = (\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s)^T, \quad \boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s \in H(\operatorname{div}, \Omega) := \{\boldsymbol{\tau} \in [L^2(\Omega)]^d : \operatorname{div} \boldsymbol{\tau} \in L^2(\Omega)\},$$

with the weak divergence fulfilling

$$\int_{\Omega} \operatorname{div} \boldsymbol{\tau} v \, d\mathbf{x} = - \int_{\Omega} \boldsymbol{\tau} \cdot \nabla v \, d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega).$$

Due to the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k) &= \int_{\Omega} (\mathbf{f}_k \cdot \mathbf{v}_k - k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k \cdot \mathbf{v}_k^\perp + \operatorname{div} \boldsymbol{\tau}_k \cdot \mathbf{v}_k + \boldsymbol{\tau}_k \cdot \nabla \mathbf{v}_k - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k \cdot \nabla \mathbf{v}_k) \, d\mathbf{x} \\ &= \int_{\Omega} (\mathbf{f}_k \cdot \mathbf{v}_k + k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k^\perp \cdot \mathbf{v}_k + \operatorname{div} \boldsymbol{\tau}_k \cdot \mathbf{v}_k + (\boldsymbol{\tau}_k - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k) \cdot \nabla \mathbf{v}_k) \, d\mathbf{x} \\ &\leq \|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} \|\nabla \mathbf{v}_k\|_{L^2(\Omega)} \\ &\leq \left(\|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 \right)^{1/2} \left(\|\mathbf{v}_k\|_{L^2(\Omega)}^2 + \|\nabla \mathbf{v}_k\|_{L^2(\Omega)}^2 \right)^{1/2} \\ &= \left(\|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 \right)^{1/2} \|\mathbf{v}_k\|_{H^1(\Omega)} \end{aligned} \quad (6.19)$$

with

$$\begin{aligned} \mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k) &= k\omega \sigma \boldsymbol{\eta}_k^\perp + \operatorname{div} \boldsymbol{\tau}_k + \mathbf{f}_k = (-k\omega \sigma \boldsymbol{\eta}_k^s + \operatorname{div} \boldsymbol{\tau}_k^c + \mathbf{f}_k^c, k\omega \sigma \boldsymbol{\eta}_k^c + \operatorname{div} \boldsymbol{\tau}_k^s + \mathbf{f}_k^s)^T \\ &= (\mathcal{R}_{1k}^c(\boldsymbol{\eta}_k^s, \boldsymbol{\tau}_k^c), \mathcal{R}_{1k}^s(\boldsymbol{\eta}_k^c, \boldsymbol{\tau}_k^s))^T \end{aligned}$$

and

$$\begin{aligned} \mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k) &= \boldsymbol{\tau}_k - \nu \nabla \boldsymbol{\eta}_k = (\boldsymbol{\tau}_k^c - \nu \nabla \boldsymbol{\eta}_k^c, \boldsymbol{\tau}_k^s - \nu \nabla \boldsymbol{\eta}_k^s)^T \\ &= (\mathcal{R}_{2k}^c(\boldsymbol{\eta}_k^c, \boldsymbol{\tau}_k^c), \mathcal{R}_{2k}^s(\boldsymbol{\eta}_k^s, \boldsymbol{\tau}_k^s))^T. \end{aligned}$$

Hence, we have derived the same results as in (6.12) and (6.13) for every mode $k = 1, \dots, N$. Using the estimate (6.19) together with the inf-sup condition (6.18), we finally arrive at the following upper bounds for every single mode $k = 1, \dots, N$:

Corollary 6.7. *Let $\boldsymbol{\eta}_k \in (H_0^1(\Omega))^2$ and the bilinear form $a_k(\cdot, \cdot)$ defined by (6.17) satisfy (6.18). Then, we obtain the estimate*

$$\|\mathbf{u}_k - \boldsymbol{\eta}_k\|_{H^1(\Omega)} \leq \frac{1}{\underline{c}_k} \left(\|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 \right)^{1/2} =: \mathcal{M}_{\|\cdot\|}^{\oplus k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k), \quad (6.20)$$

where $\underline{c}_k = \min\{\underline{\nu}, k\omega \underline{\sigma}\}$ and $\boldsymbol{\tau}_k = (\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s)^T$ with $\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s \in H(\operatorname{div}, \Omega)$.

Using the inf-sup condition

$$\begin{aligned}
\sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{a_k(\mathbf{u}_k, \mathbf{v}_k)}{|\mathbf{v}_k|_{H^1(\Omega)}} &= \sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{(\nu \nabla \mathbf{u}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{v}_k^\perp)_{L^2(\Omega)}}{|\mathbf{v}_k|_{H^1(\Omega)}} \\
&\geq \frac{(\nu \nabla \mathbf{u}_k, \nabla (\mathbf{u}_k - \mathbf{u}_k^\perp))_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, (\mathbf{u}_k - \mathbf{u}_k^\perp)^\perp)_{L^2(\Omega)}}{|\mathbf{u}_k - \mathbf{u}_k^\perp|_{H^1(\Omega)}} \\
&= \frac{(\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}}{\sqrt{2} |\mathbf{u}_k|_{H^1(\Omega)}} \\
&\geq \frac{\nu \|\nabla \mathbf{u}_k\|_{L^2(\Omega)}^2 + k\omega \sigma \|\mathbf{u}_k\|_{L^2(\Omega)}^2}{\sqrt{2} |\mathbf{u}_k|_{H^1(\Omega)}} \\
&\geq \frac{\nu \|\nabla \mathbf{u}_k\|_{L^2(\Omega)}^2 + \frac{k\omega \sigma}{C_F^2 + 1} \|\nabla \mathbf{u}_k\|_{L^2(\Omega)}^2}{\sqrt{2} |\mathbf{u}_k|_{H^1(\Omega)}} \\
&\geq \frac{\min\{\nu, \frac{k\omega \sigma}{C_F^2 + 1}\}}{\sqrt{2}} |\mathbf{u}_k|_{H^1(\Omega)}
\end{aligned} \tag{6.21}$$

together with the estimate

$$\begin{aligned}
\mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k) &\leq \|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} \|\nabla \mathbf{v}_k\|_{L^2(\Omega)} \\
&\leq (C_F \|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}) |\mathbf{v}_k|_{H^1(\Omega)}
\end{aligned}$$

yields the following error majorant for the H^1 -seminorm:

Corollary 6.8. *Let $\boldsymbol{\eta}_k \in (H_0^1(\Omega))^2$ and the bilinear form $a_k(\cdot, \cdot)$ defined by (6.17) satisfy (6.21). Then, we obtain the estimate*

$$\begin{aligned}
|\mathbf{u}_k - \boldsymbol{\eta}_k|_{H^1(\Omega)} &\leq \frac{\sqrt{2}}{\min\{\nu, \frac{k\omega \sigma}{C_F^2 + 1}\}} (C_F \|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}) \\
&=: \mathcal{M}_{|\cdot|}^{\oplus k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k),
\end{aligned} \tag{6.22}$$

where $\boldsymbol{\tau}_k = (\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s)^T$ with $\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s \in H(\text{div}, \Omega)$.

Now, let us consider the case $k = 0$. Here, we want to compute an upper bound for the error

$$e_0^c := u_0^c - \eta_0^c$$

in $H_0^1(\Omega)$. We define the bilinear form $a_0(\cdot, \cdot)$ as in (3.9), i.e.,

$$a_0(u_0^c, v_0^c) = \int_{\Omega} \nu(\mathbf{x}) \nabla u_0^c(\mathbf{x}) \cdot \nabla v_0^c(\mathbf{x}) \, d\mathbf{x}. \tag{6.23}$$

From the proof of Theorem 3.1, we know that

$$\sup_{0 \neq v_0^c \in H_0^1(\Omega)} \frac{a_0(u_0^c - \eta_0^c, v_0^c)}{\|v_0^c\|_{H^1(\Omega)}} \geq \underline{c}_0 \|u_0^c - \eta_0^c\|_{H^1(\Omega)} \tag{6.24}$$

with the inf-sup constant $\underline{c}_0 = \frac{\nu}{C_F^2 + 1}$. Moreover, one can easily show that

$$\sup_{0 \neq v_0^c \in H_0^1(\Omega)} \frac{a_0(u_0^c - \eta_0^c, v_0^c)}{|v_0^c|_{H^1(\Omega)}} \geq \frac{a_0(u_0^c - \eta_0^c, u_0^c - \eta_0^c)}{|u_0^c - \eta_0^c|_{H^1(\Omega)}} \geq \nu |u_0^c - \eta_0^c|_{H^1(\Omega)}, \tag{6.25}$$

since ν satisfies the assumptions (2.29). We denote the right-hand side of (6.16) by $\mathcal{F}_{\eta_0^c}(v_0^c)$, i.e.,

$$\mathcal{F}_{\eta_0^c}(v_0^c) = \int_{\Omega} (f_0^c(\mathbf{x}) v_0^c(\mathbf{x}) - \nu(\mathbf{x}) \nabla \eta_0^c(\mathbf{x}) \cdot \nabla v_0^c(\mathbf{x})) \, d\mathbf{x},$$

and need to find an upper bound of

$$\sup_{0 \neq v_0^c \in H_0^1(\Omega)} \frac{\mathcal{F}_{\eta_0^c}(v_0^c)}{\|v_0^c\|_{H^1(\Omega)}} \quad \text{and} \quad \sup_{0 \neq v_0^c \in H_0^1(\Omega)} \frac{\mathcal{F}_{\eta_0^c}(v_0^c)}{|v_0^c|_{H^1(\Omega)}}.$$

Again, we introduce a vector-valued function

$$\boldsymbol{\tau}_0^c \in H(\operatorname{div}, \Omega).$$

Due to the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathcal{F}_{\eta_0^c}(v_0^c) &= \int_{\Omega} (f_0^c v_0^c + \operatorname{div} \boldsymbol{\tau}_0^c v_0^c + \boldsymbol{\tau}_0^c \cdot \nabla v_0^c - \nu(\mathbf{x}) \nabla \eta_0^c \cdot \nabla v_0^c) \, d\mathbf{x} \\ &= \int_{\Omega} ((f_0^c + \operatorname{div} \boldsymbol{\tau}_0^c) v_0^c + (\boldsymbol{\tau}_0^c - \nu(\mathbf{x}) \nabla \eta_0^c) \cdot \nabla v_0^c) \, d\mathbf{x} \\ &\leq \|\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c)\|_{L^2(\Omega)} \|v_0^c\|_{L^2(\Omega)} + \|\mathcal{R}_{20}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)} \|\nabla v_0^c\|_{L^2(\Omega)} \\ &\leq \left(\|\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{20}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 \right)^{1/2} \|v_0^c\|_{H^1(\Omega)} \end{aligned} \quad (6.26)$$

and

$$\begin{aligned} \mathcal{F}_{\eta_0^c}(v_0^c) &\leq \|\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c)\|_{L^2(\Omega)} \|v_0^c\|_{L^2(\Omega)} + \|\mathcal{R}_{20}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)} \|\nabla v_0^c\|_{L^2(\Omega)} \\ &\leq (C_F \|\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c)\|_{L^2(\Omega)} + \|\mathcal{R}_{20}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}) |v_0^c|_{H^1(\Omega)} \end{aligned} \quad (6.27)$$

with

$$\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c) = f_0^c + \operatorname{div} \boldsymbol{\tau}_0^c \quad \text{and} \quad \mathcal{R}_{20}^c(\eta_0^c, \boldsymbol{\tau}_0^c) = \boldsymbol{\tau}_0^c - \nu \nabla \eta_0^c.$$

Hence, we have derived the same results as in (6.12) and (6.13) for $k = 0$. Using the estimates (6.26) and (6.27) together with the inf-sup conditions (6.24) and (6.25), we finally arrive at the following upper bounds for the case $k = 0$, which correspond to the H^1 -norm and seminorm, respectively:

Corollary 6.9. *Let $\eta_0^c \in H_0^1(\Omega)$ and the bilinear form $a_0(\cdot, \cdot)$ defined as in (6.23) satisfy (6.24). Then,*

$$\|u_0^c - \eta_0^c\|_{H^1(\Omega)} \leq \frac{1}{\underline{c}_0} \left(\|\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{20}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 \right)^{1/2} =: \mathcal{M}_{\|\cdot\|}^{\oplus_0}(\eta_0^c, \boldsymbol{\tau}_0^c), \quad (6.28)$$

where $\underline{c}_0 = \frac{\nu}{C_F^2 + 1}$ and $\boldsymbol{\tau}_0^c \in H(\operatorname{div}, \Omega)$.

Corollary 6.10. *Let $\eta_0^c \in H_0^1(\Omega)$ and the bilinear form $a_0(\cdot, \cdot)$ defined as in (6.23) satisfy (6.25). Then,*

$$|u_0^c - \eta_0^c|_{H^1(\Omega)} \leq \frac{1}{\underline{\nu}} (C_F \|\mathcal{R}_{10}^c(\boldsymbol{\tau}_0^c)\|_{L^2(\Omega)} + \|\mathcal{R}_{20}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}) =: \mathcal{M}_{|\cdot|}^{\oplus_0}(\eta_0^c, \boldsymbol{\tau}_0^c), \quad (6.29)$$

where $\boldsymbol{\tau}_0^c \in H(\operatorname{div}, \Omega)$.

Now, we go back to the beginning of the chapter, but consider error estimates regarding approximations η for the solution u that are less regular than $H_{0,per}^{1,1}(Q_T)$.

A second a posteriori error result

In the following, we deduce another upper bound of the error $e := u - \eta$, which is valid for approximations that are less regular with respect to the time, i.e.,

$$\eta \in H_0^{1, \frac{1}{2}}(Q_T).$$

In fact, we will choose a multiharmonic finite element approximation u_{Nh} as η , which is, of course, more regular in time, but the abstract functional a posteriori error estimates, which we obtain, can be used in a more general setting.

Let us again consider the functional

$$\mathcal{F}_\eta(v) = \int_0^T \int_\Omega \left(f v - \sigma(\mathbf{x}) \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp - \nu(\mathbf{x}) \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt$$

defined for all $v \in H_0^{1, \frac{1}{2}}(Q_T)$. Besides the vector-valued function $\boldsymbol{\tau} \in H(\text{div}, Q_T)$, let us introduce the function

$$\kappa \in H^{0, \frac{1}{2}}(Q_T),$$

which fulfills the identity

$$\int_0^T \kappa \partial_t^{1/2} v^\perp dt = - \int_0^T \partial_t^{1/2} \kappa^\perp v dt \quad (6.30)$$

for all $v \in H^{0, \frac{1}{2}}(Q_T)$. This identity is defined in the Fourier space according to Definition 3.2, i.e.,

$$(\kappa, \partial_t^{1/2} v)_{L^2(Q_T)} := \frac{T}{2} \sum_{k=1}^{\infty} (k\omega)^{1/2} (\boldsymbol{\kappa}_k, \mathbf{v}_k)_{L^2(\Omega)} \quad (6.31)$$

yielding the following definitions in the Fourier space:

$$\begin{aligned} \partial_t^{1/2} \kappa(\mathbf{x}, t) &:= \sum_{k=1}^{\infty} (k\omega)^{1/2} [\kappa_k^c(\mathbf{x}) \cos(k\omega t) + \kappa_k^s(\mathbf{x}) \sin(k\omega t)] \\ &= \sum_{k=1}^{\infty} (k\omega)^{1/2} \underbrace{(\kappa_k^c(\mathbf{x}), \kappa_k^s(\mathbf{x}))}_{= \boldsymbol{\kappa}_k^T} \cdot \begin{pmatrix} \cos(k\omega t) \\ \sin(k\omega t) \end{pmatrix}, \\ \partial_t^{1/2} \kappa^\perp(\mathbf{x}, t) &:= \sum_{k=1}^{\infty} (k\omega)^{1/2} [-\kappa_k^s(\mathbf{x}) \cos(k\omega t) + \kappa_k^c(\mathbf{x}) \sin(k\omega t)] \\ &= \sum_{k=1}^{\infty} (k\omega)^{1/2} \underbrace{(-\kappa_k^s(\mathbf{x}), \kappa_k^c(\mathbf{x}))}_{= (\boldsymbol{\kappa}_k^\perp)^T} \cdot \begin{pmatrix} \cos(k\omega t) \\ \sin(k\omega t) \end{pmatrix}. \end{aligned}$$

Hence,

$$(\kappa, \partial_t^{1/2} v^\perp)_{L^2(Q_T)} = \frac{T}{2} \sum_{k=1}^{\infty} (k\omega)^{1/2} (\boldsymbol{\kappa}_k, \mathbf{v}_k^\perp)_{L^2(\Omega)} = -(\partial_t^{1/2} \kappa, v^\perp)_{L^2(Q_T)}$$

and

$$\begin{aligned} (\kappa, \partial_t^{1/2} v^\perp)_{L^2(Q_T)} &= \frac{T}{2} \sum_{k=1}^{\infty} (k\omega)^{1/2} (\boldsymbol{\kappa}_k, \mathbf{v}_k^\perp)_{L^2(\Omega)} = \frac{T}{2} \sum_{k=1}^{\infty} (k\omega)^{1/2} (-\boldsymbol{\kappa}_k^\perp, \mathbf{v}_k)_{L^2(\Omega)} \\ &= -\frac{T}{2} \sum_{k=1}^{\infty} (k\omega)^{1/2} (\boldsymbol{\kappa}_k^\perp, \mathbf{v}_k)_{L^2(\Omega)} = -(\partial_t^{1/2} \kappa^\perp, v)_{L^2(Q_T)}. \end{aligned}$$

Remark 6.11. Identity (6.30) and definition (6.31) coincide with the identities (3.10) like, e.g.,

$$(\partial_t^{1/2} u, \partial_t^{1/2} v^\perp)_{L^2(Q_T)} = (\partial_t u, v)_{L^2(Q_T)} \quad \forall u \in H_{per}^{0,1}(Q_T) \quad \forall v \in H^{0, \frac{1}{2}}(Q_T),$$

since we have that

$$(\partial_t^{1/2} u, \partial_t^{1/2} v^\perp)_{L^2(Q_T)} = \frac{T}{2} \sum_{k=1}^{\infty} (k\omega)(\mathbf{u}_k, \mathbf{v}_k^\perp)_{L^2(\Omega)} = \frac{T}{2} \sum_{k=1}^{\infty} (k\omega)(-\mathbf{u}_k^\perp, \mathbf{v}_k)_{L^2(\Omega)} = (\partial_t u, v)_{L^2(Q_T)}.$$

We rearrange the functional $\mathcal{F}_\eta(v)$ and write it as

$$\begin{aligned} \mathcal{F}_\eta(v) &= \int_0^T \int_\Omega \left(f v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(f v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp + \sigma \partial_t^{1/2} \kappa^\perp v + \operatorname{div} \boldsymbol{\tau} v \right. \\ &\quad \left. + \boldsymbol{\tau} \cdot \nabla v - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(f v + \operatorname{div} \boldsymbol{\tau} v + \sigma \partial_t^{1/2} \kappa^\perp v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp \right. \\ &\quad \left. + \boldsymbol{\tau} \cdot \nabla v - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left((f + \operatorname{div} \boldsymbol{\tau} + \sigma \partial_t^{1/2} \kappa^\perp) v + (\sigma(-\partial_t^{1/2} \eta + \kappa)) \partial_t^{1/2} v^\perp + (\boldsymbol{\tau} - \nu \nabla \eta) \cdot \nabla v \right) d\mathbf{x} dt \end{aligned}$$

for all $v \in H_0^{1, \frac{1}{2}}(Q_T)$.

Remark 6.12. We can interpret $\boldsymbol{\tau}$ as “an image” of $\nu \nabla u$ and κ as “an image” of $\partial_t^{1/2} u$.

Let

$$\begin{aligned} \mathcal{R}_1(\boldsymbol{\tau}, \kappa) &= f + \operatorname{div} \boldsymbol{\tau} + \sigma \partial_t^{1/2} \kappa^\perp, \\ \mathcal{R}_2(\boldsymbol{\tau}, \eta) &= \boldsymbol{\tau} - \nu \nabla \eta, \\ \mathcal{R}_3(\kappa, \eta) &= \sigma(\kappa - \partial_t^{1/2} \eta). \end{aligned}$$

Then, the functional $\mathcal{F}_\eta(v)$ can be estimated from above as follows

$$\begin{aligned} \mathcal{F}_\eta(v) &\leq \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)} \\ &\leq \left(\|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2} \|v\|_{H^{1, \frac{1}{2}}(Q_T)}, \end{aligned}$$

using the Cauchy-Schwarz inequality and $\|\partial_t^{1/2} v^\perp\|_{L^2(Q_T)} = \|\partial_t^{1/2} v\|_{L^2(Q_T)}$, since

$$\|\partial_t^{1/2} v^\perp\|_{L^2(Q_T)}^2 = \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{v}_k^\perp\|_{L^2(\Omega)}^2 = \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{v}_k\|_{L^2(\Omega)}^2 = \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2.$$

Altogether, we obtain the upper bound

$$\begin{aligned} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}} &\leq \left(\|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2}, \end{aligned} \tag{6.32}$$

and, finally, deduce the following theorem:

Theorem 6.13. Let $\eta \in H_0^{1, \frac{1}{2}}(Q_T)$ and the bilinear form $a(\cdot, \cdot)$ defined in (3.17) satisfy (6.3). Then,

$$\begin{aligned} \|u - \eta\|_{H^{1, \frac{1}{2}}(Q_T)} &\leq \frac{1}{\mu_1} \left(\|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &=: \mathcal{M}_{\|\cdot\|}^{\oplus}(\eta, \boldsymbol{\tau}, \kappa), \end{aligned} \quad (6.33)$$

where $\boldsymbol{\tau} \in H(\operatorname{div}, Q_T)$, $\kappa \in H^{0, \frac{1}{2}}(Q_T)$ and $\mu_1 = \min\{\frac{\nu}{C_F^2 + 1}, \underline{\sigma}\}$.

Proof. From (6.3) follows that

$$\|u - \eta\|_{H^{1, \frac{1}{2}}(Q_T)} \leq \frac{1}{\mu_1} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u - \eta, v)}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}} = \frac{1}{\mu_1} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{\|v\|_{H^{1, \frac{1}{2}}(Q_T)}},$$

which leads together with (6.32) to the final estimate (6.33). \square

Remark 6.14. If $\mathcal{R}_1(\boldsymbol{\tau}, \kappa) = 0$, $\mathcal{R}_2(\boldsymbol{\tau}, \eta) = 0$ and $\mathcal{R}_3(\kappa, \eta) = 0$, then

$$\begin{aligned} -\sigma \partial_t^{1/2} \kappa^\perp - \operatorname{div} \boldsymbol{\tau} &= f, \\ \boldsymbol{\tau} &= \nu \nabla \eta, \\ \kappa &= \partial_t^{1/2} \eta. \end{aligned}$$

Since η satisfies the Dirichlet condition on Σ_T , η is the solution. In other words, $\mathcal{M}_{\|\cdot\|}^{\oplus}(\eta, \boldsymbol{\tau}, \kappa)$ vanishes if and only if η is the exact solution, $\boldsymbol{\tau}$ is the exact flux and κ is the exact half time derivative of the solution. Moreover, if $\eta \in H_{0, \text{per}}^{1, 1}(Q_T)$, we derive the original equation (3.1)

$$\sigma \partial_t \eta - \operatorname{div}(\nu \nabla \eta) = f$$

in the weak sense, due to

$$-(\sigma \partial_t^{1/2} (\partial_t^{1/2} \eta)^\perp, v)_{L^2(Q_T)} = (\sigma \partial_t^{1/2} \eta, \partial_t^{1/2} v^\perp)_{L^2(Q_T)} = (\sigma \partial_t \eta, v)_{L^2(Q_T)}$$

using the σ -weighted counterparts of the identities (6.30) and (3.10), cf. (3.13) as well.

It is obvious that similar results to the ones obtained in Theorem 6.2 for the $H^{1, \frac{1}{2}}$ -seminorm can be shown here together with the estimate

$$\begin{aligned} \mathcal{F}_\eta(v) &\leq \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)} \\ &\leq C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)} \\ &= (C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)}) \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)} \\ &\leq \left((C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)})^2 + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2} |v|_{H^{1, \frac{1}{2}}(Q_T)}. \end{aligned} \quad (6.34)$$

Hence, we obtain the following theorem:

Theorem 6.15. Let $\eta \in H_0^{1, \frac{1}{2}}(Q_T)$ and the bilinear form $a(\cdot, \cdot)$ defined by (3.17) satisfy (6.4). Then,

$$\begin{aligned} |u - \eta|_{H^{1, \frac{1}{2}}(Q_T)} &\leq \frac{1}{\mu_1} \left((C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)})^2 + \|\mathcal{R}_3(\kappa, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &=: \mathcal{M}_{|\cdot|}^{\oplus}(\eta, \boldsymbol{\tau}, \kappa), \end{aligned} \quad (6.35)$$

where $\boldsymbol{\tau} \in H(\operatorname{div}, Q_T)$, $\kappa \in H^{0, \frac{1}{2}}(Q_T)$ and $\mu_1 = \min\{\underline{\nu}, \underline{\sigma}\}$.

Proof. The estimate is analogously proven as the one of Theorem 6.13 by using (6.4) and (6.34). \square

Similarly, we can prove a posteriori error estimates using the seminorm $|\cdot|_{V_0}$ introduced in Chapter 3, which is a weighted $H^{1, \frac{1}{2}}$ -seminorm, i.e.,

$$\begin{aligned} |u|_{V_0}^2 &= (\nu \nabla u, \nabla u)_{L^2(Q_T)} + (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)}, \\ &= T (\nu \nabla u_0^c, \nabla u_0^c)_{L^2(\Omega)} + \frac{T}{2} \sum_{k=1}^{\infty} [(\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}], \end{aligned}$$

and, as already mentioned, in fact, a norm due to the Friedrichs inequality. We obtain the following inf-sup and sup-sup conditions regarding the V_0 -seminorm:

Lemma 6.16. *The space-time bilinear form $a(\cdot, \cdot)$ defined by (3.17) fulfills the following inf-sup and sup-sup conditions:*

$$\mu_1 |u|_{V_0} \leq \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u, v)}{|v|_{V_0}} \leq \mu_2 |u|_{V_0} \quad (6.36)$$

for all $u \in H_0^{1, \frac{1}{2}}(Q_T)$ with constants $\mu_1 = 1/\sqrt{2}$ and $\mu_2 = 1$.

Proof. We start with the proof of the sup-sup condition. Using the triangle inequality and the σ - and ν -weighted counterparts of the Cauchy-Schwarz inequalities (3.14) and (3.15), we obtain the estimate

$$\begin{aligned} |a(u, v)| &= \left| \int_0^T \int_{\Omega} (\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} v^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla v) \, d\mathbf{x} \, dt \right| \\ &\leq \left| \int_0^T \int_{\Omega} \sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} v^\perp \, d\mathbf{x} \, dt \right| + \left| \int_0^T \int_{\Omega} \nu(\mathbf{x}) \nabla u \cdot \nabla v \, d\mathbf{x} \, dt \right| \\ &\leq (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v^\perp, \partial_t^{1/2} v^\perp)_{L^2(Q_T)}^{1/2} + (\nu \nabla u, \nabla u)_{L^2(Q_T)}^{1/2} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2} \\ &= (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}^{1/2} + (\nu \nabla u, \nabla u)_{L^2(Q_T)}^{1/2} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2}, \end{aligned}$$

since

$$\begin{aligned} (\sigma \partial_t^{1/2} v^\perp, \partial_t^{1/2} v^\perp)_{L^2(Q_T)} &= \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\sigma \mathbf{v}_k^\perp, \mathbf{v}_k^\perp)_{L^2(\Omega)} \\ &= \frac{T}{2} \sum_{k=1}^{\infty} k\omega (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)} = (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}. \end{aligned}$$

Finally, we prove the sup-sup condition by

$$\begin{aligned} |a(u, v)| &\leq (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}^{1/2} + (\nu \nabla u, \nabla u)_{L^2(Q_T)}^{1/2} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2} \\ &\leq \left((\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)} + (\nu \nabla u, \nabla u)_{L^2(Q_T)} \right)^{1/2} \left((\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)} + (\nu \nabla v, \nabla v)_{L^2(Q_T)} \right)^{1/2} \\ &= \mu_2 |u|_{V_0} |v|_{V_0} \end{aligned}$$

with the constant $\mu_2 = 1$. Next, we prove the inf-sup condition by choosing the test function $v = u - u^\perp$ and using the σ - and ν -weighted orthogonality relations (3.13). With the choice $v = u - u^\perp$, we obtain

$$\begin{aligned} a(u, u) &= \int_0^T \int_{\Omega} (\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} u^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla u) \, d\mathbf{x} \, dt = (\nu \nabla u, \nabla u)_{L^2(Q_T)}, \\ a(u, -u^\perp) &= \int_0^T \int_{\Omega} (\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} u + \nu(\mathbf{x}) \nabla u \cdot \nabla u^\perp) \, d\mathbf{x} \, dt = (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)}, \\ a(u, u - u^\perp) &= |u|_{V_0}^2. \end{aligned}$$

By using the σ - and ν -weighted orthogonalities (3.13) again, we get that $|v|_{V_0} = \sqrt{2}|u|_{V_0}$, i.e.,

$$\begin{aligned}
|v|_{V_0}^2 &= |u - u^\perp|_{V_0}^2 = (\nu \nabla(u - u^\perp), \nabla(u - u^\perp))_{L^2(Q_T)} + (\sigma \partial_t^{1/2}(u - u^\perp), \partial_t^{1/2}(u - u^\perp))_{L^2(Q_T)} \\
&= (\nu \nabla u, \nabla u)_{L^2(Q_T)} - (\nu \nabla u^\perp, \nabla u)_{L^2(Q_T)} - (\nu \nabla u, \nabla u^\perp)_{L^2(Q_T)} + (\nu \nabla u^\perp, \nabla u^\perp)_{L^2(Q_T)} \\
&\quad + (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)} - (\sigma \partial_t^{1/2} u^\perp, \partial_t^{1/2} u)_{L^2(Q_T)} \\
&\quad - (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u^\perp)_{L^2(Q_T)} + (\sigma \partial_t^{1/2} u^\perp, \partial_t^{1/2} u^\perp)_{L^2(Q_T)} \\
&= (\nu \nabla u, \nabla u)_{L^2(Q_T)} + (\nu \nabla u, \nabla u)_{L^2(Q_T)} + (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)} + (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)} \\
&= 2|u|_{V_0}^2.
\end{aligned}$$

Altogether, we arrive at the following estimate of the supremum from below:

$$\sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u, v)}{|v|_{V_0}} \geq \frac{a(u, u - u^\perp)}{|u - u^\perp|_{V_0}} = \frac{|u|_{V_0}^2}{\sqrt{2}|u|_{V_0}} = \frac{1}{\sqrt{2}}|u|_{V_0},$$

which finally yields the inf-sup constant $\mu_1 = 1/\sqrt{2}$. \square

Now, we want to find an upper bound for

$$\sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{|v|_{V_0}}.$$

Hence, we estimate the functional $\mathcal{F}_\eta(v)$ as follows

$$\begin{aligned}
\mathcal{F}_\eta(v) &= \int_0^T \int_\Omega \left(f v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \\
&= \int_0^T \int_\Omega \left(f v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp + \sigma \partial_t^{1/2} \kappa^\perp v \right. \\
&\quad \left. + \operatorname{div} \boldsymbol{\tau} v + \boldsymbol{\tau} \cdot \nabla v - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \\
&= \int_0^T \int_\Omega \left(f v + \operatorname{div} \boldsymbol{\tau} v + \sigma \partial_t^{1/2} \kappa^\perp v + \boldsymbol{\tau} \cdot \nabla v - \nu \nabla \eta \cdot \nabla v \right. \\
&\quad \left. - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp \right) d\mathbf{x} dt \\
&= \int_0^T \int_\Omega \left((f + \operatorname{div} \boldsymbol{\tau} + \sigma \partial_t^{1/2} \kappa^\perp) v + (\boldsymbol{\tau} - \nu \nabla \eta) \cdot \nabla v \right. \\
&\quad \left. + \sigma (-\partial_t^{1/2} \eta + \kappa) \partial_t^{1/2} v^\perp \right) d\mathbf{x} dt \tag{6.37} \\
&\leq \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\
&\quad + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}^{1/2} \\
&\leq C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\
&\quad + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}^{1/2} \\
&= (C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)}) \|\nabla v\|_{L^2(Q_T)} \\
&\quad + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}^{1/2} \\
&\leq \left((C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)})^2 + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)} \right)^{1/2} \\
&\quad \times \left(\|\nabla v\|_{L^2(Q_T)}^2 + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)} \right)^{1/2}
\end{aligned}$$

for all $v \in H_0^{1, \frac{1}{2}}(Q_T)$, where

$$\begin{aligned}\mathcal{R}_1(\boldsymbol{\tau}, \kappa) &= f + \operatorname{div} \boldsymbol{\tau} + \sigma \partial_t^{1/2} \kappa^\perp, \\ \mathcal{R}_2(\boldsymbol{\tau}, \eta) &= \boldsymbol{\tau} - \nu \nabla \eta, \\ \mathcal{R}_3(\kappa, \eta) &= \kappa - \partial_t^{1/2} \eta.\end{aligned}$$

Hence, it follows an a posteriori error result for the V_0 -seminorm:

Theorem 6.17. *Let $\eta \in H_0^{1, \frac{1}{2}}(Q_T)$ and the bilinear form $a(\cdot, \cdot)$ satisfy (6.36). Then,*

$$\begin{aligned}|u - \eta|_{V_0} \leq \mu_1 &\left((C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)})^2 \right. \\ &\left. + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)} \right)^{1/2} =: \mathcal{M}_{|\cdot|_{V_0}}^\oplus(\eta, \boldsymbol{\tau}, \kappa),\end{aligned}\tag{6.38}$$

where $\boldsymbol{\tau} \in H(\operatorname{div}, Q_T)$, $\kappa \in H^{0, \frac{1}{2}}(Q_T)$ and $\mu_1 = \frac{\sqrt{2}}{\min\{\sqrt{\underline{\nu}}, 1\}}$.

Proof. Using the left inequality of (6.36) and the estimate (6.37), we obtain the following upper bound using the notation $\|\cdot\|_{L^2} = \|\cdot\|_{L^2(Q_T)}$:

$$\begin{aligned}|u - \eta|_{V_0} &\leq \sqrt{2} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u - \eta, v)}{|v|_{V_0}} = \sqrt{2} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{|v|_{V_0}} \\ &\leq \sqrt{2} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\left((C_F \|\mathcal{R}_1\|_{L^2} + \|\mathcal{R}_2\|_{L^2})^2 + (\sigma \mathcal{R}_3, \mathcal{R}_3)_{L^2} \right)^{1/2} \left(\|\nabla v\|_{L^2}^2 + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2} \right)^{1/2}}{\left((\nu \nabla v, \nabla v)_{L^2} + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2} \right)^{1/2}} \\ &\leq \sqrt{2} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\left((C_F \|\mathcal{R}_1\|_{L^2} + \|\mathcal{R}_2\|_{L^2})^2 + (\sigma \mathcal{R}_3, \mathcal{R}_3)_{L^2} \right)^{1/2} \left(\|\nabla v\|_{L^2}^2 + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2} \right)^{1/2}}{\left(\underline{\nu} \|\nabla v\|_{L^2}^2 + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2} \right)^{1/2}} \\ &\leq \sqrt{2} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\left((C_F \|\mathcal{R}_1\|_{L^2} + \|\mathcal{R}_2\|_{L^2})^2 + (\sigma \mathcal{R}_3, \mathcal{R}_3)_{L^2} \right)^{1/2} \left(\|\nabla v\|_{L^2}^2 + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2} \right)^{1/2}}{\min\{\sqrt{\underline{\nu}}, 1\} \left(\|\nabla v\|_{L^2}^2 + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2} \right)^{1/2}} \\ &= \frac{\sqrt{2}}{\min\{\sqrt{\underline{\nu}}, 1\}} \left((C_F \|\mathcal{R}_1(\boldsymbol{\tau}, \kappa)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \eta)\|_{L^2(Q_T)})^2 + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)} \right)^{1/2},\end{aligned}$$

which immediately leads to (6.38). \square

In order to derive a posteriori estimates for the full weighted $H^{1, \frac{1}{2}}$ -norm, which we define as

$$\|v\|_{V_0}^2 = \|v\|_{L^2(Q_T)}^2 + (\nu \nabla v, \nabla v)_{L^2(Q_T)} + (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)},$$

we have to rearrange the functional $\mathcal{F}_\eta(v)$ again as follows

$$\begin{aligned}\mathcal{F}_\eta(v) &= \int_0^T \int_\Omega \left(f v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(f v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp + \sigma \partial_t^{1/2} \kappa^\perp v \right. \\ &\quad \left. + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}) v + (\nu \tilde{\boldsymbol{\tau}}) \cdot \nabla v - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(f v + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}) v + \sigma \partial_t^{1/2} \kappa^\perp v - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp \right. \\ &\quad \left. + (\nu \tilde{\boldsymbol{\tau}}) \cdot \nabla v - \nu \nabla \eta \cdot \nabla v \right) d\mathbf{x} dt,\end{aligned}$$

which finally yields

$$\mathcal{F}_\eta(v) = \int_0^T \int_\Omega \left((f + \operatorname{div}(\nu \tilde{\tau}) + \sigma \partial_t^{1/2} \kappa^\perp) v + \sigma (-\partial_t^{1/2} \eta + \kappa) \partial_t^{1/2} v^\perp + \nu (\tilde{\tau} - \nabla \eta) \cdot \nabla v \right) d\mathbf{x} dt$$

for all $v \in H_0^{1, \frac{1}{2}}(Q_T)$. Here, we have introduced a vector-valued function $\tilde{\tau}$ fulfilling the identity

$$\int_\Omega \operatorname{div}(\nu \tilde{\tau}) v d\mathbf{x} = - \int_\Omega (\nu \tilde{\tau}) \cdot \nabla v d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega).$$

Let

$$\begin{aligned} \mathcal{R}_1(\tilde{\tau}, \kappa) &= f + \operatorname{div}(\nu \tilde{\tau}) + \sigma \partial_t^{1/2} \kappa^\perp, \\ \mathcal{R}_2(\tilde{\tau}, \eta) &= \tilde{\tau} - \nabla \eta, \\ \mathcal{R}_3(\kappa, \eta) &= \kappa - \partial_t^{1/2} \eta. \end{aligned}$$

Then, the functional $\mathcal{F}_\eta(v)$ can be estimated from above as follows

$$\begin{aligned} \mathcal{F}_\eta(v) &\leq \|\mathcal{R}_1(\tilde{\tau}, \kappa)\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + (\nu \mathcal{R}_2(\tilde{\tau}, \eta), \mathcal{R}_2(\tilde{\tau}, \eta))_{L^2(Q_T)}^{1/2} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2} \\ &\quad + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}^{1/2} \\ &\leq \left(\|\mathcal{R}_1(\tilde{\tau}, \kappa)\|_{L^2(Q_T)}^2 + (\nu \mathcal{R}_2(\tilde{\tau}, \eta), \mathcal{R}_2(\tilde{\tau}, \eta))_{L^2(Q_T)} + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)} \right)^{1/2} \|v\|_{V_0}, \end{aligned}$$

which leads to the upper bound

$$\begin{aligned} \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{\mathcal{F}_\eta(v)}{\|v\|_{V_0}} & \leq \left(\|\mathcal{R}_1(\tilde{\tau}, \kappa)\|_{L^2(Q_T)}^2 + (\nu \mathcal{R}_2(\tilde{\tau}, \eta), \mathcal{R}_2(\tilde{\tau}, \eta))_{L^2(Q_T)} + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)} \right)^{1/2}. \end{aligned} \quad (6.39)$$

Moreover, we can prove the following inf-sup conditions:

Lemma 6.18. *The space-time bilinear form $a(\cdot, \cdot)$ defined by (3.17) fulfills the following inf-sup and sup-sup conditions:*

$$\mu_1 \|u\|_{V_0} \leq \sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u, v)}{\|v\|_{V_0}} \leq \mu_2 \|u\|_{V_0} \quad (6.40)$$

for all $u \in H_0^{1, \frac{1}{2}}(Q_T)$ with constants $\mu_1 = \min\{1, \frac{\nu}{C_F^2}\}/\sqrt{5}$ and $\mu_2 = 1$.

Proof. The sup-sup condition is analogously proven as it is done in Lemma 6.16 with the final result

$$|a(u, v)| \leq |u|_{V_0} |v|_{V_0} \leq \mu_2 \|u\|_{V_0} \|v\|_{V_0},$$

where $\mu_2 = 1$. The inf-sup condition is proven by choosing the test function $v = u + u - u^\perp$ and using the σ - and ν -weighted orthogonality relations (3.13) as well as the Friedrichs inequality (3.19) in the Fourier space. With the choice $v = 2u - u^\perp$, we obtain

$$\begin{aligned} a(u, 2u) &= \int_0^T \int_\Omega \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} (2u)^\perp + \nu(\mathbf{x}) \nabla u \cdot \nabla (2u) \right) d\mathbf{x} dt = 2(\nu \nabla u, \nabla u)_{L^2(Q_T)}, \\ a(u, -u^\perp) &= \int_0^T \int_\Omega \left(\sigma(\mathbf{x}) \partial_t^{1/2} u \partial_t^{1/2} u + \nu(\mathbf{x}) \nabla u \cdot \nabla u^\perp \right) d\mathbf{x} dt = (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)}, \\ a(u, 2u - u^\perp) &= |u|_{V_0}^2 + (\nu \nabla u, \nabla u)_{L^2(Q_T)} \geq |u|_{V_0}^2 + \nu \|\nabla u\|_{L^2(Q_T)}^2 \\ &\geq |u|_{V_0}^2 + \frac{\nu}{C_F^2} \|u\|_{L^2(Q_T)}^2 \geq \min\{1, \frac{\nu}{C_F^2}\} \|u\|_{V_0}^2. \end{aligned}$$

By using orthogonalities (3.11), (3.12) and (3.13), we get that $\|v\|_{V_0} = \sqrt{5} \|u\|_{V_0}$, since

$$\begin{aligned}
 \|v\|_{V_0}^2 &= \|2u - u^\perp\|_{V_0}^2 = \|2u - u^\perp\|_{L^2(Q_T)}^2 + (\nu \nabla(2u - u^\perp), \nabla(2u - u^\perp))_{L^2(Q_T)} \\
 &\quad + (\sigma \partial_t^{1/2}(2u - u^\perp), \partial_t^{1/2}(2u - u^\perp))_{L^2(Q_T)} \\
 &= \|2u\|_{L^2(Q_T)}^2 + \|u^\perp\|_{L^2(Q_T)}^2 + (\nu \nabla(2u), \nabla(2u))_{L^2(Q_T)} + (\nu \nabla u^\perp, \nabla u^\perp)_{L^2(Q_T)} \\
 &\quad + (\sigma \partial_t^{1/2}(2u), \partial_t^{1/2}(2u))_{L^2(Q_T)} + (\sigma \partial_t^{1/2} u^\perp, \partial_t^{1/2} u^\perp)_{L^2(Q_T)} \\
 &= 5 \|u\|_{L^2(Q_T)}^2 + 5 (\nu \nabla u, \nabla u)_{L^2(Q_T)} + 5 (\sigma \partial_t^{1/2} u, \partial_t^{1/2} u)_{L^2(Q_T)} \\
 &= 5 \|u\|_{V_0}^2.
 \end{aligned}$$

Altogether, we arrive at the following estimate of the supremum from below:

$$\sup_{0 \neq v \in H_0^{1, \frac{1}{2}}(Q_T)} \frac{a(u, v)}{\|v\|_{V_0}} \geq \frac{a(u, 2u - u^\perp)}{\|2u - u^\perp\|_{V_0}} \geq \frac{\min\{1, \frac{\nu}{C_F^2}\} \|u\|_{V_0}^2}{\sqrt{5} \|u\|_{V_0}} = \mu_1 \|u\|_{V_0},$$

which finally yields the inf-sup constant $\mu_1 = \min\{1, \frac{\nu}{C_F^2}\} / \sqrt{5}$. \square

Altogether, we obtain the following a posteriori error result for the full V_0 -norm:

Theorem 6.19. *Let $\eta \in H_0^{1, \frac{1}{2}}(Q_T)$ and the bilinear form $a(\cdot, \cdot)$ satisfy (6.40). Then,*

$$\begin{aligned}
 \|u - \eta\|_{V_0} \leq \frac{1}{\mu_1} \left(\|\mathcal{R}_1(\tilde{\tau}, \kappa)\|_{L^2(Q_T)}^2 + (\nu \mathcal{R}_2(\tilde{\tau}, \eta), \mathcal{R}_2(\tilde{\tau}, \eta))_{L^2(Q_T)} \right. \\
 \left. + (\sigma \mathcal{R}_3(\kappa, \eta), \mathcal{R}_3(\kappa, \eta))_{L^2(Q_T)} \right)^{1/2} =: \mathcal{M}_{\|\cdot\|_{V_0}}^\oplus(\eta, \tilde{\tau}, \kappa),
 \end{aligned} \tag{6.41}$$

where $(\nu \tilde{\tau}) \in H(\operatorname{div}, Q_T)$, $\kappa \in H^{0, \frac{1}{2}}(Q_T)$ and $\mu_1 = \min\{1, \frac{\nu}{C_F^2}\} / \sqrt{5}$.

Proof. The a posteriori error estimate immediately follows from (6.39) and (6.40). \square

Now, let us discuss again a posteriori error estimates for the Fourier coefficients using the multiharmonic approximation and the \mathcal{P} -norm (3.33) introduced in Section 3.3, i.e.,

$$\|\mathbf{u}_k\|_{\mathcal{P}}^2 = (\nu \nabla \mathbf{u}_k, \nabla \mathbf{u}_k)_{L^2(\Omega)} + k\omega (\sigma \mathbf{u}_k, \mathbf{u}_k)_{L^2(\Omega)}.$$

We proved the following inf-sup condition:

$$\sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{a_k(\mathbf{u}_k - \boldsymbol{\eta}_k, \mathbf{v}_k)}{\|\mathbf{v}_k\|_{\mathcal{P}}} \geq \underline{c} \|\mathbf{u}_k - \boldsymbol{\eta}_k\|_{\mathcal{P}} \tag{6.42}$$

with the parameter-independent constant $\underline{c} = 1/\sqrt{2}$. Hence, we want to find an upper bound of

$$\sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{\mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k)}{\|\mathbf{v}_k\|_{\mathcal{P}}}.$$

Besides introducing the functions $\boldsymbol{\tau}_k = (\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s)^T$ of vector-valued functions

$$\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s \in H(\operatorname{div}, \Omega) := \{\boldsymbol{\tau} \in [L^2(\Omega)]^d : \operatorname{div} \boldsymbol{\tau} \in L^2(\Omega)\},$$

where the weak divergence fulfills the identity

$$\int_{\Omega} \operatorname{div} \boldsymbol{\tau} v \, dx = - \int_{\Omega} \boldsymbol{\tau} \cdot \nabla v \, dx \quad \forall v \in C_0^\infty(\Omega),$$

we introduce the functions

$$\boldsymbol{\kappa}_k = (\boldsymbol{\kappa}_k^c, \boldsymbol{\kappa}_k^s)^T \in (H_0^1(\Omega))^2$$

fulfilling the identity

$$\int_{\Omega} k\omega \sigma(\mathbf{x}) \boldsymbol{\kappa}_k \cdot \mathbf{v}^\perp d\mathbf{x} = - \int_{\Omega} k\omega \sigma(\mathbf{x}) \boldsymbol{\kappa}_k^\perp \cdot \mathbf{v} d\mathbf{x} \quad \forall \mathbf{v} \in (C_0^\infty(\Omega))^2,$$

which is a simple orthogonality relation. Due to the Cauchy-Schwarz and Friedrichs inequalities, we obtain

$$\begin{aligned} \mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k) &= \int_{\Omega} (\mathbf{f}_k \cdot \mathbf{v}_k - k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k \cdot \mathbf{v}_k^\perp + \mathbf{div} \boldsymbol{\tau}_k \cdot \mathbf{v}_k + \boldsymbol{\tau}_k \cdot \nabla \mathbf{v}_k + k\omega \sigma(\mathbf{x}) \boldsymbol{\kappa}_k \cdot \mathbf{v}_k^\perp \\ &\quad + k\omega \sigma(\mathbf{x}) \boldsymbol{\kappa}_k^\perp \cdot \mathbf{v}_k - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k \cdot \nabla \mathbf{v}_k) d\mathbf{x} \\ &= \int_{\Omega} (\mathbf{f}_k \cdot \mathbf{v}_k + k\omega \sigma(\mathbf{x}) \boldsymbol{\kappa}_k^\perp \cdot \mathbf{v}_k + \mathbf{div} \boldsymbol{\tau}_k \cdot \mathbf{v}_k + (\boldsymbol{\tau}_k - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k) \cdot \nabla \mathbf{v}_k \\ &\quad - k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k \cdot \mathbf{v}_k^\perp + k\omega \sigma(\mathbf{x}) \boldsymbol{\kappa}_k \cdot \mathbf{v}_k^\perp) d\mathbf{x} \\ &\leq \|\mathcal{R}_{1k}(\boldsymbol{\kappa}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} \|\nabla \mathbf{v}_k\|_{L^2(\Omega)} \\ &\quad + \sqrt{k\omega} (\sigma \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k), \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k))_{L^2(\Omega)}^{1/2} (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)}^{1/2} \\ &\leq (C_F \|\mathcal{R}_{1k}(\boldsymbol{\kappa}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}) \|\nabla \mathbf{v}_k\|_{L^2(\Omega)} \\ &\quad + \sqrt{k\omega} (\sigma \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k), \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k))_{L^2(\Omega)}^{1/2} (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)}^{1/2} \\ &\leq \left((C_F \|\mathcal{R}_{1k}(\boldsymbol{\kappa}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)})^2 \right. \\ &\quad \left. + (\sigma \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k), \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k))_{L^2(\Omega)} \right)^{1/2} \left(\|\nabla \mathbf{v}_k\|_{L^2(\Omega)}^2 + k\omega (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)} \right)^{1/2} \end{aligned}$$

with

$$\begin{aligned} \mathcal{R}_{1k}(\boldsymbol{\kappa}_k, \boldsymbol{\tau}_k) &= k\omega \sigma \boldsymbol{\kappa}_k^\perp + \mathbf{div} \boldsymbol{\tau}_k + \mathbf{f}_k = (-k\omega \sigma \boldsymbol{\kappa}_k^s + \mathbf{div} \boldsymbol{\tau}_k^c + \mathbf{f}_k^c, k\omega \sigma \boldsymbol{\kappa}_k^c + \mathbf{div} \boldsymbol{\tau}_k^s + \mathbf{f}_k^s)^T \\ &= (\mathcal{R}_{1k}^c(\boldsymbol{\kappa}_k^s, \boldsymbol{\tau}_k^c), \mathcal{R}_{1k}^s(\boldsymbol{\kappa}_k^c, \boldsymbol{\tau}_k^s))^T, \end{aligned}$$

$$\begin{aligned} \mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k) &= \boldsymbol{\tau}_k - \nu \nabla \boldsymbol{\eta}_k = (\boldsymbol{\tau}_k^c - \nu \nabla \boldsymbol{\eta}_k^c, \boldsymbol{\tau}_k^s - \nu \nabla \boldsymbol{\eta}_k^s)^T \\ &= (\mathcal{R}_{2k}^c(\boldsymbol{\eta}_k^c, \boldsymbol{\tau}_k^c), \mathcal{R}_{2k}^s(\boldsymbol{\eta}_k^s, \boldsymbol{\tau}_k^s))^T \end{aligned}$$

and

$$\begin{aligned} \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k) &= \boldsymbol{\kappa}_k - \boldsymbol{\eta}_k = (\boldsymbol{\kappa}_k^c - \boldsymbol{\eta}_k^c, \boldsymbol{\kappa}_k^s - \boldsymbol{\eta}_k^s)^T \\ &= (\mathcal{R}_{3k}^c(\boldsymbol{\eta}_k^c, \boldsymbol{\kappa}_k^c), \mathcal{R}_{3k}^s(\boldsymbol{\eta}_k^s, \boldsymbol{\kappa}_k^s))^T. \end{aligned}$$

Now, we obtain the following upper bound using the notation $\|\cdot\|_{L^2} = \|\cdot\|_{L^2(\Omega)}$:

$$\begin{aligned} &\sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{\mathcal{F}_{\boldsymbol{\eta}_k}(\mathbf{v}_k)}{\|\mathbf{v}_k\|_{\mathcal{P}}} \\ &\leq \sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{\left((C_F \|\mathcal{R}_{1k}\|_{L^2} + \|\mathcal{R}_{2k}\|_{L^2})^2 + (\sigma \mathcal{R}_{3k}, \mathcal{R}_{3k})_{L^2} \right)^{1/2} \left(\|\nabla \mathbf{v}_k\|_{L^2}^2 + k\omega (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2} \right)^{1/2}}{\left((\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2} + k\omega (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2} \right)^{1/2}} \\ &\leq \sup_{0 \neq \mathbf{v}_k \in (H_0^1(\Omega))^2} \frac{\left((C_F \|\mathcal{R}_{1k}\|_{L^2} + \|\mathcal{R}_{2k}\|_{L^2})^2 + (\sigma \mathcal{R}_{3k}, \mathcal{R}_{3k})_{L^2} \right)^{1/2} \left(\|\nabla \mathbf{v}_k\|_{L^2}^2 + k\omega (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2} \right)^{1/2}}{\left(\nu \|\nabla \mathbf{v}_k\|_{L^2} + k\omega (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2} \right)^{1/2}} \\ &\leq \frac{1}{\min\{\sqrt{\nu}, 1\}} \left((C_F \|\mathcal{R}_{1k}\|_{L^2} + \|\mathcal{R}_{2k}\|_{L^2})^2 + (\sigma \mathcal{R}_{3k}, \mathcal{R}_{3k})_{L^2} \right)^{1/2}. \end{aligned} \tag{6.43}$$

Hence, we have derived another result as in Corollary 6.7 and Corollary 6.8 for every mode $k = 1, \dots, N$ but now with the \mathcal{P} -norm. Using the estimate (6.43) together with the inf-sup condition (6.42), we finally arrive at the following upper bounds for every single mode $k = 1, \dots, N$:

Corollary 6.20. *Let $\boldsymbol{\eta}_k \in H_0^1(\Omega)$ and the bilinear form $a_k(\cdot, \cdot)$ defined by (6.17) satisfy (6.42). Then, it follows together with (6.43) that*

$$\begin{aligned} \|\mathbf{u}_k - \boldsymbol{\eta}_k\|_{\mathcal{P}} &\leq \frac{\sqrt{2}}{\min\{\sqrt{\mathcal{L}}, 1\}} \left((C_F \|\mathcal{R}_{1k}(\boldsymbol{\kappa}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)} + \|\mathcal{R}_{2k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)})^2 \right. \\ &\quad \left. + (\sigma \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k), \mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\kappa}_k))_{L^2(\Omega)} \right)^{1/2} \\ &=: \mathcal{M}_{\|\cdot\|_{\mathcal{P}}}^{\oplus k}(\boldsymbol{\eta}_k, \boldsymbol{\tau}_k, \boldsymbol{\kappa}_k), \end{aligned} \quad (6.44)$$

where $\boldsymbol{\tau}_k = (\boldsymbol{\tau}_k^c, \boldsymbol{\tau}_k^s)^T \in (H(\operatorname{div}, \Omega))^2$ and $\boldsymbol{\kappa}_k = (\boldsymbol{\kappa}_k^c, \boldsymbol{\kappa}_k^s)^T \in (H_0^1(\Omega))^2$.

In the case $k = 0$, we obtain the same result as presented in Corollary 6.10.

Remark 6.21. *The construction of $\boldsymbol{\eta}$ and $\boldsymbol{\tau}$ is an important issue in order to obtain final bounds from the majorants \mathcal{M}_*^{\oplus} (replacing $*$ with the different seminorms and norms) in practice. As already discussed, we can choose for $\boldsymbol{\eta}$ a multiharmonic finite element approximation, solve the discretized problem and then reconstruct the flux. A good reconstruction of the flux is an important and non-trivial topic. For instance, by choosing piecewise linear finite element approximations in space, their gradients are only piecewise constant and so do not belong to $H(\operatorname{div}, Q_T)$. Hence, it is important to regularize $\boldsymbol{\tau}$ by a post-processing operator which maps the L^2 -functions into $H(\operatorname{div}, Q_T)$, see [153]. There are various techniques for realizing these post-processing steps such as, e.g., local post-processing by an elementwise averaging procedure or by using Raviart-Thomas elements, see, e.g., [148, 168] and [153, 123].*

Remark 6.22. *Another very important topic is the construction of a so-called adaptive multiharmonic finite element method (AMhFEM). In addition to constructing an adaptive finite element method (AFEM), we can compute the finite element approximated Fourier coefficients parallel on different meshes, since the computations of the Fourier coefficients corresponding to every single mode $k = 0, 1 \dots$ are decoupled. Then, by prescribing certain bounds, we can filter out the Fourier coefficients, which are important for the (numerical) solution of the problem. Altogether, such an AMhFEM yields adaptivity in space and time.*

6.2 Functional a posteriori error estimates for parabolic time-periodic optimal control problems

As starting point let us consider the variational problem (4.7) of the reduced optimality system: Given the desired state $y_d \in L^2(Q_T)$, find y and p from $H_0^{1, \frac{1}{2}}(Q_T)$ such that

$$\begin{aligned} \int_0^T \int_{\Omega} \left(yv - \nu(\mathbf{x}) \nabla p \cdot \nabla v + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} v^\perp \right) d\mathbf{x} dt &= \int_0^T \int_{\Omega} y_d v d\mathbf{x} dt, \\ \int_0^T \int_{\Omega} \left(\nu(\mathbf{x}) \nabla y \cdot \nabla q + \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} q^\perp + \lambda^{-1} p q \right) d\mathbf{x} dt &= 0, \end{aligned} \quad (6.45)$$

for all test functions $v, q \in H_0^{1, \frac{1}{2}}(Q_T)$, where all functions are given in their Fourier series expansion in time according to Definition 3.2. Let us define the space-time bilinear form

$$\begin{aligned} \mathcal{B}((y, p), (v, q)) &= \int_0^T \int_{\Omega} \left(yv - \nu(\mathbf{x}) \nabla p \cdot \nabla v + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} v^\perp \right. \\ &\quad \left. + \nu(\mathbf{x}) \nabla y \cdot \nabla q + \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} q^\perp + \lambda^{-1} p q \right) d\mathbf{x} dt. \end{aligned} \quad (6.46)$$

This space-time bilinear form is analogously defined as the bilinear form (4.23) for every single mode k . Let (η, ζ) be an approximation of (y, p) , e.g., the multiharmonic finite element approximations $\eta = y_{Nh}$ and $\zeta = p_{Nh}$, where $u_{Nh} = -\lambda^{-1}p_{Nh}$, of the state and the adjoint state, respectively.

A first a posteriori error result

First, we assume again that η and ζ are a bit more regular than the state y and the adjoint state p with respect to the time variable, i.e., $\eta, \zeta \in H_{0,per}^{1,1}(Q_T)$, that is clearly true for the multiharmonic finite element approximations. Our goal is to deduce a computable upper bound of the error

$$e := (y, p) - (\eta, \zeta) = (y - \eta, p - \zeta) \in \left(H_0^{1, \frac{1}{2}}(Q_T)\right)^2.$$

From (6.45) it follows that

$$\begin{aligned} & \int_0^T \int_{\Omega} \left((y - \eta) v - \nu(\mathbf{x}) \nabla(p - \zeta) \cdot \nabla v + \sigma(\mathbf{x}) \partial_t^{1/2}(p - \zeta) \partial_t^{1/2} v^\perp \right. \\ & \quad \left. + \nu(\mathbf{x}) \nabla(y - \eta) \cdot \nabla q + \sigma(\mathbf{x}) \partial_t^{1/2}(y - \eta) \partial_t^{1/2} q^\perp + \lambda^{-1}(p - \zeta) q \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(y_d v - \eta v + \nu(\mathbf{x}) \nabla \zeta \cdot \nabla v - \sigma(\mathbf{x}) \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp \right. \\ & \quad \left. - \nu(\mathbf{x}) \nabla \eta \cdot \nabla q - \sigma(\mathbf{x}) \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \end{aligned} \quad (6.47)$$

for all $v, q \in H_0^{1, \frac{1}{2}}(Q_T)$. The left hand side of (6.47) is obviously

$$\mathcal{B}((y - \eta, p - \zeta), (v, q)).$$

Let us now prove inf-sup and sup-sup conditions for the bilinear form $\mathcal{B}(\cdot, \cdot)$. First, we define the following $H^{1, \frac{1}{2}}$ -norm for $(y, p) \in H_0^{1, \frac{1}{2}}(Q_T) \times H_0^{1, \frac{1}{2}}(Q_T)$:

$$\|(y, p)\|_{H^{1, \frac{1}{2}}(Q_T)} = \left(\|y\|_{H^{1, \frac{1}{2}}(Q_T)}^2 + \|p\|_{H^{1, \frac{1}{2}}(Q_T)}^2 \right)^{1/2}.$$

Lemma 6.23. *The space-time bilinear form $\mathcal{B}(\cdot, \cdot)$ defined by (6.46) fulfills the following inf-sup and sup-sup conditions:*

$$\mu_1 \|(y, p)\|_{H^{1, \frac{1}{2}}(Q_T)} \leq \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}((y, p), (v, q))}{\|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)}} \leq \mu_2 \|(y, p)\|_{H^{1, \frac{1}{2}}(Q_T)} \quad (6.48)$$

for all $y, p \in H_0^{1, \frac{1}{2}}(Q_T)$ with the positive constants $\mu_1 = \min\{1, \frac{1}{\lambda}, \frac{\nu}{\sqrt{\lambda}}, \underline{\nu}\sqrt{\lambda}, \frac{\sigma}{\sqrt{\lambda}}, \underline{\sigma}\sqrt{\lambda}\}$ and $\mu_2 = \max\{1, \bar{\nu}, \bar{\sigma}, \lambda^{-1}\}$.

Proof. We start with the proof of the sup-sup condition. Using (2.29), i.e., the boundedness of the coefficients σ and ν , as well as the triangle and the Cauchy-Schwarz inequalities (3.14) and (3.15),

we obtain the estimate

$$\begin{aligned}
|\mathcal{B}((y, p), (v, q))| &= \left| \int_0^T \int_{\Omega} \left(yv - \nu(\mathbf{x}) \nabla p \cdot \nabla v + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} v^\perp \right. \right. \\
&\quad \left. \left. + \nu(\mathbf{x}) \nabla y \cdot \nabla q + \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} q^\perp + \lambda^{-1} p q \right) d\mathbf{x} dt \right| \\
&\leq \left| \int_0^T \int_{\Omega} yv d\mathbf{x} dt \right| + \left| \int_0^T \int_{\Omega} \nu(\mathbf{x}) \nabla p \cdot \nabla v d\mathbf{x} dt \right| + \left| \int_0^T \int_{\Omega} \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} v^\perp d\mathbf{x} dt \right| \\
&\quad + \left| \int_0^T \int_{\Omega} \nu(\mathbf{x}) \nabla y \cdot \nabla q d\mathbf{x} dt \right| + \left| \int_0^T \int_{\Omega} \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} q^\perp d\mathbf{x} dt \right| + \lambda^{-1} \left| \int_0^T \int_{\Omega} p q d\mathbf{x} dt \right| \\
&\leq \int_0^T \int_{\Omega} |yv| d\mathbf{x} dt + \bar{\nu} \int_0^T \int_{\Omega} |\nabla p| |\nabla v| d\mathbf{x} dt + \bar{\sigma} \int_0^T \int_{\Omega} |\partial_t^{1/2} p| |\partial_t^{1/2} v^\perp| d\mathbf{x} dt \\
&\quad + \bar{\nu} \int_0^T \int_{\Omega} |\nabla y| |\nabla q| d\mathbf{x} dt + \bar{\sigma} \int_0^T \int_{\Omega} |\partial_t^{1/2} y| |\partial_t^{1/2} q^\perp| d\mathbf{x} dt + \lambda^{-1} \int_0^T \int_{\Omega} |p q| d\mathbf{x} dt \\
&\leq \|y\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \bar{\nu} \|\nabla p\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\
&\quad + \bar{\sigma} \|\partial_t^{1/2} p\|_{L^2(Q_T)} \|\partial_t^{1/2} v^\perp\|_{L^2(Q_T)} + \bar{\nu} \|\nabla y\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} \\
&\quad + \bar{\sigma} \|\partial_t^{1/2} y\|_{L^2(Q_T)} \|\partial_t^{1/2} q^\perp\|_{L^2(Q_T)} + \lambda^{-1} \|p\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)}.
\end{aligned}$$

Since

$$\|\partial_t^{1/2} v^\perp\|_{L^2(Q_T)}^2 = \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{v}_k^\perp\|_{L^2(\Omega)}^2 = \frac{T}{2} \sum_{k=1}^{\infty} k\omega \|\mathbf{v}_k\|_{L^2(\Omega)}^2 = \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2,$$

we have that

$$\begin{aligned}
|\mathcal{B}((y, p), (v, q))| &\leq \|y\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \bar{\nu} \|\nabla p\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\
&\quad + \bar{\sigma} \|\partial_t^{1/2} p\|_{L^2(Q_T)} \|\partial_t^{1/2} v^\perp\|_{L^2(Q_T)} + \bar{\nu} \|\nabla y\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} \\
&\quad + \bar{\sigma} \|\partial_t^{1/2} y\|_{L^2(Q_T)} \|\partial_t^{1/2} q^\perp\|_{L^2(Q_T)} + \lambda^{-1} \|p\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)} \\
&= \|y\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \bar{\nu} \|\nabla p\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} + \bar{\sigma} \|\partial_t^{1/2} p\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)} \\
&\quad + \bar{\nu} \|\nabla y\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} + \bar{\sigma} \|\partial_t^{1/2} y\|_{L^2(Q_T)} \|\partial_t^{1/2} q\|_{L^2(Q_T)} + \lambda^{-1} \|p\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)} \\
&\leq \mu_2 \left(\|y\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\nabla p\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} + \|\partial_t^{1/2} p\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)} \right. \\
&\quad \left. + \|\nabla y\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} + \|\partial_t^{1/2} y\|_{L^2(Q_T)} \|\partial_t^{1/2} q\|_{L^2(Q_T)} + \|p\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)} \right) \\
&\leq \mu_2 \left(\|y\|_{L^2(Q_T)}^2 + \|\nabla p\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} p\|_{L^2(Q_T)}^2 + \|\nabla y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 + \|p\|_{L^2(Q_T)}^2 \right)^{1/2} \\
&\quad \times \left(\|v\|_{L^2(Q_T)}^2 + \|\nabla v\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2 + \|\nabla q\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} q\|_{L^2(Q_T)}^2 + \|q\|_{L^2(Q_T)}^2 \right)^{1/2} \\
&= \mu_2 \left(\|y\|_{L^2(Q_T)}^2 + \|\nabla y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 + \|p\|_{L^2(Q_T)}^2 + \|\nabla p\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} p\|_{L^2(Q_T)}^2 \right)^{1/2} \\
&\quad \times \left(\|v\|_{L^2(Q_T)}^2 + \|\nabla v\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2 + \|q\|_{L^2(Q_T)}^2 + \|\nabla q\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} q\|_{L^2(Q_T)}^2 \right)^{1/2} \\
&= \mu_2 \left(\|y\|_{H^{1, \frac{1}{2}}(Q_T)}^2 + \|p\|_{H^{1, \frac{1}{2}}(Q_T)}^2 \right)^{1/2} \left(\|v\|_{H^{1, \frac{1}{2}}(Q_T)}^2 + \|q\|_{H^{1, \frac{1}{2}}(Q_T)}^2 \right)^{1/2} \\
&= \mu_2 \|(y, p)\|_{H^{1, \frac{1}{2}}(Q_T)} \|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)}
\end{aligned}$$

with $\mu_2 = \max\{1, \bar{\nu}, \bar{\sigma}, \lambda^{-1}\}$. Next, we prove the inf-sup condition. By choosing the test function

$$(v, q) = \left(y - \frac{1}{\sqrt{\lambda}}p - \frac{1}{\sqrt{\lambda}}p^\perp, p + \sqrt{\lambda}y - \sqrt{\lambda}y^\perp\right)$$

and using the σ - and ν -weighted orthogonality relations (3.13), we get the following relations:

$$\begin{aligned} \mathcal{B}((y, p), (y, p)) &= \int_0^T \int_\Omega \left(y y - \nu(\mathbf{x}) \nabla p \cdot \nabla y + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} y^\perp \right. \\ &\quad \left. + \nu(\mathbf{x}) \nabla y \cdot \nabla p + \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} p^\perp + \lambda^{-1} p p \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(y y + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} y^\perp - \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} y^\perp + \lambda^{-1} p p \right) d\mathbf{x} dt \\ &= \|y\|_{L^2(Q_T)}^2 + \lambda^{-1} \|p\|_{L^2(Q_T)}^2, \end{aligned}$$

$$\begin{aligned} \mathcal{B}((y, p), (-\frac{1}{\sqrt{\lambda}}p, \sqrt{\lambda}y)) &= \int_0^T \int_\Omega \left(y(-\frac{1}{\sqrt{\lambda}}p) - \nu(\mathbf{x}) \nabla p \cdot \nabla(-\frac{1}{\sqrt{\lambda}}p) + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2}(-\frac{1}{\sqrt{\lambda}}p)^\perp \right. \\ &\quad \left. + \nu(\mathbf{x}) \nabla y \cdot \nabla(\sqrt{\lambda}y) + \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2}(\sqrt{\lambda}y)^\perp + \lambda^{-1} p(\sqrt{\lambda}y) \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(\frac{1}{\sqrt{\lambda}} \nu(\mathbf{x}) \nabla p \cdot \nabla p - \frac{1}{\sqrt{\lambda}} \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} p^\perp \right. \\ &\quad \left. + \sqrt{\lambda} \nu(\mathbf{x}) \nabla y \cdot \nabla y + \sqrt{\lambda} \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} y^\perp \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(\frac{1}{\sqrt{\lambda}} \nu(\mathbf{x}) \nabla p \cdot \nabla p + \sqrt{\lambda} \nu(\mathbf{x}) \nabla y \cdot \nabla y \right) d\mathbf{x} dt \\ &= \frac{1}{\sqrt{\lambda}} (\nu \nabla p, \nabla p)_{L^2(Q_T)} + \sqrt{\lambda} (\nu \nabla y, \nabla y)_{L^2(Q_T)} \end{aligned}$$

and

$$\begin{aligned} \mathcal{B}((y, p), (-\frac{1}{\sqrt{\lambda}}p^\perp, -\sqrt{\lambda}y^\perp)) &= \int_0^T \int_\Omega \left(y(-\frac{1}{\sqrt{\lambda}}p^\perp) - \nu(\mathbf{x}) \nabla p \cdot \nabla(-\frac{1}{\sqrt{\lambda}}p^\perp) \right. \\ &\quad \left. + \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2}(-\frac{1}{\sqrt{\lambda}}p^\perp)^\perp + \nu(\mathbf{x}) \nabla y \cdot \nabla(-\sqrt{\lambda}y^\perp) \right. \\ &\quad \left. + \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2}(-\sqrt{\lambda}y^\perp)^\perp + \lambda^{-1} p(-\sqrt{\lambda}y^\perp) \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(-\frac{1}{\sqrt{\lambda}} y p^\perp + \frac{1}{\sqrt{\lambda}} \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2}(-p^\perp)^\perp + \sqrt{\lambda} \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2}(-y^\perp)^\perp - \frac{1}{\sqrt{\lambda}} p y^\perp \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(-\frac{1}{\sqrt{\lambda}} y p^\perp + \frac{1}{\sqrt{\lambda}} \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} p + \sqrt{\lambda} \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} y + \frac{1}{\sqrt{\lambda}} y p^\perp \right) d\mathbf{x} dt \\ &= \int_0^T \int_\Omega \left(\frac{1}{\sqrt{\lambda}} \sigma(\mathbf{x}) \partial_t^{1/2} p \partial_t^{1/2} p + \sqrt{\lambda} \sigma(\mathbf{x}) \partial_t^{1/2} y \partial_t^{1/2} y \right) d\mathbf{x} dt \\ &= \frac{1}{\sqrt{\lambda}} (\sigma \partial_t^{1/2} p, \partial_t^{1/2} p)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)}. \end{aligned}$$

From these relations, we easily obtain the estimates

$$\begin{aligned} \mathcal{B}((y, p), (y, p)) &= \|y\|_{L^2(Q_T)}^2 + \lambda^{-1} \|p\|_{L^2(Q_T)}^2 \\ &\geq \min\{1, \lambda^{-1}\} \left(\|y\|_{L^2(Q_T)}^2 + \|p\|_{L^2(Q_T)}^2 \right) \\ &= \min\{1, \lambda^{-1}\} \|(y, p)\|_{L^2(Q_T)}^2, \end{aligned}$$

$$\begin{aligned}
 \mathcal{B}((y, p), (-\frac{1}{\sqrt{\lambda}}p, \sqrt{\lambda}y)) &= \frac{1}{\sqrt{\lambda}}(\nu \nabla p, \nabla p)_{L^2(Q_T)} + \sqrt{\lambda}(\nu \nabla y, \nabla y)_{L^2(Q_T)} \\
 &\geq \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \underline{\nu} \left(\|\nabla y\|_{L^2(Q_T)}^2 + \|\nabla p\|_{L^2(Q_T)}^2 \right) \\
 &= \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \underline{\nu} \|(\nabla y, \nabla p)\|_{L^2(Q_T)}^2
 \end{aligned}$$

and

$$\begin{aligned}
 \mathcal{B}((y, p), (-\frac{1}{\sqrt{\lambda}}p^\perp, -\sqrt{\lambda}y^\perp)) &= \frac{1}{\sqrt{\lambda}}(\sigma \partial_t^{1/2} p, \partial_t^{1/2} p)_{L^2(Q_T)} + \sqrt{\lambda}(\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)} \\
 &\geq \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \underline{\sigma} \left(\|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} p\|_{L^2(Q_T)}^2 \right) \\
 &= \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \underline{\sigma} \|(\partial_t^{1/2} y, \partial_t^{1/2} p)\|_{L^2(Q_T)}^2.
 \end{aligned}$$

Altogether, we have the following estimate from below:

$$\begin{aligned}
 \mathcal{B}((y, p), (y - \frac{1}{\sqrt{\lambda}}p - \frac{1}{\sqrt{\lambda}}p^\perp, p + \sqrt{\lambda}y - \sqrt{\lambda}y^\perp)) \\
 \geq \min\{1, \frac{1}{\lambda}, \frac{\underline{\nu}}{\sqrt{\lambda}}, \underline{\nu}\sqrt{\lambda}, \frac{\underline{\sigma}}{\sqrt{\lambda}}, \underline{\sigma}\sqrt{\lambda}\} \left(\|(y, p)\|_{L^2(Q_T)}^2 + \|(\nabla y, \nabla p)\|_{L^2(Q_T)}^2 + \|(\partial_t^{1/2} y, \partial_t^{1/2} p)\|_{L^2(Q_T)}^2 \right) \\
 = \mu_1 \|(y, p)\|_{H^{1, \frac{1}{2}}(Q_T)}^2
 \end{aligned}$$

with the constant

$$\mu_1 = \min\{1, \frac{1}{\lambda}, \frac{\underline{\nu}}{\sqrt{\lambda}}, \underline{\nu}\sqrt{\lambda}, \frac{\underline{\sigma}}{\sqrt{\lambda}}, \underline{\sigma}\sqrt{\lambda}\}.$$

□

We now know from Lemma 6.23 that

$$\sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}((y - \eta, p - \zeta), (v, q))}{\|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)}} \geq \mu_1 \|(y - \eta, p - \zeta)\|_{H^{1, \frac{1}{2}}(Q_T)} \quad (6.49)$$

with a positive constant μ_1 . For completeness, we similarly prove inf-sup and sup-sup conditions with the $H^{1, \frac{1}{2}}$ -seminorm, which is, in fact, an equivalent norm due to the Friedrichs inequality.

Lemma 6.24. *The space-time bilinear form $\mathcal{B}(\cdot, \cdot)$ defined by (6.46) fulfills the following inf-sup and sup-sup conditions:*

$$\mu_1 |(y, p)|_{H^{1, \frac{1}{2}}(Q_T)} \leq \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}((y, p), (v, q))}{|(v, q)|_{H^{1, \frac{1}{2}}(Q_T)}} \leq \mu_2 |(y, p)|_{H^{1, \frac{1}{2}}(Q_T)} \quad (6.50)$$

for all $y, p \in H_0^{1, \frac{1}{2}}(Q_T)$ with positive constants

$$\mu_1 = \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \min\{\underline{\nu}, \underline{\sigma}\} \quad \text{and} \quad \mu_2 = \max\{1, \bar{\nu}, \bar{\sigma}, \lambda^{-1}\} \max\{1, 1 + C_F^2\},$$

where C_F is the Friedrichs constant as in Theorem 2.5.

Proof. We prove the sup-sup condition by using the assumptions (2.29) as well as the triangle and the Cauchy-Schwarz inequalities (3.14) and (3.15). As in the proof of Lemma 6.23, we obtain the estimate

$$\begin{aligned} |\mathcal{B}((y, p), (v, q))| &\leq \max\{1, \bar{\nu}, \bar{\sigma}, \lambda^{-1}\} \left(\|y\|_{L^2(Q_T)}^2 + \|\nabla y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|p\|_{L^2(Q_T)}^2 + \|\nabla p\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} p\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &\quad \times \left(\|v\|_{L^2(Q_T)}^2 + \|\nabla v\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|q\|_{L^2(Q_T)}^2 + \|\nabla q\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} q\|_{L^2(Q_T)}^2 \right)^{1/2}. \end{aligned}$$

Using Friedrichs inequality (2.17), see Theorem 2.5, which can be written in the Fourier space by (3.19), we get

$$\begin{aligned} |\mathcal{B}((y, p), (v, q))| &\leq \max\{1, \bar{\nu}, \bar{\sigma}, \lambda^{-1}\} \left(C_F^2 \|\nabla y\|_{L^2(Q_T)}^2 + \|\nabla y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + C_F^2 \|\nabla p\|_{L^2(Q_T)}^2 + \|\nabla p\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} p\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &\quad \times \left(C_F^2 \|\nabla v\|_{L^2(Q_T)}^2 + \|\nabla v\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + C_F^2 \|\nabla q\|_{L^2(Q_T)}^2 + \|\nabla q\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} q\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &= \max\{1, \bar{\nu}, \bar{\sigma}, \lambda^{-1}\} \max\{1, 1 + C_F^2\} \left(\|\nabla y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\nabla p\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} p\|_{L^2(Q_T)}^2 \right)^{1/2} \left(\|\nabla v\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\nabla q\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} q\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &\leq \mu_2 \left(|y|_{H^{1, \frac{1}{2}}(Q_T)}^2 + |p|_{H^{1, \frac{1}{2}}(Q_T)}^2 \right)^{1/2} \left(|v|_{H^{1, \frac{1}{2}}(Q_T)}^2 + |q|_{H^{1, \frac{1}{2}}(Q_T)}^2 \right)^{1/2} \\ &= \mu_2 |(y, p)|_{H^{1, \frac{1}{2}}} |(v, q)|_{H^{1, \frac{1}{2}}} \end{aligned}$$

with the constant $\mu_2 = \max\{1, \bar{\nu}, \bar{\sigma}, \lambda^{-1}\} \max\{1, 1 + C_F^2\}$. Next, we prove the inf-sup condition by choosing the test function

$$(v, q) = \left(-\frac{1}{\sqrt{\lambda}} p - \frac{1}{\sqrt{\lambda}} p^\perp, \sqrt{\lambda} y - \sqrt{\lambda} y^\perp \right),$$

and using the σ - and ν -weighted orthogonality relations (3.13). We get the same relations and estimates as in the proof of Lemma 6.23, i.e.,

$$\begin{aligned} \mathcal{B}((y, p), \left(-\frac{1}{\sqrt{\lambda}} p, \sqrt{\lambda} y \right)) &= \frac{1}{\sqrt{\lambda}} (\nu \nabla p, \nabla p)_{L^2(Q_T)} + \sqrt{\lambda} (\nu \nabla y, \nabla y)_{L^2(Q_T)} \\ &\geq \min\left\{ \frac{1}{\sqrt{\lambda}}, \sqrt{\lambda} \right\} \underline{\nu} \left(\|\nabla y\|_{L^2(Q_T)}^2 + \|\nabla p\|_{L^2(Q_T)}^2 \right) \\ &= \min\left\{ \frac{1}{\sqrt{\lambda}}, \sqrt{\lambda} \right\} \underline{\nu} \|\nabla(y, p)\|_{L^2(Q_T)}^2 \end{aligned}$$

and

$$\begin{aligned} \mathcal{B}((y, p), \left(-\frac{1}{\sqrt{\lambda}} p^\perp, -\sqrt{\lambda} y^\perp \right)) &= \frac{1}{\sqrt{\lambda}} (\sigma \partial_t^{1/2} p, \partial_t^{1/2} p)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)} \\ &\geq \min\left\{ \frac{1}{\sqrt{\lambda}}, \sqrt{\lambda} \right\} \underline{\sigma} \left(\|\partial_t^{1/2} y\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} p\|_{L^2(Q_T)}^2 \right) \\ &= \min\left\{ \frac{1}{\sqrt{\lambda}}, \sqrt{\lambda} \right\} \underline{\sigma} \|\partial_t^{1/2}(y, p)\|_{L^2(Q_T)}^2. \end{aligned}$$

Altogether, we have the following estimate from below:

$$\begin{aligned} & \mathcal{B}((y, p), (-\frac{1}{\sqrt{\lambda}}p - \frac{1}{\sqrt{\lambda}}p^\perp, \sqrt{\lambda}y - \sqrt{\lambda}y^\perp)) \\ & \geq \min\{\frac{\underline{\nu}}{\sqrt{\lambda}}, \underline{\nu}\sqrt{\lambda}, \frac{\underline{\sigma}}{\sqrt{\lambda}}, \underline{\sigma}\sqrt{\lambda}\} \left(\|(\nabla y, \nabla p)\|_{L^2(Q_T)}^2 + \|(\partial_t^{1/2}y, \partial_t^{1/2}p)\|_{L^2(Q_T)}^2 \right) \\ & = \mu_1 |(y, p)|_{H^{1, \frac{1}{2}}(Q_T)}^2 \end{aligned}$$

with the constant $\mu_1 = \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \min\{\underline{\nu}, \underline{\sigma}\}$. \square

By using the weighted $H^{1, \frac{1}{2}}$ -norms $\|\cdot\|_{V_0}$, i.e.,

$$\begin{aligned} \|y\|_{V_0} & := \left(\|y\|_{L^2(Q_T)}^2 + \sqrt{\lambda}(\nu \nabla y, \nabla y)_{L^2(Q_T)} + \sqrt{\lambda}(\sigma \partial_t^{1/2}y, \partial_t^{1/2}y)_{L^2(Q_T)} \right)^{1/2}, \\ \|(y, p)\|_{V_0} & := \left(\|y\|_{V_0}^2 + \lambda^{-1}\|p\|_{V_0}^2 \right)^{1/2}, \end{aligned}$$

which were introduced in Chapter 4, we can get rid of all parameters from the inf-sup and sup-sup constants μ_1 and μ_2 . Hence, we obtain the following lemma:

Lemma 6.25. *The space-time bilinear form $\mathcal{B}(\cdot, \cdot)$ defined by (6.46) fulfills the following inf-sup and sup-sup conditions:*

$$\mu_1 \|(y, p)\|_{V_0} \leq \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}((y, p), (v, q))}{\|(v, q)\|_{V_0}} \leq \mu_2 \|(y, p)\|_{V_0} \quad (6.51)$$

for all $y, p \in H_0^{1, \frac{1}{2}}(Q_T)$ with $\mu_1 = 1/\sqrt{3}$ and $\mu_2 = 1$.

Proof. We start with the proof of the sup-sup condition. Using the triangle inequality and the σ - and ν -weighted counterparts of the Cauchy-Schwarz inequalities (3.14) and (3.15), we obtain the estimate

$$\begin{aligned} |\mathcal{B}((y, p), (v, q))| & = \left| \int_0^T \int_\Omega \left(yv - \nu(\mathbf{x}) \nabla p \cdot \nabla v + \sigma(\mathbf{x}) \partial_t^{1/2}p \partial_t^{1/2}v^\perp \right. \right. \\ & \quad \left. \left. + \nu(\mathbf{x}) \nabla y \cdot \nabla q + \sigma(\mathbf{x}) \partial_t^{1/2}y \partial_t^{1/2}q^\perp + \lambda^{-1}pq \right) d\mathbf{x} dt \right| \\ & \leq \left| \int_0^T \int_\Omega yv d\mathbf{x} dt \right| + \left| \int_0^T \int_\Omega \nu(\mathbf{x}) \nabla p \cdot \nabla v d\mathbf{x} dt \right| \\ & \quad + \left| \int_0^T \int_\Omega \sigma(\mathbf{x}) \partial_t^{1/2}p \partial_t^{1/2}v^\perp d\mathbf{x} dt \right| + \left| \int_0^T \int_\Omega \nu(\mathbf{x}) \nabla y \cdot \nabla q d\mathbf{x} dt \right| \\ & \quad + \left| \int_0^T \int_\Omega \sigma(\mathbf{x}) \partial_t^{1/2}y \partial_t^{1/2}q^\perp d\mathbf{x} dt \right| + \left| \int_0^T \int_\Omega \lambda^{-1}pq d\mathbf{x} dt \right| \\ & \leq \|y\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + (\nu \nabla p, \nabla p)_{L^2(Q_T)}^{1/2} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2} \\ & \quad + (\sigma \partial_t^{1/2}p, \partial_t^{1/2}p)_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2}v, \partial_t^{1/2}v)_{L^2(Q_T)}^{1/2} + (\nu \nabla y, \nabla y)_{L^2(Q_T)}^{1/2} (\nu \nabla q, \nabla q)_{L^2(Q_T)}^{1/2} \\ & \quad + (\sigma \partial_t^{1/2}y, \partial_t^{1/2}y)_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2}q, \partial_t^{1/2}q)_{L^2(Q_T)}^{1/2} + \lambda^{-1} \|p\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)}. \end{aligned}$$

With a proper weighting with λ , we finally prove the sup-sup condition as follows

$$\begin{aligned}
|\mathcal{B}((y, p), (v, q))| &\leq \|y\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \lambda^{-1/4} (\nu \nabla p, \nabla p)_{L^2(Q_T)}^{1/2} \lambda^{1/4} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2} \\
&\quad + \lambda^{-1/4} (\sigma \partial_t^{1/2} p, \partial_t^{1/2} p)_{L^2(Q_T)}^{1/2} \lambda^{1/4} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}^{1/2} \\
&\quad + \lambda^{1/4} (\nu \nabla y, \nabla y)_{L^2(Q_T)}^{1/2} \lambda^{-1/4} (\nu \nabla q, \nabla q)_{L^2(Q_T)}^{1/2} \\
&\quad + \lambda^{1/4} (\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)}^{1/2} \lambda^{-1/4} (\sigma \partial_t^{1/2} q, \partial_t^{1/2} q)_{L^2(Q_T)}^{1/2} \\
&\quad + \lambda^{-1/2} \|p\|_{L^2(Q_T)} \lambda^{-1/2} \|q\|_{L^2(Q_T)} \\
&\leq \left(\|y\|_{L^2(Q_T)}^2 + \lambda^{-1/2} (\nu \nabla p, \nabla p)_{L^2(Q_T)} + \lambda^{-1/2} (\sigma \partial_t^{1/2} p, \partial_t^{1/2} p)_{L^2(Q_T)} \right. \\
&\quad \left. + \lambda^{1/2} (\nu \nabla y, \nabla y)_{L^2(Q_T)} + \lambda^{1/2} (\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)} + \lambda^{-1} \|p\|_{L^2(Q_T)}^2 \right)^{1/2} \\
&\quad \times \left(\|v\|_{L^2(Q_T)}^2 + \lambda^{1/2} (\nu \nabla v, \nabla v)_{L^2(Q_T)} + \lambda^{1/2} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)} \right. \\
&\quad \left. + \lambda^{-1/2} (\nu \nabla q, \nabla q)_{L^2(Q_T)} + \lambda^{-1/2} (\sigma \partial_t^{1/2} q, \partial_t^{1/2} q)_{L^2(Q_T)} + \lambda^{-1} \|q\|_{L^2(Q_T)}^2 \right)^{1/2} \\
&= \left(\|y\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\nu \nabla y, \nabla y)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)} \right. \\
&\quad \left. + \lambda^{-1} (\|p\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\nu \nabla p, \nabla p)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} p, \partial_t^{1/2} p)_{L^2(Q_T)}) \right)^{1/2} \\
&\quad \times \left(\|v\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\nu \nabla v, \nabla v)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)} \right. \\
&\quad \left. + \lambda^{-1} (\|q\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\nu \nabla q, \nabla q)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} q, \partial_t^{1/2} q)_{L^2(Q_T)}) \right)^{1/2} \\
&= \left(\|y\|_{\mathbb{V}_0}^2 + \lambda^{-1} \|p\|_{\mathbb{V}_0}^2 \right)^{1/2} \left(\|v\|_{\mathbb{V}_0}^2 + \lambda^{-1} \|q\|_{\mathbb{V}_0}^2 \right)^{1/2} \\
&= \mu_2 \| (y, p) \|_{\mathbb{V}_0} \| (v, q) \|_{\mathbb{V}_0}
\end{aligned}$$

with $\mu_2 = 1$. Next, we prove the inf-sup condition. By choosing the same test function as in the proof of Lemma 6.23, i.e.,

$$(v, q) = \left(y - \frac{1}{\sqrt{\lambda}} p - \frac{1}{\sqrt{\lambda}} p^\perp, p + \sqrt{\lambda} y - \sqrt{\lambda} y^\perp \right)$$

and using the σ - and ν -weighted orthogonality relations (3.13), we get the following relations:

$$\begin{aligned}
\mathcal{B}((y, p), (y, p)) &= \|y\|_{L^2(Q_T)}^2 + \lambda^{-1} \|p\|_{L^2(Q_T)}^2, \\
\mathcal{B}((y, p), (-\frac{1}{\sqrt{\lambda}} p, \sqrt{\lambda} y)) &= \frac{1}{\sqrt{\lambda}} (\nu \nabla p, \nabla p)_{L^2(Q_T)} + \sqrt{\lambda} (\nu \nabla y, \nabla y)_{L^2(Q_T)}, \\
\mathcal{B}((y, p), (-\frac{1}{\sqrt{\lambda}} p^\perp, -\sqrt{\lambda} y^\perp)) &= \frac{1}{\sqrt{\lambda}} (\sigma \partial_t^{1/2} p, \partial_t^{1/2} p)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} y, \partial_t^{1/2} y)_{L^2(Q_T)},
\end{aligned}$$

which were already proven, see Lemma 6.23. Altogether, we obtain

$$\mathcal{B}((y, p), (y - \frac{1}{\sqrt{\lambda}} p - \frac{1}{\sqrt{\lambda}} p^\perp, p + \sqrt{\lambda} y - \sqrt{\lambda} y^\perp)) = \|y\|_{\mathbb{V}_0}^2 + \lambda^{-1} \|p\|_{\mathbb{V}_0}^2 = \|(y, p)\|_{\mathbb{V}_0}^2.$$

Due to

$$\|(v, q)\|_{\mathbb{V}_0} = \sqrt{3} \|(y, p)\|_{\mathbb{V}_0},$$

we arrive at the following estimate of the supremum from below:

$$\begin{aligned} \sup_{0 \neq (v,q) \in (H_0^{1,\frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}((y,p), (v,q))}{\|(v,q)\|_{V_0}} &\geq \frac{\mathcal{B}((y,p), (y - \frac{1}{\sqrt{\lambda}}p - \frac{1}{\sqrt{\lambda}}p^\perp, p + \sqrt{\lambda}y - \sqrt{\lambda}y^\perp))}{\|(y - \frac{1}{\sqrt{\lambda}}p - \frac{1}{\sqrt{\lambda}}p^\perp, p + \sqrt{\lambda}y - \sqrt{\lambda}y^\perp)\|_{V_0}} \\ &= \frac{\|(y,p)\|_{V_0}^2}{\sqrt{3}\|(y,p)\|_{V_0}} = \mu_1 \|(y,p)\|_{V_0}, \end{aligned}$$

which finally yields the inf-sup constant $\mu_1 = 1/\sqrt{3}$. \square

We denote the right-hand side of (6.47) by $\mathcal{F}_{(\eta,\zeta)}(v,q)$. Indeed,

$$\begin{aligned} \mathcal{F}_{(\eta,\zeta)}(v,q) &= \int_0^T \int_\Omega \left(y_d v - \eta v + \nu(\mathbf{x}) \nabla \zeta \cdot \nabla v - \sigma(\mathbf{x}) \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp \right. \\ &\quad \left. - \nu(\mathbf{x}) \nabla \eta \cdot \nabla q - \sigma(\mathbf{x}) \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp - \lambda^{-1} \zeta q \right) dx dt \end{aligned}$$

is a linear functional defined on $v, q \in H_0^{1,\frac{1}{2}}(Q_T)$. We need to find an upper bound of

$$\sup_{0 \neq (v,q) \in (H_0^{1,\frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta,\zeta)}(v,q)}{\|(v,q)\|_{H^{1,\frac{1}{2}}(Q_T)}}, \quad \sup_{0 \neq (v,q) \in (H_0^{1,\frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta,\zeta)}(v,q)}{|(v,q)|_{H^{1,\frac{1}{2}}(Q_T)}} \quad \text{or} \quad (6.52)$$

$$\sup_{0 \neq (v,q) \in (H_0^{1,\frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta,\zeta)}(v,q)}{\|(v,q)\|_{V_0}}, \quad (6.53)$$

where (6.53) will be discussed later. For getting such upper bounds of (6.52), we need to reconstruct $\mathcal{F}_{(\eta,\zeta)}$. First, we note that the σ -weighted identity (6.6), i.e.,

$$(\sigma \partial_t^{1/2} \eta, \partial_t^{1/2} v^\perp)_{L^2(Q_T)} = (\sigma \partial_t \eta, v)_{L^2(Q_T)}$$

is valid for $\eta \in H_{0,per}^{1,1}(Q_T)$ and $v \in H_0^{1,\frac{1}{2}}(Q_T)$ (and, hence, also for $\zeta \in H_{0,per}^{1,1}(Q_T)$ and $q \in H_0^{1,\frac{1}{2}}(Q_T)$), see also (3.10). Let

$$\boldsymbol{\tau}, \boldsymbol{\rho} \in H(\operatorname{div}, Q_T) := \{\boldsymbol{\tau} \in [L^2(Q_T)]^d : \operatorname{div} \boldsymbol{\tau}(\cdot, t) \in L^2(\Omega) \text{ for a.e. } t \in (0, T)\}$$

be two arbitrary vector-valued functions, where $\operatorname{div} = \operatorname{div}_{\mathbf{x}}$ again denotes the weak spatial divergence. Then, using

$$\int_\Omega \operatorname{div} \boldsymbol{\tau} v dx = - \int_\Omega \boldsymbol{\tau} \cdot \nabla v dx \quad \text{and} \quad \int_\Omega \operatorname{div} \boldsymbol{\rho} q dx = - \int_\Omega \boldsymbol{\rho} \cdot \nabla q dx \quad \forall v, q \in C_0^\infty(\Omega),$$

we obtain

$$\begin{aligned} \mathcal{F}_{(\eta,\zeta)}(v,q) &= \int_{Q_T} \left(y_d v - \eta v + \nu(\mathbf{x}) \nabla \zeta \cdot \nabla v - \sigma(\mathbf{x}) \partial_t \zeta v - \nu(\mathbf{x}) \nabla \eta \cdot \nabla q - \sigma(\mathbf{x}) \partial_t \eta q - \lambda^{-1} \zeta q \right) dx dt \\ &= \int_{Q_T} \left(y_d v - \eta v + \nu(\mathbf{x}) \nabla \zeta \cdot \nabla v - \sigma(\mathbf{x}) \partial_t \zeta v + \operatorname{div} \boldsymbol{\tau} v + \boldsymbol{\tau} \cdot \nabla v \right. \\ &\quad \left. - \nu(\mathbf{x}) \nabla \eta \cdot \nabla q - \sigma(\mathbf{x}) \partial_t \eta q - \lambda^{-1} \zeta q + \operatorname{div} \boldsymbol{\rho} q + \boldsymbol{\rho} \cdot \nabla q \right) dx dt \\ &= \int_{Q_T} \left((y_d - \sigma(\mathbf{x}) \partial_t \zeta - \eta + \operatorname{div} \boldsymbol{\tau}) v + (\nu(\mathbf{x}) \nabla \zeta + \boldsymbol{\tau}) \cdot \nabla v \right. \\ &\quad \left. + (-\sigma(\mathbf{x}) \partial_t \eta - \lambda^{-1} \zeta + \operatorname{div} \boldsymbol{\rho}) q + (\boldsymbol{\rho} - \nu(\mathbf{x}) \nabla \eta) \cdot \nabla q \right) dx dt \\ &\leq \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} \end{aligned}$$

with

$$\begin{aligned}\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau}) &= \sigma \partial_t \zeta + \eta - \operatorname{div} \boldsymbol{\tau} - y_d, & \mathcal{R}_2(\zeta, \boldsymbol{\tau}) &= \boldsymbol{\tau} + \nu \nabla \zeta, \\ \mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho}) &= \sigma \partial_t \eta + \lambda^{-1} \zeta - \operatorname{div} \boldsymbol{\rho}, & \mathcal{R}_4(\eta, \boldsymbol{\rho}) &= \boldsymbol{\rho} - \nu \nabla \eta,\end{aligned}$$

where we have applied the Cauchy-Schwarz inequality in the last estimate. The following two theorems provide some upper bounds for the error $e = (y - \eta, p - \zeta)$ in the $H^{1, \frac{1}{2}}$ -seminorm and $H^{1, \frac{1}{2}}$ -norm.

Theorem 6.26. *Let $\eta, \zeta \in H_{0, \text{per}}^{1,1}(Q_T)$ and the bilinear form $\mathcal{B}(\cdot, \cdot)$ defined by (6.46) satisfy the inf-sup condition (6.50). Then,*

$$\begin{aligned}|e|_{H^{1, \frac{1}{2}}(Q_T)} &\leq \frac{1}{\mu_1} (C_F \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \\ &\quad + C_F \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)}) \\ &=: \mathcal{M}_{|\cdot|}^{\oplus}(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}),\end{aligned}\tag{6.54}$$

where $e = (y - \eta, p - \zeta) \in (H_0^{1, \frac{1}{2}}(Q_T))^2$, $\mu_1 = \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \min\{\underline{\nu}, \underline{\sigma}\}$ and $\boldsymbol{\tau}, \boldsymbol{\rho} \in H(\operatorname{div}, Q_T)$.

Proof. Due to the Friedrichs inequality (2.17), see Theorem 2.5, which can be written in the Fourier space by (3.19), we obtain

$$\begin{aligned}\mathcal{F}_{(\eta, \zeta)}(v, q) &\leq \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} \\ &\leq \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} C_F \|\nabla v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)} C_F \|\nabla q\|_{L^2(Q_T)} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} \\ &= \left(C_F \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \right) \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \left(C_F \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)} \right) \|\nabla q\|_{L^2(Q_T)}.\end{aligned}$$

Hence,

$$\begin{aligned}&\sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta, \zeta)}(v, q)}{|(v, q)|_{H^{1, \frac{1}{2}}(Q_T)}} \\ &\leq \sup_{(v, q)} \frac{\left(C_F \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2} \right) \|\nabla v\|_{L^2} + \left(C_F \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2} \right) \|\nabla q\|_{L^2}}{(|v|_{H^{1, \frac{1}{2}}}^2 + |q|_{H^{1, \frac{1}{2}}}^2)^{1/2}} \\ &= \sup_{(v, q)} \frac{\left(C_F \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2} \right) \|\nabla v\|_{L^2} + \left(C_F \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2} \right) \|\nabla q\|_{L^2}}{(\|\nabla v\|_{L^2}^2 + \|\partial_t^{1/2} v\|_{L^2}^2 + \|\nabla q\|_{L^2}^2 + \|\partial_t^{1/2} q\|_{L^2}^2)^{1/2}} \\ &\leq C_F \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} + C_F \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)},\end{aligned}$$

where

$$\|\cdot\|_{L^2} = \|\cdot\|_{L^2(Q_T)} \quad \text{and} \quad |\cdot|_{H^{1, \frac{1}{2}}} = |\cdot|_{H^{1, \frac{1}{2}}(Q_T)}.$$

From the inf-sup condition (6.50) follows that

$$|e|_{H^{1, \frac{1}{2}}(Q_T)} \leq \frac{1}{\mu_1} \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}(e, (v, q))}{|(v, q)|_{H^{1, \frac{1}{2}}(Q_T)}} = \frac{1}{\mu_1} \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta, \zeta)}(v, q)}{|(v, q)|_{H^{1, \frac{1}{2}}(Q_T)}}.$$

Together with the estimates before, this yields the final estimate (6.54). \square

Remark 6.27. For computational reasons, it is common to reformulate majorants as, e.g., the majorant $\mathcal{M}_{|\cdot|}^{\oplus}(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho})$, in such a way that they are given by quadratic functionals, see, e.g., [59]. This is done by introducing some parameters $\alpha, \beta, \gamma > 0$, i.e.,

$$\begin{aligned} \mathcal{M}_{|\cdot|}^{\oplus}(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho})^2 &\leq \mathcal{M}_{|\cdot|}^{\oplus}(\alpha, \beta, \gamma; \eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho})^2 \\ &= \frac{1}{\mu_1^2} \left(C_F^2 (1 + \alpha)(1 + \beta) \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \frac{(1 + \alpha)(1 + \beta)}{\beta} \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + C_F^2 \frac{(1 + \alpha)(1 + \gamma)}{\alpha} \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 + \frac{(1 + \alpha)(1 + \gamma)}{\alpha\gamma} \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 \right). \end{aligned}$$

A similar estimate as (6.54) for the seminorm can be proven for the norm using the inf-sup condition (6.49).

Theorem 6.28. Let $\eta, \zeta \in H_{0,per}^{1,1}(Q_T)$ and the bilinear form $\mathcal{B}(\cdot, \cdot)$ defined by (6.46) satisfy (6.49). Then,

$$\begin{aligned} \|e\|_{H^{1,\frac{1}{2}}(Q_T)} &\leq \frac{1}{\mu_1} \left(\|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &=: \mathcal{M}_{\|\cdot\|}^{\oplus}(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}), \end{aligned} \quad (6.55)$$

where $e = (y - \eta, p - \zeta) \in (H_0^{1,\frac{1}{2}}(Q_T))^2$, $\mu_1 = \min\{1, \frac{1}{\lambda}, \frac{\nu}{\sqrt{\lambda}}, \nu\sqrt{\lambda}, \frac{\sigma}{\sqrt{\lambda}}, \sigma\sqrt{\lambda}\}$ and $\boldsymbol{\tau}, \boldsymbol{\rho} \in H(\text{div}, Q_T)$.

Proof. Applying the Cauchy-Schwarz inequality yields the estimate

$$\begin{aligned} \mathcal{F}_{(\eta,\zeta)}(v, q) &\leq \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2} \|v\|_{L^2} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2} \|\nabla v\|_{L^2} \\ &\quad + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2} \|q\|_{L^2} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2} \|\nabla q\|_{L^2} \\ &\leq \left(\|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2}^2 + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2}^2 + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2}^2 + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2}^2 \right)^{1/2} \\ &\quad \times \left(\|v\|_{L^2}^2 + \|\nabla v\|_{L^2}^2 + \|q\|_{L^2}^2 + \|\nabla q\|_{L^2}^2 \right)^{1/2} \\ &= \left(\|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2}^2 + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2}^2 + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2}^2 + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2}^2 \right)^{1/2} \\ &\quad \times \left(\|v\|_{H^{1,0}}^2 + \|q\|_{H^{1,0}}^2 \right)^{1/2} \\ &= \left(\|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2}^2 + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2}^2 + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2}^2 + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2}^2 \right)^{1/2} \|(v, q)\|_{H^{1,0}} \end{aligned}$$

with

$$\|\cdot\|_{L^2} = \|\cdot\|_{L^2(Q_T)} \quad \text{and} \quad \|\cdot\|_{H^{1,0}} = \|\cdot\|_{H^{1,0}(Q_T)}.$$

From the inf-sup condition (6.49) follows that

$$\begin{aligned} \|e\|_{H^{1,\frac{1}{2}}(Q_T)} &\leq \frac{1}{\mu_1} \sup_{0 \neq (v,q) \in (H_0^{1,\frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}(e, (v, q))}{\|(v, q)\|_{H^{1,\frac{1}{2}}(Q_T)}} = \frac{1}{\mu_1} \sup_{0 \neq (v,q) \in (H_0^{1,\frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta,\zeta)}(v, q)}{\|(v, q)\|_{H^{1,\frac{1}{2}}(Q_T)}} \\ &\leq \frac{1}{\mu_1} \left(\|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 \right)^{1/2}, \end{aligned}$$

where $\mu_1 = \min\{1, \frac{1}{\lambda}, \frac{\nu}{\sqrt{\lambda}}, \nu\sqrt{\lambda}, \frac{\sigma}{\sqrt{\lambda}}, \sigma\sqrt{\lambda}\}$ as proven in Lemma 6.23. \square

The multiharmonic approximation

Since we assume that the desired state y_d is from $L^2(Q_T)$, we expand it into a Fourier series in time. Moreover, we choose our approximations η and ζ to the exact state y and adjoint state p , respectively, as well as the vector-valued functions $\boldsymbol{\tau}$ and $\boldsymbol{\rho}$ to be some truncated Fourier series, e.g.,

$$\eta(\mathbf{x}, t) = \eta_0^c(\mathbf{x}) + \sum_{k=1}^N [\eta_k^c(\mathbf{x}) \cos(k\omega t) + \eta_k^s(\mathbf{x}) \sin(k\omega t)],$$

where all Fourier coefficients are at least from the space $L^2(\Omega)$ and are defined as, e.g.,

$$\begin{aligned} \eta_0^c(\mathbf{x}) &= \frac{1}{T} \int_0^T \eta(\mathbf{x}, t) dt, \\ \eta_k^c(\mathbf{x}) &= \frac{2}{T} \int_0^T \eta(\mathbf{x}, t) \cos(k\omega t) dt, \\ \eta_k^s(\mathbf{x}) &= \frac{2}{T} \int_0^T \eta(\mathbf{x}, t) \sin(k\omega t) dt. \end{aligned}$$

We need to compute the $L^2(Q_T)$ -norms of the functions

$$\begin{aligned} \mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau}) &= \sigma \partial_t \zeta + \eta - \operatorname{div} \boldsymbol{\tau} - y_d, & \mathcal{R}_2(\zeta, \boldsymbol{\tau}) &= \boldsymbol{\tau} + \nu \nabla \zeta, \\ \mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho}) &= \sigma \partial_t \eta + \lambda^{-1} \zeta - \operatorname{div} \boldsymbol{\rho}, & \mathcal{R}_4(\eta, \boldsymbol{\rho}) &= \boldsymbol{\rho} - \nu \nabla \eta. \end{aligned}$$

Due to the orthogonalities of the cosine and sine functions (2.9), the integrals in time can be computed easily. Moreover, remember that

$$\begin{aligned} \partial_t \eta(\mathbf{x}, t) &= \sum_{k=1}^N [k\omega \eta_k^s(\mathbf{x}) \cos(k\omega t) - k\omega \eta_k^c(\mathbf{x}) \sin(k\omega t)], \\ \nabla \eta(\mathbf{x}, t) &= \nabla \eta_0^c(\mathbf{x}) + \sum_{k=1}^N [\nabla \eta_k^c(\mathbf{x}) \cos(k\omega t) + \nabla \eta_k^s(\mathbf{x}) \sin(k\omega t)], \\ \operatorname{div} \boldsymbol{\tau}(\mathbf{x}, t) &= \operatorname{div} \boldsymbol{\tau}_0^c(\mathbf{x}) + \sum_{k=1}^N [\operatorname{div} \boldsymbol{\tau}_k^c(\mathbf{x}) \cos(k\omega t) + \operatorname{div} \boldsymbol{\tau}_k^s(\mathbf{x}) \sin(k\omega t)]. \end{aligned}$$

We will start with considering the L^2 -norm of \mathcal{R}_2 and \mathcal{R}_4 , and, then, of \mathcal{R}_3 and \mathcal{R}_1 . The time integrals are computed similarly as in Section 6.1. We obtain

$$\begin{aligned} \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 &= \int_0^T \int_{\Omega} |\boldsymbol{\tau} + \nu \nabla \zeta|^2 d\mathbf{x} dt \\ &= T \|\boldsymbol{\tau}_0^c + \nu \nabla \zeta_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N [\|\boldsymbol{\tau}_k^c + \nu \nabla \zeta_k^c\|_{L^2(\Omega)}^2 + \|\boldsymbol{\tau}_k^s + \nu \nabla \zeta_k^s\|_{L^2(\Omega)}^2] \\ &= T \|\boldsymbol{\tau}_0^c + \nu \nabla \zeta_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|\boldsymbol{\tau}_k + \nu \nabla \zeta_k\|_{L^2(\Omega)}^2, \end{aligned}$$

$$\begin{aligned} \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 &= \int_0^T \int_{\Omega} |\boldsymbol{\rho} - \nu \nabla \eta|^2 d\mathbf{x} dt \\ &= T \|\boldsymbol{\rho}_0^c - \nu \nabla \eta_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N [\|\boldsymbol{\rho}_k^c - \nu \nabla \eta_k^c\|_{L^2(\Omega)}^2 + \|\boldsymbol{\rho}_k^s - \nu \nabla \eta_k^s\|_{L^2(\Omega)}^2] \\ &= T \|\boldsymbol{\rho}_0^c - \nu \nabla \eta_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|\boldsymbol{\rho}_k - \nu \nabla \eta_k\|_{L^2(\Omega)}^2. \end{aligned}$$

and

$$\begin{aligned}
\|\mathcal{R}_3(\eta, \zeta, \rho)\|_{L^2(Q_T)}^2 &= \int_0^T \int_{\Omega} (\sigma \partial_t \eta + \lambda^{-1} \zeta - \operatorname{div} \rho)^2 dx dt \\
&= T \|\lambda^{-1} \zeta_0^c - \operatorname{div} \rho_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N [\|k\omega \sigma \eta_k^s + \lambda^{-1} \zeta_k^c - \operatorname{div} \rho_k^c\|_{L^2(\Omega)}^2 \\
&\quad + \|-k\omega \sigma \eta_k^c + \lambda^{-1} \zeta_k^s - \operatorname{div} \rho_k^s\|_{L^2(\Omega)}^2] \\
&= T \|\lambda^{-1} \zeta_0^c - \operatorname{div} \rho_0^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|-k\omega \sigma \eta_k^\perp + \lambda^{-1} \zeta_k - \operatorname{div} \rho_k\|_{L^2(\Omega)}^2.
\end{aligned}$$

Finally, we obtain

$$\begin{aligned}
\|\mathcal{R}_1(\eta, \zeta, \tau)\|_{L^2(Q_T)}^2 &= \int_0^T \int_{\Omega} (\sigma \partial_t \zeta + \eta - \operatorname{div} \tau - y_d)^2 dx dt \\
&= T \|\eta_0^c - \operatorname{div} \tau_0^c - y_{d0}^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N [\|k\omega \sigma \zeta_k^s + \eta_k^c - \operatorname{div} \tau_k^c - y_{dk}^c\|_{L^2(\Omega)}^2 \\
&\quad + \|-k\omega \sigma \zeta_k^c + \eta_k^s - \operatorname{div} \tau_k^s - y_{dk}^s\|_{L^2(\Omega)}^2] + \frac{T}{2} \sum_{k=N+1}^{\infty} [\|y_{dk}^c\|_{L^2(\Omega)}^2 + \|y_{dk}^s\|_{L^2(\Omega)}^2] \\
&= T \|\eta_0^c - \operatorname{div} \tau_0^c - y_{d0}^c\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|-k\omega \sigma \zeta_k^\perp + \eta_k - \operatorname{div} \tau_k - y_{dk}\|_{L^2(\Omega)}^2 \\
&\quad + \frac{T}{2} \sum_{k=N+1}^{\infty} \|y_{dk}\|_{L^2(\Omega)}^2.
\end{aligned}$$

Again, the term

$$\frac{T}{2} \sum_{k=N+1}^{\infty} \|y_{dk}\|_{L^2(\Omega)}^2 = \frac{T}{2} \sum_{k=N+1}^{\infty} [\|y_{dk}^c\|_{L^2(\Omega)}^2 + \|y_{dk}^s\|_{L^2(\Omega)}^2]$$

corresponds to the high oscillatory part of the given data, which is here the given desired state y_d , and, hence, this term can be controlled due to the knowledge on the given data.

All the L^2 -norms of \mathcal{R}_1 , \mathcal{R}_2 , \mathcal{R}_3 and \mathcal{R}_4 corresponding to every single mode k are decoupled. Here, we define the following functions for every mode $k = 1, \dots, N$:

$$\begin{aligned}
\mathcal{R}_{1k}(\eta_k, \zeta_k, \tau_k) &= -k\omega \sigma \zeta_k^\perp + \eta_k - \operatorname{div} \tau_k - y_{dk} = (\mathcal{R}_{1k}^c(\eta_k^c, \zeta_k^s, \tau_k^c), \mathcal{R}_{1k}^s(\eta_k^s, \zeta_k^c, \tau_k^s))^T \\
&= (k\omega \sigma \zeta_k^s + \eta_k^c - \operatorname{div} \tau_k^c - y_{dk}^c, -k\omega \sigma \zeta_k^c + \eta_k^s - \operatorname{div} \tau_k^s - y_{dk}^s)^T, \\
\mathcal{R}_{2k}(\zeta_k, \tau_k) &= \tau_k + \nu \nabla \zeta_k = (\mathcal{R}_{2k}^c(\zeta_k^c, \tau_k^c), \mathcal{R}_{2k}^s(\zeta_k^s, \tau_k^s))^T = (\tau_k^c + \nu \nabla \zeta_k^c, \tau_k^s + \nu \nabla \zeta_k^s)^T, \\
\mathcal{R}_{3k}(\eta_k, \zeta_k, \rho_k) &= -k\omega \sigma \eta_k^\perp + \lambda^{-1} \zeta_k - \operatorname{div} \rho_k = (\mathcal{R}_{3k}^c(\eta_k^s, \zeta_k^c, \rho_k^c), \mathcal{R}_{3k}^s(\eta_k^c, \zeta_k^s, \rho_k^s))^T \\
&= (k\omega \sigma \eta_k^s + \lambda^{-1} \zeta_k^c - \operatorname{div} \rho_k^c, -k\omega \sigma \eta_k^c + \lambda^{-1} \zeta_k^s - \operatorname{div} \rho_k^s)^T, \\
\mathcal{R}_{4k}(\eta_k, \rho_k) &= \rho_k - \nu \nabla \eta_k = (\mathcal{R}_{4k}^c(\eta_k^c, \rho_k^c), \mathcal{R}_{4k}^s(\eta_k^s, \rho_k^s))^T = (\rho_k^c - \nu \nabla \eta_k^c, \rho_k^s - \nu \nabla \eta_k^s)^T,
\end{aligned} \tag{6.56}$$

and, for $k = 0$, we define

$$\begin{aligned}
\mathcal{R}_{10}^c(\eta_0^c, \tau_0^c) &= \eta_0^c - \operatorname{div} \tau_0^c - y_{d0}^c, & \mathcal{R}_{20}^c(\zeta_0^c, \tau_0^c) &= \tau_0^c + \nu \nabla \zeta_0^c, \\
\mathcal{R}_{30}^c(\zeta_0^c, \rho_0^c) &= \lambda^{-1} \zeta_0^c - \operatorname{div} \rho_0^c, & \mathcal{R}_{40}^c(\eta_0^c, \rho_0^c) &= \rho_0^c - \nu \nabla \eta_0^c.
\end{aligned} \tag{6.57}$$

Altogether, we obtain the following error majorants corresponding to the optimal control problem:

Corollary 6.29. *The error majorants $\mathcal{M}_{|\cdot|}^\oplus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho})$ and $\mathcal{M}_{\|\cdot\|}^\oplus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho})$ are given by*

$$\begin{aligned} \mathcal{M}_{|\cdot|}^\oplus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}) &= \frac{1}{\mu_{1,|\cdot|}} \left(C_F \|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)} \right. \\ &\quad \left. + C_F \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)} + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)} \right) \\ &= \frac{1}{\mu_{1,|\cdot|}} \left(C_F (T \|\mathcal{R}_{10}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=N+1}^{\infty} \|\mathbf{y}_{dk}\|_{L^2(\Omega)}^2)^{1/2} \right. \\ &\quad \left. + (T \|\mathcal{R}_{20}^c(\zeta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|\mathcal{R}_{2k}(\boldsymbol{\zeta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2)^{1/2} \right. \\ &\quad \left. + C_F (T \|\mathcal{R}_{30}^c(\zeta_0^c, \boldsymbol{\rho}_0^c)\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|\mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \boldsymbol{\rho}_k)\|_{L^2(\Omega)}^2)^{1/2} \right. \\ &\quad \left. + (T \|\mathcal{R}_{40}^c(\eta_0^c, \boldsymbol{\rho}_0^c)\|_{L^2(\Omega)}^2 + \frac{T}{2} \sum_{k=1}^N \|\mathcal{R}_{4k}(\boldsymbol{\eta}_k, \boldsymbol{\rho}_k)\|_{L^2(\Omega)}^2)^{1/2} \right), \end{aligned}$$

where $\mu_{1,|\cdot|} = \min\{\frac{1}{\sqrt{\lambda}}, \sqrt{\lambda}\} \min\{\underline{\nu}, \underline{\sigma}\}$, and

$$\begin{aligned} \mathcal{M}_{\|\cdot\|}^\oplus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}) &= \frac{1}{\mu_{1,\|\cdot\|}} \left(\|\mathcal{R}_1(\eta, \zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\zeta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\mathcal{R}_3(\eta, \zeta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 + \|\mathcal{R}_4(\eta, \boldsymbol{\rho})\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &= \frac{1}{\mu_{1,\|\cdot\|}} \left(T (\|\mathcal{R}_{10}^c(\eta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{20}^c(\zeta_0^c, \boldsymbol{\tau}_0^c)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{30}^c(\zeta_0^c, \boldsymbol{\rho}_0^c)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{40}^c(\eta_0^c, \boldsymbol{\rho}_0^c)\|_{L^2(\Omega)}^2) \right. \\ &\quad \left. + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{2k}(\boldsymbol{\zeta}_k, \boldsymbol{\tau}_k)\|_{L^2(\Omega)}^2 + \|\mathcal{R}_{3k}(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \boldsymbol{\rho}_k)\|_{L^2(\Omega)}^2 \right. \\ &\quad \left. + \|\mathcal{R}_{4k}(\boldsymbol{\eta}_k, \boldsymbol{\rho}_k)\|_{L^2(\Omega)}^2] + \frac{T}{2} \sum_{k=N+1}^{\infty} \|\mathbf{y}_{dk}\|_{L^2(\Omega)}^2 \right)^{1/2}, \end{aligned}$$

where $\mu_{1,\|\cdot\|} = \min\{1, \frac{1}{\lambda}, \frac{\underline{\nu}}{\sqrt{\lambda}}, \underline{\nu}\sqrt{\lambda}, \frac{\underline{\sigma}}{\sqrt{\lambda}}, \underline{\sigma}\sqrt{\lambda}\}$, respectively.

Remark 6.30. *It is easy to see that η is the exact state and ζ the adjoint state of problem (6.45) and $\boldsymbol{\tau}, \boldsymbol{\rho}$ are their exact fluxes if and only if the error majorants vanish, i.e.,*

$$\mathcal{R}_{j_0}^c = 0 \quad \text{and} \quad \mathcal{R}_{jk} = 0 \quad \forall k = 1, \dots, N, \quad \forall j \in \{1, 2, 3, 4\}, \quad (6.58)$$

i.e.,

$$\begin{aligned} k\omega \sigma \zeta_k^s + \eta_k^c - \operatorname{div} \boldsymbol{\tau}_k^c &= y_{dk}^c, & -k\omega \sigma \zeta_k^c + \eta_k^s - \operatorname{div} \boldsymbol{\tau}_k^s &= y_{dk}^s, \\ k\omega \sigma \eta_k^s + \lambda^{-1} \zeta_k^c - \operatorname{div} \boldsymbol{\rho}_k^c &= 0, & -k\omega \sigma \eta_k^c + \lambda^{-1} \zeta_k^s - \operatorname{div} \boldsymbol{\rho}_k^s &= 0, \\ \boldsymbol{\tau}_k^c &= -\nu \nabla \zeta_k^c, & \boldsymbol{\tau}_k^s &= -\nu \nabla \zeta_k^s, & \boldsymbol{\rho}_k^c &= \nu \nabla \eta_k^c, & \boldsymbol{\rho}_k^s &= \nu \nabla \eta_k^s, \end{aligned}$$

for all $k = 1, \dots, N$, and

$$\eta_0^c - \operatorname{div} \boldsymbol{\tau}_0^c = y_{d0}^c, \quad \boldsymbol{\tau}_0^c = -\nu \nabla \zeta_0^c, \quad \lambda^{-1} \zeta_0^c - \operatorname{div} \boldsymbol{\rho}_0^c = 0, \quad \boldsymbol{\rho}_0^c = \nu \nabla \eta_0^c,$$

for $k = 0$, and the given desired state y_d has a multiharmonic representation, i.e.,

$$y_d(\mathbf{x}, t) = y_{d0}^c(\mathbf{x}) + \sum_{k=1}^N [y_{dk}^c(\mathbf{x}) \cos(k\omega t) + y_{dk}^s(\mathbf{x}) \sin(k\omega t)].$$

Moreover, (η, ζ) and $(\boldsymbol{\tau}, \boldsymbol{\rho})$ converge to the exact solution and flux, respectively, if and only if $\eta = \eta_N$, $\zeta = \zeta_N$, $\boldsymbol{\tau} = \boldsymbol{\tau}_N$ and $\boldsymbol{\rho} = \boldsymbol{\rho}_N$ with N going to infinity and the error majorants corresponding to the modes $k = 0, 1, \dots$ vanish as in (6.58).

A second a posteriori error result

Let us assume that the approximations η and ζ of y and p are not from the space $H_{0,per}^{1,1}(Q_T)$, but are less regular. More precisely, we want to deduce an upper bound valid for approximations

$$\eta, \zeta \in H_0^{1, \frac{1}{2}}(Q_T).$$

Let us again consider the functional

$$\begin{aligned} \mathcal{F}_{(\eta, \zeta)}(v, q) = \int_0^T \int_{\Omega} & \left(y_d v - \eta v + \nu(\mathbf{x}) \nabla \zeta \cdot \nabla v - \sigma(\mathbf{x}) \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp \right. \\ & \left. - \nu(\mathbf{x}) \nabla \eta \cdot \nabla q - \sigma(\mathbf{x}) \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \end{aligned}$$

defined for all $v, q \in H_0^{1, \frac{1}{2}}(Q_T)$. Besides the vector-valued functions $\boldsymbol{\tau}, \boldsymbol{\rho} \in H(\text{div}, Q_T)$, let us introduce the functions

$$\kappa, \chi \in H^{0, \frac{1}{2}}(Q_T),$$

which fulfill the identity (6.30), i.e.,

$$\int_0^T \kappa \partial_t^{1/2} v^\perp dt = - \int_0^T \partial_t^{1/2} \kappa^\perp v dt$$

for all $v \in H_0^{0, \frac{1}{2}}(Q_T)$. We rearrange the functional $\mathcal{F}_{(\eta, \zeta)}(v, q)$ and write it as

$$\begin{aligned} \mathcal{F}_{(\eta, \zeta)}(v, q) &= \int_0^T \int_{\Omega} \left(y_d v - \eta v + \nu \nabla \zeta \cdot \nabla v - \sigma \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp \right. \\ & \quad \left. - \nu \nabla \eta \cdot \nabla q - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(y_d v - \eta v + \nu \nabla \zeta \cdot \nabla v + \text{div } \boldsymbol{\tau} v + \boldsymbol{\tau} \cdot \nabla v \right. \\ & \quad \left. - \sigma \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp + \sigma \partial_t^{1/2} \kappa^\perp v \right. \\ & \quad \left. - \nu \nabla \eta \cdot \nabla q + \text{div } \boldsymbol{\rho} q + \boldsymbol{\rho} \cdot \nabla q \right. \\ & \quad \left. - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp + \sigma \chi \partial_t^{1/2} q^\perp + \sigma \partial_t^{1/2} \chi^\perp q - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(y_d v - \eta v + \text{div } \boldsymbol{\tau} v + \sigma \partial_t^{1/2} \kappa^\perp v \right. \\ & \quad \left. - \sigma \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp + \nu \nabla \zeta \cdot \nabla v + \boldsymbol{\tau} \cdot \nabla v \right. \\ & \quad \left. - \nu \nabla \eta \cdot \nabla q + \boldsymbol{\rho} \cdot \nabla q - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp + \sigma \chi \partial_t^{1/2} q^\perp \right. \\ & \quad \left. + \text{div } \boldsymbol{\rho} q + \sigma \partial_t^{1/2} \chi^\perp q - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(\left(y_d - \eta + \text{div } \boldsymbol{\tau} + \sigma \partial_t^{1/2} \kappa^\perp \right) v + \left(\sigma(\kappa - \partial_t^{1/2} \zeta) \right) \partial_t^{1/2} v^\perp \right. \\ & \quad \left. + (\boldsymbol{\tau} + \nu \nabla \zeta) \cdot \nabla v + (\boldsymbol{\rho} - \nu \nabla \eta) \cdot \nabla q \right. \\ & \quad \left. + \left(\sigma(\chi - \partial_t^{1/2} \eta) \right) \partial_t^{1/2} q^\perp + \left(\text{div } \boldsymbol{\rho} + \sigma \partial_t^{1/2} \chi^\perp - \lambda^{-1} \zeta \right) q \right) d\mathbf{x} dt \end{aligned}$$

for all $v, q \in H_0^{1, \frac{1}{2}}(Q_T)$. Let

$$\begin{aligned}\mathcal{R}_1(\boldsymbol{\tau}, \boldsymbol{\kappa}, \eta) &= y_d - \eta + \operatorname{div} \boldsymbol{\tau} + \sigma \partial_t^{1/2} \boldsymbol{\kappa}^\perp, \\ \mathcal{R}_2(\boldsymbol{\tau}, \zeta) &= \boldsymbol{\tau} + \nu \nabla \zeta, \\ \mathcal{R}_3(\boldsymbol{\kappa}, \zeta) &= \sigma(\boldsymbol{\kappa} - \partial_t^{1/2} \zeta), \\ \mathcal{R}_4(\boldsymbol{\rho}, \chi, \zeta) &= \operatorname{div} \boldsymbol{\rho} + \sigma \partial_t^{1/2} \chi^\perp - \lambda^{-1} \zeta, \\ \mathcal{R}_5(\boldsymbol{\rho}, \eta) &= \boldsymbol{\rho} - \nu \nabla \eta, \\ \mathcal{R}_6(\chi, \eta) &= \sigma(\chi - \partial_t^{1/2} \eta).\end{aligned}$$

Hence, using Cauchy-Schwarz inequality and the identity $\|\partial_t^{1/2} v^\perp\|_{L^2(Q_T)} = \|\partial_t^{1/2} v\|_{L^2(Q_T)}$, which is valid for $v \in H_0^{0, \frac{1}{2}}(Q_T)$, we can estimate the functional $\mathcal{F}_{(\eta, \zeta)}(v, q)$ from above as follows

$$\begin{aligned}\mathcal{F}_{(\eta, \zeta)}(v, q) &\leq \|\mathcal{R}_1(\boldsymbol{\tau}, \boldsymbol{\kappa}, \eta)\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} + \|\mathcal{R}_2(\boldsymbol{\tau}, \zeta)\|_{L^2(Q_T)} \|\nabla v\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_3(\boldsymbol{\kappa}, \zeta)\|_{L^2(Q_T)} \|\partial_t^{1/2} v\|_{L^2(Q_T)} + \|\mathcal{R}_4(\boldsymbol{\rho}, \chi, \zeta)\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)} \\ &\quad + \|\mathcal{R}_5(\boldsymbol{\rho}, \eta)\|_{L^2(Q_T)} \|\nabla q\|_{L^2(Q_T)} + \|\mathcal{R}_6(\chi, \eta)\|_{L^2(Q_T)} \|\partial_t^{1/2} q\|_{L^2(Q_T)} \\ &\leq \left(\|\mathcal{R}_1(\boldsymbol{\tau}, \boldsymbol{\kappa}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\boldsymbol{\tau}, \zeta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_3(\boldsymbol{\kappa}, \zeta)\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\mathcal{R}_4(\boldsymbol{\rho}, \chi, \zeta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_5(\boldsymbol{\rho}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_6(\chi, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2} \|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)},\end{aligned}$$

where

$$\begin{aligned}\|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)} &= \left(\|v\|_{L^2(Q_T)}^2 + \|\nabla v\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} v\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|q\|_{L^2(Q_T)}^2 + \|\nabla q\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} q\|_{L^2(Q_T)}^2 \right)^{1/2}.\end{aligned}$$

Altogether, we obtain the upper bound

$$\begin{aligned}\sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta, \zeta)}(v, q)}{\|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)}} &\leq \left(\|\mathcal{R}_1(\boldsymbol{\tau}, \boldsymbol{\kappa}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\boldsymbol{\tau}, \zeta)\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\mathcal{R}_3(\boldsymbol{\kappa}, \zeta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_4(\boldsymbol{\rho}, \chi, \zeta)\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\mathcal{R}_5(\boldsymbol{\rho}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_6(\chi, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2},\end{aligned}\tag{6.59}$$

and, finally, deduce the following theorem:

Theorem 6.31. *Let $\eta, \zeta \in H_0^{1, \frac{1}{2}}(Q_T)$ and the bilinear form $\mathcal{B}(\cdot, \cdot)$ satisfy (6.49). Then,*

$$\begin{aligned}\|e\|_{H^{1, \frac{1}{2}}(Q_T)} &\leq \frac{1}{\mu_1} \left(\|\mathcal{R}_1(\boldsymbol{\tau}, \boldsymbol{\kappa}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_2(\boldsymbol{\tau}, \zeta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_3(\boldsymbol{\kappa}, \zeta)\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \|\mathcal{R}_4(\boldsymbol{\rho}, \chi, \zeta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_5(\boldsymbol{\rho}, \eta)\|_{L^2(Q_T)}^2 + \|\mathcal{R}_6(\chi, \eta)\|_{L^2(Q_T)}^2 \right)^{1/2} \\ &=: \mathcal{M}_{\|\cdot\|}^{\oplus}(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}, \boldsymbol{\kappa}, \chi),\end{aligned}\tag{6.60}$$

where $\boldsymbol{\tau}, \boldsymbol{\rho} \in H(\operatorname{div}, Q_T)$, $\boldsymbol{\kappa}, \chi \in H^{0, \frac{1}{2}}(Q_T)$ and $\mu_1 = \min\{1, \frac{1}{\lambda}, \frac{\nu}{\sqrt{\lambda}}, \nu\sqrt{\lambda}, \frac{\sigma}{\sqrt{\lambda}}, \sigma\sqrt{\lambda}\}$.

Proof. From (6.49) follows that

$$\|e\|_{H^{1, \frac{1}{2}}(Q_T)} \leq \frac{1}{\mu_1} \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}(e, (v, q))}{\|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)}} = \frac{1}{\mu_1} \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta, \zeta)}(v, q)}{\|(v, q)\|_{H^{1, \frac{1}{2}}(Q_T)}},$$

which leads together with (6.59) to the final estimate (6.60). \square

Remark 6.32. If $\mathcal{R}_1(\boldsymbol{\tau}, \kappa, \eta) = 0$, $\mathcal{R}_2(\boldsymbol{\tau}, \zeta) = 0$, $\mathcal{R}_3(\kappa, \zeta) = 0$, $\mathcal{R}_4(\boldsymbol{\rho}, \chi, \zeta) = 0$, $\mathcal{R}_5(\boldsymbol{\rho}, \eta) = 0$ and $\mathcal{R}_6(\chi, \eta) = 0$, then

$$\begin{aligned} -\sigma \partial_t^{1/2} \kappa^\perp - \operatorname{div} \boldsymbol{\tau} + \eta &= y_d, & -\sigma \partial_t^{1/2} \chi^\perp - \operatorname{div} \boldsymbol{\rho} + \lambda^{-1} \zeta &= 0, \\ \boldsymbol{\tau} &= -\nu \nabla \zeta, & \boldsymbol{\rho} &= \nu \nabla \eta, \\ \kappa &= \partial_t^{1/2} \zeta, & \chi &= \partial_t^{1/2} \eta. \end{aligned}$$

Since η and ζ satisfy the Dirichlet condition on Σ_T , (η, ζ) is the solution. In other words, the majorant $\mathcal{M}_{\|\cdot\|}^\oplus(\eta, \zeta, \boldsymbol{\tau}, \boldsymbol{\rho}, \kappa, \chi)$ vanishes if and only if (η, ζ) is the exact solution, $(\boldsymbol{\tau}, \boldsymbol{\rho})$ the exact flux and (κ, χ) the exact half time derivative. Moreover, if $\eta, \zeta \in H_{0,per}^{1,1}(Q_T)$, we derive the optimality system

$$\begin{aligned} \sigma \partial_t \zeta + \operatorname{div}(\nu \nabla \zeta) + \eta &= y_d & \text{in } Q_T, \\ \sigma \partial_t \eta - \operatorname{div}(\nu \nabla \eta) + \lambda^{-1} \zeta &= 0 & \text{in } Q_T, \end{aligned}$$

in the weak sense.

Let us now find an upper bound for the supremum (6.53) in the V_0 -norm case, i.e.,

$$\sup_{0 \neq (v,q) \in (H_0^{1,\frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta,\zeta)}(v,q)}{\|(v,q)\|_{V_0}},$$

where

$$\|(v,q)\|_{V_0}^2 = \|v\|_{V_0}^2 + \lambda^{-1} \|q\|_{V_0}^2,$$

with

$$\|v\|_{V_0}^2 = \|v\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\nu \nabla v, \nabla v)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)}.$$

Let us introduce some new functions $\tilde{\boldsymbol{\tau}}$ and $\tilde{\boldsymbol{\rho}}$, where $(\nu \tilde{\boldsymbol{\tau}}), (\nu \tilde{\boldsymbol{\rho}}) \in H(\operatorname{div}, Q_T)$ and the identity

$$\int_{\Omega} \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}) v \, d\mathbf{x} = - \int_{\Omega} (\nu \tilde{\boldsymbol{\tau}}) \cdot \nabla v \, d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega)$$

is fulfilled for both vector-valued functions $\tilde{\boldsymbol{\tau}}$ and $\tilde{\boldsymbol{\rho}}$. Then,

$$\begin{aligned} \mathcal{F}_{(\eta,\zeta)}(v,q) &= \int_0^T \int_{\Omega} \left(y_d v - \eta v + \nu \nabla \zeta \cdot \nabla v - \sigma \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp \right. \\ &\quad \left. - \nu \nabla \eta \cdot \nabla q - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(y_d v - \eta v + \nu \nabla \zeta \cdot \nabla v + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}) v + (\nu \tilde{\boldsymbol{\tau}}) \cdot \nabla v \right. \\ &\quad \left. - \sigma \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp + \sigma \partial_t^{1/2} \kappa^\perp v \right. \\ &\quad \left. - \nu \nabla \eta \cdot \nabla q + \operatorname{div}(\nu \tilde{\boldsymbol{\rho}}) q + (\nu \tilde{\boldsymbol{\rho}}) \cdot \nabla q \right. \\ &\quad \left. - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp + \sigma \chi \partial_t^{1/2} q^\perp + \sigma \partial_t^{1/2} \chi^\perp q - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \\ &= \int_0^T \int_{\Omega} \left(y_d v - \eta v + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}) v + \sigma \partial_t^{1/2} \kappa^\perp v \right. \\ &\quad \left. - \sigma \partial_t^{1/2} \zeta \partial_t^{1/2} v^\perp + \sigma \kappa \partial_t^{1/2} v^\perp + \nu \nabla \zeta \cdot \nabla v + (\nu \tilde{\boldsymbol{\tau}}) \cdot \nabla v \right. \\ &\quad \left. - \nu \nabla \eta \cdot \nabla q + (\nu \tilde{\boldsymbol{\rho}}) \cdot \nabla q - \sigma \partial_t^{1/2} \eta \partial_t^{1/2} q^\perp + \sigma \chi \partial_t^{1/2} q^\perp \right. \\ &\quad \left. + \operatorname{div}(\nu \tilde{\boldsymbol{\rho}}) q + \sigma \partial_t^{1/2} \chi^\perp q - \lambda^{-1} \zeta q \right) d\mathbf{x} dt \end{aligned}$$

and

$$\begin{aligned} \mathcal{F}_{(\eta, \zeta)}(v, q) = & \int_0^T \int_{\Omega} \left((y_d - \eta + \operatorname{div}(\nu \tilde{\tau}) + \sigma \partial_t^{1/2} \kappa^\perp) v + \sigma(\kappa - \partial_t^{1/2} \zeta) \partial_t^{1/2} v^\perp \right. \\ & + \nu(\tilde{\tau} + \nabla \zeta) \cdot \nabla v + \nu(\tilde{\rho} - \nabla \eta) \cdot \nabla q \\ & \left. + \sigma(\chi - \partial_t^{1/2} \eta) \partial_t^{1/2} q^\perp + (\operatorname{div}(\nu \tilde{\rho}) + \sigma \partial_t^{1/2} \chi^\perp - \lambda^{-1} \zeta) q \right) dx dt \end{aligned}$$

for all $v, q \in H_0^{1, \frac{1}{2}}(Q_T)$. Let us now define

$$\begin{aligned} \mathcal{R}_1(\tilde{\tau}, \kappa, \eta) &= y_d - \eta + \operatorname{div}(\nu \tilde{\tau}) + \sigma \partial_t^{1/2} \kappa^\perp, & \mathcal{R}_4(\tilde{\rho}, \chi, \zeta) &= \operatorname{div}(\nu \tilde{\rho}) + \sigma \partial_t^{1/2} \chi^\perp - \lambda^{-1} \zeta, \\ \mathcal{R}_2(\tilde{\tau}, \zeta) &= \tilde{\tau} + \nabla \zeta, & \mathcal{R}_5(\tilde{\rho}, \eta) &= \tilde{\rho} - \nabla \eta, \\ \mathcal{R}_3(\kappa, \zeta) &= \kappa - \partial_t^{1/2} \zeta, & \mathcal{R}_6(\chi, \eta) &= \chi - \partial_t^{1/2} \eta. \end{aligned}$$

Hence, we obtain the following a posteriori error result for the V_0 -norm:

Theorem 6.33. *Let $\eta, \zeta \in H_0^{1, \frac{1}{2}}(Q_T)$ and the bilinear form $\mathcal{B}(\cdot, \cdot)$ defined by (6.46) satisfy (6.51). Then,*

$$\begin{aligned} \|e\|_{V_0} &\leq \frac{1}{\mu_1} \left(\|\mathcal{R}_1(\tilde{\tau}, \kappa, \eta)\|_{L^2(Q_T)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_2(\tilde{\tau}, \zeta), \mathcal{R}_2(\tilde{\tau}, \zeta))_{L^2(Q_T)} \right. \\ &\quad + \frac{1}{\sqrt{\lambda}} (\sigma \mathcal{R}_3(\kappa, \zeta), \mathcal{R}_3(\kappa, \zeta))_{L^2(Q_T)} + \lambda \|\mathcal{R}_4(\tilde{\rho}, \chi, \zeta)\|_{L^2(Q_T)}^2 \\ &\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_5(\tilde{\rho}, \eta), \mathcal{R}_5(\tilde{\rho}, \eta))_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \mathcal{R}_6(\chi, \eta), \mathcal{R}_6(\chi, \eta))_{L^2(Q_T)} \right)^{1/2} \\ &=: \mathcal{M}_{\|\cdot\|_{V_0}}^\oplus(\eta, \zeta, \tilde{\tau}, \tilde{\rho}, \kappa, \chi), \end{aligned} \tag{6.61}$$

where $(\nu \tilde{\tau}), (\nu \tilde{\rho}) \in H(\operatorname{div}, Q_T)$, $\kappa, \chi \in H^{0, \frac{1}{2}}(Q_T)$ and $\mu_1 = 1/\sqrt{3}$.

Proof. Using the Cauchy-Schwarz inequality as well as a proper weighting with λ , we can estimate the functional $\mathcal{F}_{(\eta, \zeta)}(v, q)$ from above as follows

$$\begin{aligned} \mathcal{F}_{(\eta, \zeta)}(v, q) &\leq \|y_d - \eta + \operatorname{div}(\nu \tilde{\tau}) + \sigma \partial_t^{1/2} \kappa^\perp\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} \\ &\quad + (\sigma(\kappa - \partial_t^{1/2} \zeta), \kappa - \partial_t^{1/2} \zeta)_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} v^\perp, \partial_t^{1/2} v^\perp)_{L^2(Q_T)}^{1/2} \\ &\quad + (\nu(\tilde{\tau} + \nabla \zeta), \tilde{\tau} + \nabla \zeta)_{L^2(Q_T)}^{1/2} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2} \\ &\quad + (\nu(\tilde{\rho} - \nabla \eta), \tilde{\rho} - \nabla \eta)_{L^2(Q_T)}^{1/2} (\nu \nabla q, \nabla q)_{L^2(Q_T)}^{1/2} \\ &\quad + (\sigma(\chi - \partial_t^{1/2} \eta), \chi - \partial_t^{1/2} \eta)_{L^2(Q_T)}^{1/2} (\sigma \partial_t^{1/2} q^\perp, \partial_t^{1/2} q^\perp)_{L^2(Q_T)}^{1/2} \\ &\quad + \|\operatorname{div}(\nu \tilde{\rho}) + \sigma \partial_t^{1/2} \chi^\perp - \lambda^{-1} \zeta\|_{L^2(Q_T)} \|q\|_{L^2(Q_T)} \\ &= \|y_d - \eta + \operatorname{div}(\nu \tilde{\tau}) + \sigma \partial_t^{1/2} \kappa^\perp\|_{L^2(Q_T)} \|v\|_{L^2(Q_T)} \\ &\quad + \lambda^{-1/4} (\sigma(\kappa - \partial_t^{1/2} \zeta), \kappa - \partial_t^{1/2} \zeta)_{L^2(Q_T)}^{1/2} \lambda^{1/4} (\sigma \partial_t^{1/2} v^\perp, \partial_t^{1/2} v^\perp)_{L^2(Q_T)}^{1/2} \\ &\quad + \lambda^{-1/4} (\nu(\tilde{\tau} + \nabla \zeta), \tilde{\tau} + \nabla \zeta)_{L^2(Q_T)}^{1/2} \lambda^{1/4} (\nu \nabla v, \nabla v)_{L^2(Q_T)}^{1/2} \\ &\quad + \lambda^{1/4} (\nu(\tilde{\rho} - \nabla \eta), \tilde{\rho} - \nabla \eta)_{L^2(Q_T)}^{1/2} \lambda^{-1/4} (\nu \nabla q, \nabla q)_{L^2(Q_T)}^{1/2} \\ &\quad + \lambda^{1/4} (\sigma(\chi - \partial_t^{1/2} \eta), \chi - \partial_t^{1/2} \eta)_{L^2(Q_T)}^{1/2} \lambda^{-1/4} (\sigma \partial_t^{1/2} q^\perp, \partial_t^{1/2} q^\perp)_{L^2(Q_T)}^{1/2} \\ &\quad + \lambda^{1/2} \|\operatorname{div}(\nu \tilde{\rho}) + \sigma \partial_t^{1/2} \chi^\perp - \lambda^{-1} \zeta\|_{L^2(Q_T)} \lambda^{-1/2} \|q\|_{L^2(Q_T)} \end{aligned}$$

and

$$\begin{aligned}
\mathcal{F}_{(\eta, \zeta)}(v, q) &\leq \left(\|y_d - \eta + \operatorname{div}(\nu \tilde{\tau}) + \sigma \partial_t^{1/2} \kappa^\perp\|_{L^2(Q_T)}^2 + \frac{1}{\sqrt{\lambda}} (\sigma(\kappa - \partial_t^{1/2} \zeta), \kappa - \partial_t^{1/2} \zeta)_{L^2(Q_T)} \right. \\
&\quad + \frac{1}{\sqrt{\lambda}} (\nu(\tilde{\tau} + \nabla \zeta), \tilde{\tau} + \nabla \zeta)_{L^2(Q_T)} + \sqrt{\lambda} (\nu(\tilde{\rho} - \nabla \eta), \tilde{\rho} - \nabla \eta)_{L^2(Q_T)} \\
&\quad + \sqrt{\lambda} (\sigma(\chi - \partial_t^{1/2} \eta), \chi - \partial_t^{1/2} \eta)_{L^2(Q_T)} + \lambda \|\operatorname{div}(\nu \tilde{\rho}) + \sigma \partial_t^{1/2} \chi^\perp - \lambda^{-1} \zeta\|_{L^2(Q_T)}^2 \Big)^{1/2} \\
&\quad \times \left(\|v\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)} + \sqrt{\lambda} (\nu \nabla v, \nabla v)_{L^2(Q_T)} \right. \\
&\quad \left. + \frac{1}{\sqrt{\lambda}} (\nu \nabla q, \nabla q)_{L^2(Q_T)} + \frac{1}{\sqrt{\lambda}} (\sigma \partial_t^{1/2} q, \partial_t^{1/2} q)_{L^2(Q_T)} + \lambda^{-1} \|q\|_{L^2(Q_T)}^2 \right)^{1/2},
\end{aligned}$$

which yields the estimate

$$\begin{aligned}
\mathcal{F}_{(\eta, \zeta)}(v, q) &\leq \left(\|\mathcal{R}_1(\tilde{\tau}, \kappa, \eta)\|_{L^2(Q_T)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_2(\tilde{\tau}, \zeta), \mathcal{R}_2(\tilde{\tau}, \zeta))_{L^2(Q_T)} \right. \\
&\quad + \frac{1}{\sqrt{\lambda}} (\sigma \mathcal{R}_3(\kappa, \zeta), \mathcal{R}_3(\kappa, \zeta))_{L^2(Q_T)} + \lambda \|\mathcal{R}_4(\tilde{\rho}, \chi, \zeta)\|_{L^2(Q_T)}^2 \\
&\quad + \sqrt{\lambda} (\nu \mathcal{R}_5(\tilde{\rho}, \eta), \mathcal{R}_5(\tilde{\rho}, \eta))_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \mathcal{R}_6(\chi, \eta), \mathcal{R}_6(\chi, \eta))_{L^2(Q_T)} \Big)^{1/2} \\
&\quad \times \left(\|v\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\nu \nabla v, \nabla v)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} v, \partial_t^{1/2} v)_{L^2(Q_T)} \right. \\
&\quad \left. + \lambda^{-1} (\|q\|_{L^2(Q_T)}^2 + \sqrt{\lambda} (\nu \nabla q, \nabla q)_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \partial_t^{1/2} q, \partial_t^{1/2} q)_{L^2(Q_T)}) \right)^{1/2} \\
&= \left(\|\mathcal{R}_1(\tilde{\tau}, \kappa, \eta)\|_{L^2(Q_T)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_2(\tilde{\tau}, \zeta), \mathcal{R}_2(\tilde{\tau}, \zeta))_{L^2(Q_T)} \right. \\
&\quad + \frac{1}{\sqrt{\lambda}} (\sigma \mathcal{R}_3(\kappa, \zeta), \mathcal{R}_3(\kappa, \zeta))_{L^2(Q_T)} + \lambda \|\mathcal{R}_4(\tilde{\rho}, \chi, \zeta)\|_{L^2(Q_T)}^2 \\
&\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_5(\tilde{\rho}, \eta), \mathcal{R}_5(\tilde{\rho}, \eta))_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \mathcal{R}_6(\chi, \eta), \mathcal{R}_6(\chi, \eta))_{L^2(Q_T)} \right)^{1/2} \|(v, q)\|_{V_0}.
\end{aligned}$$

Altogether, we obtain the upper bound

$$\begin{aligned}
\sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta, \zeta)}(v, q)}{\|(v, q)\|_{V_0}} &\leq \left(\|\mathcal{R}_1(\tilde{\tau}, \kappa, \eta)\|_{L^2(Q_T)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_2(\tilde{\tau}, \zeta), \mathcal{R}_2(\tilde{\tau}, \zeta))_{L^2(Q_T)} \right. \\
&\quad + \frac{1}{\sqrt{\lambda}} (\sigma \mathcal{R}_3(\kappa, \zeta), \mathcal{R}_3(\kappa, \zeta))_{L^2(Q_T)} + \lambda \|\mathcal{R}_4(\tilde{\rho}, \chi, \zeta)\|_{L^2(Q_T)}^2 \\
&\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_5(\tilde{\rho}, \eta), \mathcal{R}_5(\tilde{\rho}, \eta))_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \mathcal{R}_6(\chi, \eta), \mathcal{R}_6(\chi, \eta))_{L^2(Q_T)} \right)^{1/2}.
\end{aligned} \tag{6.62}$$

From (6.51) follows that

$$\|e\|_{V_0} \leq \frac{1}{\mu_1} \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{B}(e, (v, q))}{\|(v, q)\|_{V_0}} = \frac{1}{\mu_1} \sup_{0 \neq (v, q) \in (H_0^{1, \frac{1}{2}}(Q_T))^2} \frac{\mathcal{F}_{(\eta, \zeta)}(v, q)}{\|(v, q)\|_{V_0}}$$

with $\mu_1 = 1/\sqrt{3}$, which leads together with (6.62) to the final estimate (6.61). \square

Analogously to Section 6.1, we now derive the majorants for the optimality equations which correspond to every single mode k .

We consider the problems (4.9), i.e.,

$$\begin{aligned}
& \int_{\Omega} \left((\mathbf{y}_k - \boldsymbol{\eta}_k) \cdot \mathbf{v}_k - \nu(\mathbf{x}) \nabla(\mathbf{p}_k - \boldsymbol{\zeta}_k) \cdot \nabla \mathbf{v}_k + k\omega \sigma(\mathbf{x})(\mathbf{p}_k - \boldsymbol{\zeta}_k) \cdot \mathbf{v}_k^\perp \right. \\
& \quad \left. + \nu(\mathbf{x}) \nabla(\mathbf{y}_k - \boldsymbol{\eta}_k) \cdot \nabla \mathbf{q}_k + k\omega \sigma(\mathbf{x})(\mathbf{y}_k - \boldsymbol{\eta}_k) \cdot \mathbf{q}_k^\perp + \lambda^{-1}(\mathbf{p}_k - \boldsymbol{\zeta}_k) \cdot \mathbf{q}_k \right) d\mathbf{x} \\
& = \int_{\Omega} \left(\mathbf{y}_{d_k} \cdot \mathbf{v}_k - \boldsymbol{\eta}_k \cdot \mathbf{v}_k + \nu(\mathbf{x}) \nabla \boldsymbol{\zeta}_k \cdot \nabla \mathbf{v}_k - k\omega \sigma(\mathbf{x}) \boldsymbol{\zeta}_k \cdot \mathbf{v}_k^\perp \right. \\
& \quad \left. - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k \cdot \nabla \mathbf{q}_k - k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k \cdot \mathbf{q}_k^\perp - \lambda^{-1} \boldsymbol{\zeta}_k \cdot \mathbf{q}_k \right) d\mathbf{x},
\end{aligned} \tag{6.63}$$

for every single mode $k = 1, \dots, N$, and, in the case $k = 0$, we obtain the variational problem (4.10), i.e.,

$$\begin{aligned}
& \int_{\Omega} \left((y_0^c - \eta_0^c) v_0^c - \nu(\mathbf{x}) \nabla(p_0^c - \zeta_0^c) \cdot \nabla v_0^c + \nu(\mathbf{x}) \nabla(y_0^c - \eta_0^c) \cdot \nabla q_0^c + \lambda^{-1}(p_0^c - \zeta_0^c) q_0^c \right) d\mathbf{x} \\
& = \int_{\Omega} \left(y_{d_0}^c v_0^c - \eta_0^c v_0^c + \nu(\mathbf{x}) \nabla \zeta_0^c \cdot \nabla v_0^c - \nu(\mathbf{x}) \nabla \eta_0^c \cdot \nabla q_0^c - \lambda^{-1} \zeta_0^c q_0^c \right) d\mathbf{x}.
\end{aligned} \tag{6.64}$$

We define the left hand sides of (6.63) and (6.64) by

$$\mathcal{B}_k((\mathbf{y}_k - \boldsymbol{\eta}_k, \mathbf{p}_k - \boldsymbol{\zeta}_k), (\mathbf{v}_k, \mathbf{q}_k)) \quad \text{and} \quad \mathcal{B}_0((y_0^c - \eta_0^c, p_0^c - \eta_0^c), (v_0^c, q_0^c)),$$

respectively, see (4.23) and (4.32). We start with the case $k = 1, \dots, N$. Let us consider again the \mathcal{P} -norm (4.26) introduced in Section 4.3, i.e.,

$$\|(\mathbf{y}_k, \mathbf{p}_k)\|_{\mathcal{P}}^2 = \|\mathbf{y}_k\|_{\mathbb{V}}^2 + \lambda^{-1} \|\mathbf{p}_k\|_{\mathbb{V}}^2$$

with

$$\|\mathbf{y}_k\|_{\mathbb{V}}^2 = \sqrt{\lambda} (\nu \nabla \mathbf{y}_k, \nabla \mathbf{y}_k)_{L^2(\Omega)} + k\omega \sqrt{\lambda} (\sigma \mathbf{y}_k, \mathbf{y}_k)_{L^2(\Omega)} + \|\mathbf{y}_k\|_{L^2(\Omega)}^2.$$

Let us compute an upper bound for the errors

$$\mathbf{e}_k := (\mathbf{y}_k - \boldsymbol{\eta}_k, \mathbf{p}_k - \boldsymbol{\zeta}_k)^T \quad \forall k = 1, \dots, N$$

in $(H_0^1(\Omega))^4$. From (4.27), it follows that

$$\sup_{0 \neq (\mathbf{v}_k, \mathbf{q}_k) \in (H_0^1(\Omega))^4} \frac{\mathcal{B}_k(\mathbf{e}_k, (\mathbf{v}_k, \mathbf{q}_k))}{\|(\mathbf{v}_k, \mathbf{q}_k)\|_{\mathcal{P}}} \geq \underline{c} \|\mathbf{e}_k\|_{\mathcal{P}} \tag{6.65}$$

with the parameter-independent constant $\underline{c} = 1/\sqrt{3}$. We denote the right-hand side of (6.63) by $\mathcal{F}_{(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k)}(\mathbf{v}_k, \mathbf{q}_k)$, i.e.,

$$\begin{aligned}
\mathcal{F}_{(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k)}(\mathbf{v}_k, \mathbf{q}_k) & = \int_{\Omega} \left(\mathbf{y}_{d_k} \cdot \mathbf{v}_k - \boldsymbol{\eta}_k \cdot \mathbf{v}_k + \nu(\mathbf{x}) \nabla \boldsymbol{\zeta}_k \cdot \nabla \mathbf{v}_k - k\omega \sigma(\mathbf{x}) \boldsymbol{\zeta}_k \cdot \mathbf{v}_k^\perp \right. \\
& \quad \left. - \nu(\mathbf{x}) \nabla \boldsymbol{\eta}_k \cdot \nabla \mathbf{q}_k - k\omega \sigma(\mathbf{x}) \boldsymbol{\eta}_k \cdot \mathbf{q}_k^\perp - \lambda^{-1} \boldsymbol{\zeta}_k \cdot \mathbf{q}_k \right) d\mathbf{x},
\end{aligned}$$

and need to find an upper bound of

$$\sup_{0 \neq (\mathbf{v}_k, \mathbf{q}_k) \in (H_0^1(\Omega))^4} \frac{\mathcal{F}_{(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k)}(\mathbf{v}_k, \mathbf{q}_k)}{\|(\mathbf{v}_k, \mathbf{q}_k)\|_{\mathcal{P}}}.$$

We introduce the functions $\tilde{\boldsymbol{\tau}}_k = (\tilde{\boldsymbol{\tau}}_k^c, \tilde{\boldsymbol{\tau}}_k^s)^T$ and $\tilde{\boldsymbol{\rho}}_k = (\tilde{\boldsymbol{\rho}}_k^c, \tilde{\boldsymbol{\rho}}_k^s)^T$ of vector-valued functions with

$$(\nu \tilde{\boldsymbol{\tau}}_k^c), (\nu \tilde{\boldsymbol{\tau}}_k^s), (\nu \tilde{\boldsymbol{\rho}}_k^c), (\nu \tilde{\boldsymbol{\rho}}_k^s) \in H(\operatorname{div}, \Omega) := \{\boldsymbol{\tau} \in [L^2(\Omega)]^d : \operatorname{div} \boldsymbol{\tau} \in L^2(\Omega)\}$$

with the weak divergence fulfilling

$$\int_{\Omega} \operatorname{div}(\nu(\mathbf{x})\tilde{\boldsymbol{\tau}}(\mathbf{x}))v(\mathbf{x})d\mathbf{x} = - \int_{\Omega} (\nu(\mathbf{x})\tilde{\boldsymbol{\tau}}(\mathbf{x})) \cdot \nabla v(\mathbf{x})d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega).$$

Moreover, we simply introduce the functions

$$\boldsymbol{\kappa}_k = (\kappa_k^c, \kappa_k^s)^T, \boldsymbol{\chi}_k = (\chi_k^c, \chi_k^s)^T \in (H_0^1(\Omega))^2,$$

which fulfill both the identity, i.e., the orthogonality relation, as follows

$$\int_{\Omega} k\omega \sigma(\mathbf{x})\boldsymbol{\kappa}_k \cdot \mathbf{v}^\perp d\mathbf{x} = - \int_{\Omega} k\omega \sigma(\mathbf{x})\boldsymbol{\kappa}_k^\perp \cdot \mathbf{v} d\mathbf{x} \quad \forall \mathbf{v} \in (C_0^\infty(\Omega))^2.$$

Due to the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathcal{F}_{(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k)}(\mathbf{v}_k, \mathbf{q}_k) &= \int_{\Omega} (\mathbf{y}_{d_k} \cdot \mathbf{v}_k - \boldsymbol{\eta}_k \cdot \mathbf{v}_k + \nu \nabla \boldsymbol{\zeta}_k \cdot \nabla \mathbf{v}_k + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}_k) \cdot \mathbf{v}_k + (\nu \tilde{\boldsymbol{\tau}}_k) \cdot \nabla \mathbf{v}_k \\ &\quad - k\omega \sigma \boldsymbol{\zeta}_k \cdot \mathbf{v}_k^\perp + k\omega \sigma \boldsymbol{\kappa}_k \cdot \mathbf{v}_k^\perp + k\omega \sigma \boldsymbol{\kappa}_k^\perp \cdot \mathbf{v}_k \\ &\quad - \nu \nabla \boldsymbol{\eta}_k \cdot \nabla \mathbf{q}_k + \operatorname{div}(\nu \tilde{\boldsymbol{\rho}}_k) \cdot \mathbf{q}_k + (\nu \tilde{\boldsymbol{\rho}}_k) \cdot \nabla \mathbf{q}_k \\ &\quad - k\omega \sigma \boldsymbol{\eta}_k \cdot \mathbf{q}_k^\perp + k\omega \sigma \boldsymbol{\chi}_k \cdot \mathbf{q}_k^\perp + k\omega \sigma \boldsymbol{\chi}_k^\perp \cdot \mathbf{q}_k - \lambda^{-1} \boldsymbol{\zeta}_k \cdot \mathbf{q}_k) d\mathbf{x}, \\ &= \int_{\Omega} ((\mathbf{y}_{d_k} - \boldsymbol{\eta}_k + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}_k) + k\omega \sigma \boldsymbol{\kappa}_k^\perp) \cdot \mathbf{v}_k + \nu (\nabla \boldsymbol{\zeta}_k + \tilde{\boldsymbol{\tau}}_k) \cdot \nabla \mathbf{v}_k \\ &\quad + k\omega \sigma (\boldsymbol{\kappa}_k - \boldsymbol{\zeta}_k) \cdot \mathbf{v}_k^\perp + (\operatorname{div}(\nu \tilde{\boldsymbol{\rho}}_k) + k\omega \sigma \boldsymbol{\chi}_k^\perp - \lambda^{-1} \boldsymbol{\zeta}_k) \cdot \mathbf{q}_k \\ &\quad + \nu (\tilde{\boldsymbol{\rho}}_k - \nabla \boldsymbol{\eta}_k) \cdot \nabla \mathbf{q}_k + k\omega \sigma (\boldsymbol{\chi}_k - \boldsymbol{\eta}_k) \cdot \mathbf{q}_k^\perp) d\mathbf{x}, \\ &\leq \|\mathcal{R}_{1k}(\boldsymbol{\eta}_k, \tilde{\boldsymbol{\tau}}_k, \boldsymbol{\kappa}_k)\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} + (\nu \mathcal{R}_{2k}(\boldsymbol{\zeta}_k, \tilde{\boldsymbol{\tau}}_k), \mathcal{R}_{2k}(\boldsymbol{\zeta}_k, \tilde{\boldsymbol{\tau}}_k))_{L^2(\Omega)}^{1/2} (\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)}^{1/2} \\ &\quad + k\omega (\sigma \mathcal{R}_{3k}(\boldsymbol{\zeta}_k, \boldsymbol{\kappa}_k), \mathcal{R}_{3k}(\boldsymbol{\zeta}_k, \boldsymbol{\kappa}_k))_{L^2(\Omega)}^{1/2} (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)}^{1/2} + \|\mathcal{R}_{4k}(\boldsymbol{\zeta}_k, \tilde{\boldsymbol{\rho}}_k, \boldsymbol{\chi}_k)\|_{L^2(\Omega)} \|\mathbf{q}_k\|_{L^2(\Omega)} \\ &\quad + (\nu \mathcal{R}_{5k}(\boldsymbol{\eta}_k, \tilde{\boldsymbol{\rho}}_k), \mathcal{R}_{5k}(\boldsymbol{\eta}_k, \tilde{\boldsymbol{\rho}}_k))_{L^2(\Omega)}^{1/2} (\nu \nabla \mathbf{q}_k, \nabla \mathbf{q}_k)_{L^2(\Omega)}^{1/2} \\ &\quad + k\omega (\sigma \mathcal{R}_{6k}(\boldsymbol{\eta}_k, \boldsymbol{\chi}_k), \mathcal{R}_{6k}(\boldsymbol{\eta}_k, \boldsymbol{\chi}_k))_{L^2(\Omega)}^{1/2} (\sigma \mathbf{q}_k, \mathbf{q}_k)_{L^2(\Omega)}^{1/2} \end{aligned}$$

with

$$\begin{aligned} \mathcal{R}_{1k}(\boldsymbol{\eta}_k, \tilde{\boldsymbol{\tau}}_k, \boldsymbol{\kappa}_k) &= \mathbf{y}_{d_k} - \boldsymbol{\eta}_k + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}_k) + k\omega \sigma \boldsymbol{\kappa}_k^\perp, \\ \mathcal{R}_{2k}(\boldsymbol{\zeta}_k, \tilde{\boldsymbol{\tau}}_k) &= \tilde{\boldsymbol{\tau}}_k + \nabla \boldsymbol{\zeta}_k, \\ \mathcal{R}_{3k}(\boldsymbol{\zeta}_k, \boldsymbol{\kappa}_k) &= \boldsymbol{\kappa}_k - \boldsymbol{\zeta}_k, \\ \mathcal{R}_{4k}(\boldsymbol{\zeta}_k, \tilde{\boldsymbol{\rho}}_k, \boldsymbol{\chi}_k) &= \operatorname{div}(\nu \tilde{\boldsymbol{\rho}}_k) + k\omega \sigma \boldsymbol{\chi}_k^\perp - \lambda^{-1} \boldsymbol{\zeta}_k, \\ \mathcal{R}_{5k}(\boldsymbol{\eta}_k, \tilde{\boldsymbol{\rho}}_k) &= \tilde{\boldsymbol{\rho}}_k - \nabla \boldsymbol{\eta}_k, \\ \mathcal{R}_{6k}(\boldsymbol{\eta}_k, \boldsymbol{\chi}_k) &= \boldsymbol{\chi}_k - \boldsymbol{\eta}_k. \end{aligned}$$

Applying the Cauchy-Schwarz inequality together with a proper weighting with λ , we get

$$\begin{aligned} \mathcal{F}_{(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k)}(\mathbf{v}_k, \mathbf{q}_k) &\leq \|\mathcal{R}_{1k}\|_{L^2(\Omega)} \|\mathbf{v}_k\|_{L^2(\Omega)} + \lambda^{-1/4} (\nu \mathcal{R}_{2k}, \mathcal{R}_{2k})_{L^2(\Omega)}^{1/2} \lambda^{1/4} (\nu \nabla \mathbf{v}_k, \nabla \mathbf{v}_k)_{L^2(\Omega)}^{1/2} \\ &\quad + \lambda^{-1/4} \sqrt{k\omega} (\sigma \mathcal{R}_{3k}, \mathcal{R}_{3k})_{L^2(\Omega)}^{1/2} \lambda^{1/4} \sqrt{k\omega} (\sigma \mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)}^{1/2} \\ &\quad + \lambda^{1/2} \|\mathcal{R}_{4k}\|_{L^2(\Omega)} \lambda^{-1/2} \|\mathbf{q}_k\|_{L^2(\Omega)} \\ &\quad + \lambda^{1/4} (\nu \mathcal{R}_{5k}, \mathcal{R}_{5k})_{L^2(\Omega)}^{1/2} \lambda^{-1/4} (\nu \nabla \mathbf{q}_k, \nabla \mathbf{q}_k)_{L^2(\Omega)}^{1/2} \\ &\quad + \lambda^{1/4} \sqrt{k\omega} (\sigma \mathcal{R}_{6k}, \mathcal{R}_{6k})_{L^2(\Omega)}^{1/2} \lambda^{-1/4} \sqrt{k\omega} (\sigma \mathbf{q}_k, \mathbf{q}_k)_{L^2(\Omega)}^{1/2} \end{aligned}$$

and

$$\begin{aligned}
\mathcal{F}_{(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k)}(\mathbf{v}_k, \mathbf{q}_k) &\leq \left(\|\mathcal{R}_{1k}\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}}(\nu\mathcal{R}_{2k}, \mathcal{R}_{2k})_{L^2(\Omega)} + \frac{1}{\sqrt{\lambda}}k\omega(\sigma\mathcal{R}_{3k}, \mathcal{R}_{3k})_{L^2(\Omega)} \right. \\
&\quad \left. + \lambda\|\mathcal{R}_{4k}\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu\mathcal{R}_{5k}, \mathcal{R}_{5k})_{L^2(\Omega)} + \sqrt{\lambda}k\omega(\sigma\mathcal{R}_{6k}, \mathcal{R}_{6k})_{L^2(\Omega)} \right)^{1/2} \\
&\quad \times \left(\|\mathbf{v}_k\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu\nabla\mathbf{v}_k, \nabla\mathbf{v}_k)_{L^2(\Omega)} + \sqrt{\lambda}k\omega(\sigma\mathbf{v}_k, \mathbf{v}_k)_{L^2(\Omega)} \right. \\
&\quad \left. + \lambda^{-1}(\|\mathbf{q}_k\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu\nabla\mathbf{q}_k, \nabla\mathbf{q}_k)_{L^2(\Omega)} + \sqrt{\lambda}k\omega(\sigma\mathbf{q}_k, \mathbf{q}_k)_{L^2(\Omega)}) \right)^{1/2}.
\end{aligned} \tag{6.66}$$

Hence, we obtain the upper bound

$$\begin{aligned}
\sup_{0 \neq (\mathbf{v}_k, \mathbf{q}_k) \in (H_0^1(\Omega))^4} \frac{\mathcal{F}_{(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k)}(\mathbf{v}_k, \mathbf{q}_k)}{\|(\mathbf{v}_k, \mathbf{q}_k)\|_{\mathcal{P}}} &\leq \left(\|\mathcal{R}_{1k}\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}}(\nu\mathcal{R}_{2k}, \mathcal{R}_{2k})_{L^2(\Omega)} \right. \\
&\quad \left. + \frac{1}{\sqrt{\lambda}}k\omega(\sigma\mathcal{R}_{3k}, \mathcal{R}_{3k})_{L^2(\Omega)} + \lambda\|\mathcal{R}_{4k}\|_{L^2(\Omega)}^2 \right. \\
&\quad \left. + \sqrt{\lambda}(\nu\mathcal{R}_{5k}, \mathcal{R}_{5k})_{L^2(\Omega)} + \sqrt{\lambda}k\omega(\sigma\mathcal{R}_{6k}, \mathcal{R}_{6k})_{L^2(\Omega)} \right)^{1/2},
\end{aligned}$$

which yields together with (6.65) the following a posteriori error result corresponding to every single mode $k = 1, \dots, N$:

Corollary 6.34. *Let $\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k \in H_0^1(\Omega)$ and the bilinear form $\mathcal{B}_k(\cdot, \cdot)$ in (6.63) satisfy (6.65). Then,*

$$\begin{aligned}
\|e_k\|_{\mathcal{P}} &\leq \sqrt{3} \left(\|\mathcal{R}_{1k}\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}}(\nu\mathcal{R}_{2k}, \mathcal{R}_{2k})_{L^2(\Omega)} + \frac{1}{\sqrt{\lambda}}k\omega(\sigma\mathcal{R}_{3k}, \mathcal{R}_{3k})_{L^2(\Omega)} \right. \\
&\quad \left. + \lambda\|\mathcal{R}_{4k}\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu\mathcal{R}_{5k}, \mathcal{R}_{5k})_{L^2(\Omega)} + \sqrt{\lambda}k\omega(\sigma\mathcal{R}_{6k}, \mathcal{R}_{6k})_{L^2(\Omega)} \right)^{1/2} \\
&=: \mathcal{M}_{\|\cdot\|_{\mathcal{P}}}^{\oplus k}(\boldsymbol{\eta}_k, \boldsymbol{\zeta}_k, \tilde{\boldsymbol{\tau}}_k, \tilde{\boldsymbol{\rho}}_k, \boldsymbol{\kappa}_k, \boldsymbol{\chi}_k),
\end{aligned} \tag{6.67}$$

where $(\nu\tilde{\boldsymbol{\tau}}_k), (\nu\tilde{\boldsymbol{\rho}}_k) \in (H(\operatorname{div}, \Omega))^2$ and $\boldsymbol{\kappa}_k, \boldsymbol{\chi}_k \in (H_0^1(\Omega))^2$.

Now, let us consider the case $k = 0$, where

$$\begin{aligned}
\mathcal{B}_0((y_0^c - \eta_0^c, p_0^c - \eta_0^c), (v_0^c, q_0^c)) &= \int_{\Omega} ((y_0^c - \eta_0^c) v_0^c - \nu(\mathbf{x})\nabla(p_0^c - \zeta_0^c) \cdot \nabla v_0^c \\
&\quad + \nu(\mathbf{x})\nabla(y_0^c - \eta_0^c) \cdot \nabla q_0^c + \lambda^{-1}(p_0^c - \zeta_0^c) q_0^c) d\mathbf{x}
\end{aligned}$$

Using the definitions

$$((y, p), (v, q))_{\mathcal{P}} = (y, v)_{L^2(\Omega)} + \sqrt{\lambda}(\nu\nabla y, \nabla v)_{L^2(\Omega)} + \lambda^{-1}((p, q)_{L^2(\Omega)} + \sqrt{\lambda}(\nu\nabla p, \nabla q)_{L^2(\Omega)})$$

and

$$\|(y, p)\|_{\mathcal{P}}^2 = \|y\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu\nabla y, \nabla y)_{L^2(\Omega)} + \lambda^{-1}(\|p\|_{L^2(\Omega)}^2 + \sqrt{\lambda}(\nu\nabla p, \nabla p)_{L^2(\Omega)})$$

of the \mathcal{P} -inner product and norm in the case $k = 0$, we proved the following inf-sup condition in Section 4.3 with inf-sup constant $\underline{c} = 1/\sqrt{2}$:

$$\sup_{0 \neq (v_0^c, q_0^c) \in (H_0^1(\Omega))^2} \frac{\mathcal{B}_0(e_0^c, (v_0^c, q_0^c))}{\|(v_0^c, q_0^c)\|_{\mathcal{P}}} \geq \underline{c} \|e_0^c\|_{\mathcal{P}} \tag{6.68}$$

for all $(y_0^c, p_0^c) \in \mathbb{V}$ with $e_0^c := (y_0^c - \eta_0^c, p_0^c - \zeta_0^c)^T$. We denote the right-hand side of (6.64) by

$$\mathcal{F}_{(\boldsymbol{\eta}_0^c, \boldsymbol{\zeta}_0^c)}(v_0^c, q_0^c) = \int_{\Omega} (y_{d0}^c v_0^c - \eta_0^c v_0^c + \nu(\mathbf{x})\nabla\zeta_0^c \cdot \nabla v_0^c - \nu(\mathbf{x})\nabla\eta_0^c \cdot \nabla q_0^c - \lambda^{-1}\zeta_0^c q_0^c) d\mathbf{x}$$

and need to find an upper bound of

$$\sup_{0 \neq (v_0^c, q_0^c) \in (H_0^1(\Omega))^2} \frac{\mathcal{F}_{(\eta_0^c, \zeta_0^c)}(v_0^c, q_0^c)}{\|(v_0^c, q_0^c)\|_{\mathcal{P}}}.$$

Again, we introduce the vector-valued functions $\tilde{\boldsymbol{\tau}}_0^c$ and $\tilde{\boldsymbol{\rho}}_0^c$, where

$$(\nu \tilde{\boldsymbol{\tau}}_0^c), (\nu \tilde{\boldsymbol{\rho}}_0^c) \in H(\operatorname{div}, \Omega) := \{\boldsymbol{\tau} \in [L^2(\Omega)]^d : \operatorname{div} \boldsymbol{\tau} \in L^2(\Omega)\}.$$

Due to the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \mathcal{F}_{(\eta_0^c, \zeta_0^c)}(v_0^c, q_0^c) &= \int_{\Omega} (y_{d0}^c v_0^c - \eta_0^c v_0^c + \nu \nabla \zeta_0^c \cdot \nabla v_0^c + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}_0^c) v_0^c + (\nu \tilde{\boldsymbol{\tau}}_0^c) \cdot \nabla v_0^c \\ &\quad + \operatorname{div}(\nu \tilde{\boldsymbol{\rho}}_0^c) q_0^c + (\nu \tilde{\boldsymbol{\rho}}_0^c) \cdot \nabla q_0^c - \nu \nabla \eta_0^c \cdot \nabla q_0^c - \lambda^{-1} \zeta_0^c q_0^c) dx \\ &= \int_{\Omega} ((y_{d0}^c - \eta_0^c + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}_0^c)) v_0^c + \nu (\tilde{\boldsymbol{\tau}}_0^c + \nabla \zeta_0^c) \cdot \nabla v_0^c \\ &\quad + \nu (\tilde{\boldsymbol{\rho}}_0^c - \nabla \eta_0^c) \cdot \nabla q_0^c + (\operatorname{div}(\nu \tilde{\boldsymbol{\rho}}_0^c) - \lambda^{-1} \zeta_0^c) q_0^c) dx \\ &\leq \|\mathcal{R}_{10}^c(\eta_0^c, \tilde{\boldsymbol{\tau}}_0^c)\|_{L^2(\Omega)} \|v_0^c\|_{L^2(\Omega)} \\ &\quad + \lambda^{-1/4} (\nu \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c), \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c))_{L^2(\Omega)}^{1/2} \lambda^{1/4} (\nu \nabla v_0^c, \nabla v_0^c)_{L^2(\Omega)}^{1/2} \\ &\quad + \lambda^{1/4} (\nu \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c), \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c))_{L^2(\Omega)}^{1/2} \lambda^{-1/4} (\nu \nabla q_0^c, \nabla q_0^c)_{L^2(\Omega)}^{1/2} \\ &\quad + \lambda^{1/2} \|\mathcal{R}_{40}^c(\zeta_0^c, \tilde{\boldsymbol{\rho}}_0^c)\|_{L^2(\Omega)} \lambda^{-1/2} \|q_0^c\|_{L^2(\Omega)} \\ &\leq \left(\|\mathcal{R}_{10}^c(\eta_0^c, \tilde{\boldsymbol{\tau}}_0^c)\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c), \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c))_{L^2(\Omega)} \right. \\ &\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c), \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c))_{L^2(\Omega)} + \lambda \|\mathcal{R}_{40}^c(\zeta_0^c, \tilde{\boldsymbol{\rho}}_0^c)\|_{L^2(\Omega)}^2 \right)^{1/2} \\ &\quad \times \left(\|v_0^c\|_{L^2(\Omega)}^2 + \sqrt{\lambda} (\nu \nabla v_0^c, \nabla v_0^c)_{L^2(\Omega)} \right. \\ &\quad \left. + \frac{1}{\sqrt{\lambda}} (\nu \nabla q_0^c, \nabla q_0^c)_{L^2(\Omega)} + \lambda^{-1} \|q_0^c\|_{L^2(\Omega)}^2 \right)^{1/2} \\ &= \left(\|\mathcal{R}_{10}^c(\eta_0^c, \tilde{\boldsymbol{\tau}}_0^c)\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c), \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c))_{L^2(\Omega)} \right. \\ &\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c), \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c))_{L^2(\Omega)} + \lambda \|\mathcal{R}_{40}^c(\zeta_0^c, \tilde{\boldsymbol{\rho}}_0^c)\|_{L^2(\Omega)}^2 \right)^{1/2} \|(v_0^c, q_0^c)\|_{\mathcal{P}} \end{aligned} \quad (6.69)$$

with

$$\begin{aligned} \mathcal{R}_{10}^c(\eta_0^c, \boldsymbol{\tau}_0^c) &= y_{d0}^c - \eta_0^c + \operatorname{div}(\nu \tilde{\boldsymbol{\tau}}_0^c), \\ \mathcal{R}_{20}^c(\zeta_0^c, \boldsymbol{\tau}_0^c) &= \tilde{\boldsymbol{\tau}}_0^c + \nabla \zeta_0^c, \\ \mathcal{R}_{30}^c(\eta_0^c, \boldsymbol{\rho}_0^c) &= \tilde{\boldsymbol{\rho}}_0^c - \nabla \eta_0^c, \\ \mathcal{R}_{40}^c(\zeta_0^c, \boldsymbol{\rho}_0^c) &= \operatorname{div}(\nu \tilde{\boldsymbol{\rho}}_0^c) - \lambda^{-1} \zeta_0^c. \end{aligned}$$

This yields the upper bound

$$\begin{aligned} \sup_{0 \neq (v_0^c, q_0^c) \in (H_0^1(\Omega))^2} \frac{\mathcal{F}_{(\eta_0^c, \zeta_0^c)}(v_0^c, q_0^c)}{\|(v_0^c, q_0^c)\|_{\mathcal{P}}} &\leq \left(\|\mathcal{R}_{10}^c(\eta_0^c, \tilde{\boldsymbol{\tau}}_0^c)\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c), \mathcal{R}_{20}^c(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c))_{L^2(\Omega)} \right. \\ &\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c), \mathcal{R}_{30}^c(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c))_{L^2(\Omega)} + \lambda \|\mathcal{R}_{40}^c(\zeta_0^c, \tilde{\boldsymbol{\rho}}_0^c)\|_{L^2(\Omega)}^2 \right)^{1/2}. \end{aligned}$$

Using (6.68) and (6.69), we finally arrive at the following upper bounds for the case $k = 0$, which correspond to the \mathcal{P} -norm:

Corollary 6.35. *Let $\eta_0^c, \zeta_0^c \in H_0^1(\Omega)$ and the bilinear form $\mathcal{B}_0(\cdot, \cdot)$ in (6.64) satisfy (6.68). Then,*

$$\begin{aligned} \|e_0^c\|_{\mathcal{P}} &\leq \sqrt{2} \left(\|\mathcal{R}_{1_0^c}(\eta_0^c, \tilde{\boldsymbol{\tau}}_0^c)\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_{2_0^c}(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c), \mathcal{R}_{2_0^c}(\zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c))_{L^2(\Omega)} \right. \\ &\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_{3_0^c}(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c), \mathcal{R}_{3_0^c}(\eta_0^c, \tilde{\boldsymbol{\rho}}_0^c))_{L^2(\Omega)} + \lambda \|\mathcal{R}_{4_0^c}(\zeta_0^c, \tilde{\boldsymbol{\rho}}_0^c)\|_{L^2(\Omega)}^2 \right)^{1/2} \\ &=: \mathcal{M}_{\|\cdot\|_{\mathcal{P}}}^{\oplus_0}(\eta_0^c, \zeta_0^c, \tilde{\boldsymbol{\tau}}_0^c, \tilde{\boldsymbol{\rho}}_0^c), \end{aligned} \quad (6.70)$$

where $(\nu \tilde{\boldsymbol{\tau}}_0^c), (\nu \tilde{\boldsymbol{\rho}}_0^c) \in H(\text{div}, \Omega)$.

Altogether, we obtain the following error majorant $\mathcal{M}_{\|\cdot\|_{V_0}}^{\oplus}(\eta, \zeta, \tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}}, \kappa, \chi)$ defined in Theorem 6.33 in the Fourier space:

Corollary 6.36. *The error majorant $\mathcal{M}_{\|\cdot\|_{V_0}}^{\oplus}(\eta, \zeta, \tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}}, \kappa, \chi)$ is given by*

$$\begin{aligned} \mathcal{M}_{\|\cdot\|_{V_0}}^{\oplus}(\eta, \zeta, \tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}}, \kappa, \chi) &= \sqrt{3} \left(\|\mathcal{R}_1(\tilde{\boldsymbol{\tau}}, \kappa, \eta)\|_{L^2(Q_T)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_2(\tilde{\boldsymbol{\tau}}, \zeta), \mathcal{R}_2(\tilde{\boldsymbol{\tau}}, \zeta))_{L^2(Q_T)} \right. \\ &\quad \left. + \frac{1}{\sqrt{\lambda}} (\sigma \mathcal{R}_3(\kappa, \zeta), \mathcal{R}_3(\kappa, \zeta))_{L^2(Q_T)} + \lambda \|\mathcal{R}_4(\tilde{\boldsymbol{\rho}}, \chi, \zeta)\|_{L^2(Q_T)}^2 \right. \\ &\quad \left. + \sqrt{\lambda} (\nu \mathcal{R}_5(\tilde{\boldsymbol{\rho}}, \eta), \mathcal{R}_5(\tilde{\boldsymbol{\rho}}, \eta))_{L^2(Q_T)} + \sqrt{\lambda} (\sigma \mathcal{R}_6(\chi, \eta), \mathcal{R}_6(\chi, \eta))_{L^2(Q_T)} \right)^{1/2} \\ &= \sqrt{3} \left(T (\|\mathcal{R}_{1_0^c}\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_{2_0^c}, \mathcal{R}_{2_0^c})_{L^2(\Omega)} + \sqrt{\lambda} (\nu \mathcal{R}_{3_0^c}, \mathcal{R}_{3_0^c})_{L^2(\Omega)} + \lambda \|\mathcal{R}_{4_0^c}\|_{L^2(\Omega)}^2) \right. \\ &\quad \left. + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{1_k}\|_{L^2(\Omega)}^2 + \frac{1}{\sqrt{\lambda}} (\nu \mathcal{R}_{2_k}, \mathcal{R}_{2_k})_{L^2(\Omega)} + \frac{1}{\sqrt{\lambda}} k\omega (\sigma \mathcal{R}_{3_k}, \mathcal{R}_{3_k})_{L^2(\Omega)} \right. \\ &\quad \left. + \lambda \|\mathcal{R}_{4_k}\|_{L^2(\Omega)}^2 + \sqrt{\lambda} (\nu \mathcal{R}_{5_k}, \mathcal{R}_{5_k})_{L^2(\Omega)} + \sqrt{\lambda} k\omega (\sigma \mathcal{R}_{6_k}, \mathcal{R}_{6_k})_{L^2(\Omega)}] \right. \\ &\quad \left. + \frac{T}{2} \sum_{k=N+1}^{\infty} \|\mathbf{y}_{d_k}\|_{L^2(\Omega)}^2 \right)^{1/2}, \end{aligned}$$

where

$$\begin{aligned} \mathcal{R}_{1_0^c} &= \mathcal{R}_{1_0^c}(\eta_0^c, \boldsymbol{\tau}_0^c) = \mathbf{y}_{d_0} - \eta_0^c + \text{div}(\nu \tilde{\boldsymbol{\tau}}_0^c), \\ \mathcal{R}_{2_0^c} &= \mathcal{R}_{2_0^c}(\zeta_0^c, \boldsymbol{\tau}_0^c) = \tilde{\boldsymbol{\tau}}_0^c + \nabla \zeta_0^c, \\ \mathcal{R}_{3_0^c} &= \mathcal{R}_{3_0^c}(\eta_0^c, \boldsymbol{\rho}_0^c) = \tilde{\boldsymbol{\rho}}_0^c - \nabla \eta_0^c, \\ \mathcal{R}_{4_0^c} &= \mathcal{R}_{3_0^c}(\zeta_0^c, \boldsymbol{\rho}_0^c) = \text{div}(\nu \tilde{\boldsymbol{\rho}}_0^c) - \lambda^{-1} \zeta_0^c. \end{aligned}$$

and

$$\begin{aligned} \mathcal{R}_{1_k} &= \mathcal{R}_{1_k}(\boldsymbol{\eta}_k, \tilde{\boldsymbol{\tau}}_k, \boldsymbol{\kappa}_k) = \mathbf{y}_{d_k} - \boldsymbol{\eta}_k + \text{div}(\nu \tilde{\boldsymbol{\tau}}_k) + k\omega \sigma \boldsymbol{\kappa}_k^{\perp}, \\ \mathcal{R}_{2_k} &= \mathcal{R}_{2_k}(\boldsymbol{\zeta}_k, \tilde{\boldsymbol{\tau}}_k) = \tilde{\boldsymbol{\tau}}_k + \nabla \boldsymbol{\zeta}_k, \\ \mathcal{R}_{3_k} &= \mathcal{R}_{3_k}(\boldsymbol{\zeta}_k, \boldsymbol{\kappa}_k) = \boldsymbol{\kappa}_k - \boldsymbol{\zeta}_k, \\ \mathcal{R}_{4_k} &= \mathcal{R}_{4_k}(\boldsymbol{\zeta}_k, \tilde{\boldsymbol{\rho}}_k, \boldsymbol{\chi}_k) = \text{div}(\nu \tilde{\boldsymbol{\rho}}_k) + k\omega \sigma \boldsymbol{\chi}_k^{\perp} - \lambda^{-1} \boldsymbol{\zeta}_k, \\ \mathcal{R}_{5_k} &= \mathcal{R}_{5_k}(\boldsymbol{\eta}_k, \tilde{\boldsymbol{\rho}}_k) = \tilde{\boldsymbol{\rho}}_k - \nabla \boldsymbol{\eta}_k, \\ \mathcal{R}_{6_k} &= \mathcal{R}_{6_k}(\boldsymbol{\eta}_k, \boldsymbol{\chi}_k) = \boldsymbol{\chi}_k - \boldsymbol{\eta}_k. \end{aligned}$$

Remark 6.37. *In order to obtain final bounds from the majorants \mathcal{M}_{*}^{\oplus} (replacing $*$ with the different seminorms and norms) in practical applications, we have to consider the construction of the approximations η and ζ of the state and adjoint state, respectively, as well as the construction of $(\boldsymbol{\tau}, \boldsymbol{\rho})$ or $(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\rho}})$ for their fluxes as already discussed in Section 6.1, see Remark 6.21.*

Remark 6.38. *In the optimal control of parabolic time-periodic problems, the construction of a so-called adaptive multiharmonic finite element method (AMhFEM) is again an important issue. More precisely, in addition to constructing an adaptive finite element method (AFEM), we can compute the finite element approximated Fourier coefficients parallel on different meshes, since the computations of the Fourier coefficients corresponding to every single mode $k = 0, 1, \dots$ are decoupled. Then, by prescribing certain bounds, we can filter out the Fourier coefficients, which are important for computing the solution of the problem. Altogether, such a AMhFEM yields adaptivity in space and time.*

6.3 Functional a posteriori estimates for cost functionals of parabolic time-periodic optimal control problems

Instead of – or in addition to – determining a posteriori estimates for the optimality systems, one may derive a posteriori estimates of the cost functional as it is done, e.g., in [59, 153].

The cost functional of our optimal control problem is given by (4.1), i.e.,

$$\begin{aligned} \mathcal{J}(y(u), u) &= \frac{1}{2} \int_0^T \int_{\Omega} (y(\mathbf{x}, t) - y_d(\mathbf{x}, t))^2 d\mathbf{x} dt + \frac{\lambda}{2} \int_0^T \int_{\Omega} (u(\mathbf{x}, t))^2 d\mathbf{x} dt \\ &= \frac{1}{2} \|y - y_d\|_{L^2(Q_T)}^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \end{aligned}$$

and in the Fourier spaces by

$$\mathcal{J}(y(u), u) = T \mathcal{J}_0(y_0^c(u_0^c), u_0^c) + \frac{T}{2} \sum_{k=1}^{\infty} \mathcal{J}_k(\mathbf{y}_k(\mathbf{u}_k), \mathbf{u}_k),$$

where

$$\begin{aligned} \mathcal{J}_0(y_0^c(u_0^c), u_0^c) &= \frac{1}{2} \int_{\Omega} (y_0^c(\mathbf{x}) - y_{d0}^c(\mathbf{x}))^2 d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} (u_0^c(\mathbf{x}))^2 d\mathbf{x} \\ &= \frac{1}{2} \|y_0^c - y_{d0}^c\|_{L^2(\Omega)}^2 + \frac{\lambda}{2} \|u_0^c\|_{L^2(\Omega)}^2 \end{aligned}$$

and

$$\begin{aligned} \mathcal{J}_k(\mathbf{y}_k(\mathbf{u}_k), \mathbf{u}_k) &= \frac{1}{2} \int_{\Omega} (\mathbf{y}_k(\mathbf{x}) - \mathbf{y}_{dk}(\mathbf{x}))^2 d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} (\mathbf{u}_k(\mathbf{x}))^2 d\mathbf{x} \\ &= \frac{1}{2} \|\mathbf{y}_k - \mathbf{y}_{dk}\|_{L^2(\Omega)}^2 + \frac{\lambda}{2} \|\mathbf{u}_k\|_{L^2(\Omega)}^2 \end{aligned}$$

are defined as in (4.17) and (4.15), respectively. Hence, we can determine majorants, i.e., upper bounds, for the cost functional $\mathcal{J}(y, u)$ by using the results of Section 6.1 immediately. Adding and subtracting η as well as applying the triangle and Friedrichs inequalities yield the estimates

$$\begin{aligned} \mathcal{J}(y(u), u) &= \frac{1}{2} \|y - y_d\|_{L^2(Q_T)}^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \\ &= \frac{1}{2} \|y - \eta + \eta - y_d\|_{L^2(Q_T)}^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \\ &\leq \frac{1}{2} (\|\eta - y_d\|_{L^2(Q_T)} + \|y - \eta\|_{L^2(Q_T)})^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \\ &\leq \frac{1}{2} (\|\eta - y_d\|_{L^2(Q_T)} + C_F \|\nabla y - \nabla \eta\|_{L^2(Q_T)})^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \end{aligned}$$

and

$$\mathcal{J}(y(u), u) \leq \frac{1}{2} \left(\|\eta - y_d\|_{L^2(Q_T)} + C_F |y - \eta|_{H^{1, \frac{1}{2}}(Q_T)} \right)^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2,$$

since

$$\|\nabla y - \nabla \eta\|_{L^2(Q_T)}^2 \leq |y - \eta|_{H^{1, \frac{1}{2}}(Q_T)}^2 = \|\nabla y - \nabla \eta\|_{L^2(Q_T)}^2 + \|\partial_t^{1/2} y - \partial_t^{1/2} \eta\|_{L^2(Q_T)}^2.$$

From Theorem 6.2, we obtain

$$\begin{aligned} \mathcal{J}(y(u), u) &\leq \frac{1}{2} \left(\|\eta - y_d\|_{L^2(Q_T)} + C_F |y - \eta|_{H^{1, \frac{1}{2}}(Q_T)} \right)^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \\ &\leq \frac{1}{2} \left(\|\eta - y_d\|_{L^2(Q_T)} + \frac{C_F}{\mu_1} (C_F \|\mathcal{R}_1(\eta, \boldsymbol{\tau}, u)\|_{L^2(Q_T)} + \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}) \right)^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2 \\ &= \frac{1}{2} \left(\|\eta - y_d\|_{L^2(Q_T)} + \frac{C_F}{\mu_1} \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)} + \frac{C_F^2}{\mu_1} \|\mathcal{R}_1(\eta, \boldsymbol{\tau}, u)\|_{L^2(Q_T)} \right)^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2, \end{aligned}$$

where $\mu_1 = \min\{\underline{\nu}, \underline{\sigma}\}$, $\boldsymbol{\tau} \in H(\operatorname{div}, Q_T)$ and

$$\mathcal{R}_1(\eta, \boldsymbol{\tau}, u) = \sigma \partial_t \eta - \operatorname{div} \boldsymbol{\tau} - u, \quad \mathcal{R}_2(\eta, \boldsymbol{\tau}) = \boldsymbol{\tau} - \nu \nabla \eta.$$

By introducing parameters $\alpha, \beta > 0$, we can reformulate the estimate such that the right-hand side is given by a quadratic functional, see, e.g., [59]. Then, we obtain the following estimate:

$$\mathcal{J}(y(u), u) \leq \mathcal{J}^\oplus(\alpha, \beta; \eta, \boldsymbol{\tau}, u) \quad \forall u \in L^2(Q_T)$$

with the majorant

$$\begin{aligned} \mathcal{J}^\oplus(\alpha, \beta; \eta, \boldsymbol{\tau}, u) &:= \frac{1 + \alpha}{2} \|\eta - y_d\|_{L^2(Q_T)}^2 + \frac{(1 + \alpha)(1 + \beta)C_F^2}{2\alpha\mu_1^2} \|\mathcal{R}_2(\eta, \boldsymbol{\tau})\|_{L^2(Q_T)}^2 \\ &\quad + \frac{(1 + \alpha)(1 + \beta)C_F^4}{2\alpha\beta\mu_1^2} \|\mathcal{R}_1(\eta, \boldsymbol{\tau}, u)\|_{L^2(Q_T)}^2 + \frac{\lambda}{2} \|u\|_{L^2(Q_T)}^2. \end{aligned} \quad (6.71)$$

Moreover, we obtain

$$\inf_{\substack{\eta \in H_{0, per}^{1,1}(Q_T), \boldsymbol{\tau} \in H(\operatorname{div}, Q_T) \\ u \in L^2(Q_T), \alpha, \beta > 0}} \mathcal{J}^\oplus(\alpha, \beta; \eta, \boldsymbol{\tau}, u) = \mathcal{J}(y(\bar{u}), \bar{u}), \quad (6.72)$$

since the infimum is attained for the optimal control \bar{u} , its corresponding state $\bar{y} = y(\bar{u})$ and its exact flux $(\nu \nabla \bar{y})$, for which \mathcal{R}_1 and \mathcal{R}_2 vanish, and for α going to zero. Hence, (6.72) states that the exact lower bound of the majorant (6.71) coincides with the optimal value of the cost functional of the optimal control problem, cf. [59]. Therefore, we have

$$\mathcal{J}(\bar{y}, \bar{u}) \leq \mathcal{J}^\oplus(\alpha, \beta; \eta, \boldsymbol{\tau}, u) \quad \forall \eta \in H_{0, per}^{1,1}(Q_T), \boldsymbol{\tau} \in H(\operatorname{div}, Q_T), u \in L^2(Q_T), \alpha, \beta > 0. \quad (6.73)$$

Now, it is easy to derive an a posteriori estimate. Let η be the multiharmonic finite element approximation y_{Nh} to the state y as, for instance, constructed in Chapter 4. Since the control u can be chosen arbitrarily in (6.71), we choose a multiharmonic finite element approximation u_{Nh} for it as well. More precisely, we can compute the multiharmonic finite element approximation u_{Nh} for the control from the multiharmonic finite element approximation p_{Nh} of the adjoint state, since $u_{Nh} = -\lambda^{-1} p_{Nh}$, by solving the optimality system as presented in Chapter 4, from which we obtain y_{Nh} as well. Hence, we arrive at the estimate

$$\mathcal{J}(\bar{y}, \bar{u}) \leq \mathcal{J}^\oplus(\alpha, \beta; y_{Nh}, \boldsymbol{\tau}, u_{Nh}) \quad \forall \boldsymbol{\tau} \in H(\operatorname{div}, Q_T), \alpha, \beta > 0. \quad (6.74)$$

Next, we have to reconstruct the flux $\boldsymbol{\tau}$, which can be done by different techniques, see Remark 6.21 and Remark 6.37. For that, we first choose the vector-valued function $\boldsymbol{\tau}$ to be some multiharmonic finite element function $\boldsymbol{\tau}_{Nh}$ as well. Then the majorant $\mathcal{J}^\oplus(\alpha, \beta; \mathbf{y}_{Nh}, \boldsymbol{\tau}_{Nh}, u_{Nh})$ is given by

$$\begin{aligned} \mathcal{J}^\oplus(\alpha, \beta; \mathbf{y}_{Nh}, \boldsymbol{\tau}_{Nh}, u_{Nh}) &= \frac{1+\alpha}{2} \|y_{Nh} - y_d\|_{L^2(Q_T)}^2 + \frac{(1+\alpha)(1+\beta)C_F^2}{2\alpha\mu_1^2} \|\mathcal{R}_2(y_{Nh}, \boldsymbol{\tau}_{Nh})\|_{L^2(Q_T)}^2 \\ &\quad + \frac{(1+\alpha)(1+\beta)C_F^4}{2\alpha\beta\mu_1^2} \|\mathcal{R}_1(y_{Nh}, \boldsymbol{\tau}_{Nh}, u_{Nh})\|_{L^2(Q_T)}^2 + \frac{\lambda}{2} \|u_{Nh}\|_{L^2(Q_T)}^2 \\ &= \frac{1+\alpha}{2} (T\|y_{0h}^c - y_{d0}^c\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|y_{kh}^c - y_{dk}^c\|_{L^2}^2 + \|y_{kh}^s - y_{dk}^s\|_{L^2}^2]) + \frac{T}{2} \sum_{k=N+1}^{\infty} [\|y_{dk}^c\|_{L^2}^2 + \|y_{dk}^s\|_{L^2}^2]) \\ &\quad + \frac{(1+\alpha)(1+\beta)C_F^2}{2\alpha\mu_1^2} (T\|\mathcal{R}_{20}^c(y_{0h}^c, \boldsymbol{\tau}_{0h}^c)\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{2k}^c(y_{kh}^c, \boldsymbol{\tau}_{kh}^c)\|_{L^2}^2 + \|\mathcal{R}_{2k}^s(y_{kh}^s, \boldsymbol{\tau}_{kh}^s)\|_{L^2}^2]) \\ &\quad + \frac{(1+\alpha)(1+\beta)C_F^4}{2\alpha\beta\mu_1^2} (T\|\mathcal{R}_{10}^c(\boldsymbol{\tau}_{0h}^c, u_{0h}^c)\|_{L^2}^2 \\ &\quad\quad\quad + \frac{T}{2} \sum_{k=1}^N [\|\mathcal{R}_{1k}^c(y_{kh}^s, \boldsymbol{\tau}_{kh}^c, u_{kh}^c)\|_{L^2}^2 + \|\mathcal{R}_{1k}^s(y_{kh}^c, \boldsymbol{\tau}_{kh}^s, u_{kh}^s)\|_{L^2}^2]) \\ &\quad + \frac{\lambda}{2} (T\|u_{0h}^c\|_{L^2}^2 + \frac{T}{2} \sum_{k=1}^N [\|u_{kh}^c\|_{L^2}^2 + \|u_{kh}^s\|_{L^2}^2]) \end{aligned}$$

with $\|\cdot\|_{L^2} = \|\cdot\|_{L^2(\Omega)}$,

$$\begin{aligned} \mathcal{R}_{10}^c(\boldsymbol{\tau}_{0h}^c, u_{0h}^c) &= \operatorname{div} \boldsymbol{\tau}_{0h}^c + u_{0h}^c, \\ \mathcal{R}_{1k}(\mathbf{y}_{kh}, \boldsymbol{\tau}_{kh}, u_{kh}) &= k\omega \boldsymbol{\sigma} \mathbf{y}_{kh}^\perp + \operatorname{div} \boldsymbol{\tau}_{kh} + u_{kh} \\ &= (-k\omega \boldsymbol{\sigma} \mathbf{y}_{kh}^s + \operatorname{div} \boldsymbol{\tau}_{kh}^c + u_{kh}^c, k\omega \boldsymbol{\sigma} \mathbf{y}_{kh}^c + \operatorname{div} \boldsymbol{\tau}_{kh}^s + u_{kh}^s)^T \\ &= (\mathcal{R}_{1k}^c(y_{kh}^s, \boldsymbol{\tau}_{kh}^c, u_{kh}^c), \mathcal{R}_{1k}^s(y_{kh}^c, \boldsymbol{\tau}_{kh}^s, u_{kh}^s))^T \end{aligned}$$

and

$$\begin{aligned} \mathcal{R}_{20}^c(y_{0h}^c, \boldsymbol{\tau}_{0h}^c) &= \boldsymbol{\tau}_{0h}^c - \nu \nabla y_{0h}^c \\ \mathcal{R}_{2k}(\mathbf{y}_{kh}, \boldsymbol{\tau}_{kh}) &= \boldsymbol{\tau}_{kh} - \nu \nabla \mathbf{y}_{kh} \\ &= (\boldsymbol{\tau}_{kh}^c - \nu \nabla y_{kh}^c, \boldsymbol{\tau}_{kh}^s - \nu \nabla y_{kh}^s)^T \\ &= (\mathcal{R}_{2k}^c(y_{kh}^c, \boldsymbol{\tau}_{kh}^c), \mathcal{R}_{2k}^s(y_{kh}^s, \boldsymbol{\tau}_{kh}^s))^T. \end{aligned}$$

Note that the computations are as straightforward as using the formulation with the optimality system.

Remark 6.39. *Since all the terms corresponding to every single mode k in the majorant \mathcal{J}^\oplus are decoupled, we arrive at some majorants \mathcal{J}_k^\oplus , for which we can, of course, introduce positive parameters α_k and β_k for every single mode k as well.*

Next, we have to reconstruct the fluxes $\boldsymbol{\tau}_{0h}^c$ and $\boldsymbol{\tau}_{kh}$ for all $k = 1, \dots, N$, which we denote by

$$\boldsymbol{\tau}_{kh} = R_h^{\text{flux}}(\nu \nabla \mathbf{y}_{kh}).$$

This can be done by various techniques as already mentioned in Remarks 6.21 and 6.37, see also [153]. Hence, we obtain the reconstructed flux

$$\boldsymbol{\tau}_{Nh} = R_h^{\text{flux}}(\nu \nabla y_{Nh}).$$

After performing a simple minimization of the majorant $\mathcal{J}^\oplus(\alpha, \beta; y_{Nh}, \tau_{Nh}, u_{Nh})$ with respect to α and β , we finally arrive at the a posteriori estimate

$$\mathcal{J}(\bar{y}, \bar{u}) \leq \mathcal{J}^\oplus(\bar{\alpha}, \bar{\beta}; y_{Nh}, \tau_{Nh}, u_{Nh}), \quad (6.75)$$

where $\bar{\alpha}$ and $\bar{\beta}$ denote the optimized positive parameters. This majorant provides a guaranteed upper bound for the cost functional. Alternatively, but more costly, we can mode-wise minimize the majorant leading to an $H(\text{div})$ -problem, see [104, 153].

Remark 6.40. *In this work, we do not consider any inequality constraints imposed neither on the control nor on the state, but inequality constraints imposed on the Fourier coefficients of the control or the state can easily be included into the multiharmonic finite element approach, see [88], and, hence, may be considered in the a posteriori error analysis of parabolic time-periodic optimal control problems as well.*

Chapter 7

Numerical results

In this chapter, we present and discuss the results of numerical experiments studying the numerical behavior of the MhFE approximations as well as of the AMLI preconditioned CG and MINRES solvers, which are proposed in this work and implemented in C++.

We investigate the practical convergence behavior with respect to the space and time discretizations as well as the efficiency and the robustness of our preconditioned MINRES solver in different settings. First, we present results on using the linear AMLI preconditioner proposed in Chapter 5 and its non-linear version, which is shortly discussed in Subsection 2.8.2. Then, we go on with some numerical experiments in a very general setting using these AMLI preconditioners as well as an AMLI preconditioner, which was presented by Kraus in [99], for different inexact realizations of our block-diagonal preconditioner in the MINRES method. A part of these numerical experiments has already been presented in the papers [89, 103, 112]. The computational domain is the unit square $\Omega = (0, 1) \times (0, 1)$ that is uniformly discretized by triangles. All computations were performed on a PC with Intel(R) Core(TM) i7-2600 CPU @3.40 GHz.

Let us start with some numerical examples for solving reaction-diffusion type problems as presented in Chapter 5 by the linear AMLI method we have investigated there as well as by its nonlinear version.

7.1 Numerical experiments for heterogeneous reaction-diffusion type problems

We present numerical results using the linear AMLI preconditioned CG algorithm presented in this work for solving heterogeneous reaction-diffusion type problems of the form (5.1), which lead to linear systems of the form (cf. also (5.3))

$$\underbrace{(K_{\nu,h} + M_{\mu,h})}_{=:A_h} u_h = \underline{f}_h,$$

after a proper finite element discretization. In this section, the right-hand side is given by A_h times the vector consisting only of ones. This yields a non-trivial problem, since we consider homogeneous Dirichlet boundary conditions. The coarsest mesh has 3×3 rectangles, and the parameter ℓ corresponds to a mesh of $3^\ell \times 3^\ell$ rectangles, where one rectangle consists of two triangular finite elements. All tables present the number of iterations for reducing the initial residual by a factor of 10^{-6} . The numerical experiments are for the following parameter settings:

- (a) no jumps in the values of ν and μ ,
- (b) jumps in the values of ν and μ on the coarse mesh.

In **Example (a)**, the system matrix is given by $A_h = K_h + M_h$, i.e., the parameters $\nu = \mu = 1$. Table 7.1 presents the number of AMLI iterations for different values of polynomial degrees ν and

Table 7.1: Number of iterations for $A_h = K_h + M_h$ using the LINEAR AMLI method with additive preconditioner for the pivot block (Example (a)).

grid	$v = 1$	$v = 2$	$v = 3$	$v = 4$	$v = 5$	$v = 6$
$\ell = 2$	20	20	20	20	20	20
$\ell = 3$	36	30	27	25	24	24
$\ell = 4$	61	39	32	28	26	25
$\ell = 5$	102	49	36	29	27	25
$\ell = 6$	170	60	39	30	27	25

on grids of different mesh size. Here, we use exactly the linear AMLI method, for which we have proved robustness and optimality in Chapter 5, and we choose the parameter $\alpha = 0.95$, see also Section 5.5. Table 7.2 presents the number of iterations using the nonlinear version of the AMLI method presented in Chapter 5 (cf. Subsection 2.8.2) for different values of v_{CG} , which denotes the number of inner CG iterations, and on grids of different mesh size. In both numerical experiments, the pivot block of the system matrix A_{11} is preconditioned by the additive preconditioner C_{11} according to Chapter 5. Comparing Tables 7.1 and 7.2 shows that the linear and the nonlinear versions of our AMLI method lead to very similar results. In the case of $v = v_{CG} = 1$ and $v = v_{CG} = 2$, we do not observe, as expected, a stabilization of the iteration numbers for increasing grid sizes. Moreover, the iteration numbers are (mildly) growing for $v = 3$ and $v = 4$ as well, which all in all accompanies our theoretical results. Following the computational times also illustrates the optimality of our AMLI method in terms of the computational complexity. For instance, the setting $v = 6$ in Table 7.1 yields 25 iterations in 0.02, 0.24 and 2.40 seconds on the $3^\ell \times 3^\ell$ grids with $\ell = 4, 5$ and 6, respectively, where more than 500.000 unknowns are involved in case of $\ell = 6$. Moreover, the setting $v_{CG} = 6$ in Table 7.2 yields 24, 24 and 23 iterations in 0.03, 0.30 and 2.92 seconds on the $3^\ell \times 3^\ell$ grids with $\ell = 4, 5$ and 6, respectively, which demonstrates again optimal complexity. All the other computational times show the same behavior.

Table 7.2: Number of iterations for $A_h = K_h + M_h$ using the NONLINEAR AMLI method with additive preconditioner for the pivot block (Example (a)).

grid	$v_{CG} = 1$	$v_{CG} = 2$	$v_{CG} = 3$	$v_{CG} = 4$	$v_{CG} = 5$	$v_{CG} = 6$
$\ell = 2$	20	20	20	20	20	20
$\ell = 3$	37	25	24	24	23	24
$\ell = 4$	64	27	24	24	24	24
$\ell = 5$	137	28	24	24	24	24
$\ell = 6$	274	29	24	24	23	23

In Table 7.3 and Table 7.4, we present the number of iterations using the linear and nonlinear versions of our AMLI method, but now the pivot block A_{11} is inverted exactly. The iteration numbers, which are presented in Table 7.3, are computed with the parameter choice $\alpha = 0.5$. Here, we want to emphasize the importance of a good parameter choice of α , which is different for different problems, e.g., inverting the pivot block exactly or using an additive preconditioner for it, see Section 5.5. The stabilization of the iteration numbers as well as the computational times are similarly behaving as observed in the computations of the numerical experiments presented in Table 7.1 and Table 7.2, where we have used the additive preconditioner for the pivot block.

Table 7.3: Number of iterations for $A_h = K_h + M_h$ using the LINEAR AMLI method inverting the pivot block exactly (Example (a)).

grid	$v = 1$	$v = 2$	$v = 3$	$v = 4$	$v = 5$	$v = 6$
$\ell = 2$	12	12	12	12	12	12
$\ell = 3$	24	18	16	16	16	15
$\ell = 4$	39	22	18	17	16	15
$\ell = 5$	61	26	19	17	16	15
$\ell = 6$	96	29	19	17	16	15

Table 7.4: Number of iterations for $A_h = K_h + M_h$ using the NONLINEAR AMLI method inverting the pivot block exactly (Example (a)).

grid	$v_{CG} = 1$	$v_{CG} = 2$	$v_{CG} = 3$	$v_{CG} = 4$	$v_{CG} = 5$	$v_{CG} = 6$
$\ell = 2$	12	12	12	12	12	12
$\ell = 3$	24	16	15	15	15	15
$\ell = 4$	43	17	15	15	15	14
$\ell = 5$	67	18	15	14	14	14
$\ell = 6$	162	18	15	14	14	14

In **Example (b)**, we allow jumps in the values of the coefficients ν and μ . More precisely, $\nu = 10^{-4}$ and $\mu = 1$ on subdomain $\Omega_1 = (0, 1) \times (0, \frac{1}{3})$ and $\nu = 10^4$ and $\mu = 10^{-4}$ on $\Omega_2 = \Omega \setminus \overline{\Omega_1}$. Hence, the jumps correspond to the partitioning of the coarse mesh, which consists of 3×3 rectangles in the computational domain $\Omega = (0, 1) \times (0, 1)$. This is a requirement in order to prove the theoretical results obtained in Chapter 5, cf. (5.4). Table 7.5 and Table 7.6 present the number of iterations using the linear and nonlinear versions of our AMLI method, where the pivot block of the system matrix A_{11} is preconditioned by the additive preconditioner C_{11} according to Chapter 5. In Table 7.7 and Table 7.8, we present the number of AMLI iterations, but now the pivot block A_{11} is inverted exactly. The iteration numbers, which are presented in Table 7.5 and Table 7.7, are computed with the parameter choices $\alpha = 0.95$ and $\alpha = 0.5$, respectively, which are exactly as in the example without considering jumps in the values of the coefficients ν and μ . The stabilization of the iteration numbers is similarly behaving as observed in the tables of Example (a), i.e., Tables 7.1, 7.2, 7.3 and 7.4. Moreover, the computational times are similar to the ones of Example (a), e.g., the setting $v = 6$ in

Table 7.5: Number of iterations for $K_{\nu,h} + M_{\mu,h}$ using the LINEAR AMLI method with additive preconditioner for the pivot block and jumps in ν and μ on the coarse mesh (Example (b)).

grid	$v = 1$	$v = 2$	$v = 3$	$v = 4$	$v = 5$	$v = 6$
$\ell = 2$	22	22	22	22	22	22
$\ell = 3$	48	36	32	30	29	28
$\ell = 4$	75	47	37	28	27	25
$\ell = 5$	108	50	36	29	27	25
$\ell = 6$	180	60	39	30	27	25

Table 7.6: Number of iterations for $K_{\nu,h} + M_{\mu,h}$ using the NONLINEAR AMLI method with additive preconditioner for the pivot block and jumps in ν and μ on the coarse mesh (Example (b)).

grid	$v_{CG} = 1$	$v_{CG} = 2$	$v_{CG} = 3$	$v_{CG} = 4$	$v_{CG} = 5$	$v_{CG} = 6$
$\ell = 2$	22	22	22	22	22	22
$\ell = 3$	51	31	27	26	26	26
$\ell = 4$	87	30	24	24	24	24
$\ell = 5$	141	29	24	24	24	24
$\ell = 6$	280	29	24	24	24	24

Table 7.7: Number of iterations for $K_{\nu,h} + M_{\mu,h}$ using the LINEAR AMLI method inverting the pivot block exactly and jumps in ν and μ on the coarse mesh (Example (b)).

grid	$v = 1$	$v = 2$	$v = 3$	$v = 4$	$v = 5$	$v = 6$
$\ell = 2$	16	16	16	16	16	16
$\ell = 3$	31	23	21	20	18	18
$\ell = 4$	51	27	18	17	16	15
$\ell = 5$	70	27	19	17	16	15
$\ell = 6$	102	27	19	17	16	15

Table 7.5 yields 25 iterations in 0.02, 0.24 and 2.40 seconds on the $3^\ell \times 3^\ell$ grids with $\ell = 4, 5$ and 6, respectively, where again more than 500.000 unknowns are involved in case of $\ell = 6$. Hence, the numerical results come along with our theory having a robust solver of optimal complexity, if the jumps in the values of the coefficients ν and μ correspond to the partitioning of the coarse mesh, see Chapter 5.

Table 7.8: Number of iterations for $K_{\nu,h} + M_{\mu,h}$ using the NONLINEAR AMLI method inverting the pivot block exactly and jumps in ν and μ on the coarse mesh (Example (b)).

grid	$v_{CG} = 1$	$v_{CG} = 2$	$v_{CG} = 3$	$v_{CG} = 4$	$v_{CG} = 5$	$v_{CG} = 6$
$\ell = 2$	16	16	16	16	16	16
$\ell = 3$	31	20	17	17	17	17
$\ell = 4$	56	18	15	15	15	15
$\ell = 5$	82	17	15	15	14	14
$\ell = 6$	167	18	14	14	14	14

Altogether, we have presented first numerical results using the linear AMLI preconditioned CG algorithm in case of a 3-refinement together with the additive preconditioner for the pivot block of the two-by-two splitting, which has been all discussed in Chapter 5. This linear AMLI preconditioned CG algorithm leads to a robust solver of optimal complexity for heterogeneous reaction-diffusion type problems including problems with jumps in the values of the reaction and diffusion coefficients μ and ν on the coarse mesh.

In the next section, we present numerical results for solving the saddle point systems presented especially in Chapter 4 by the AMLI preconditioned MINRES method, cf. Section 5.6.

7.2 Numerical experiments for saddle point systems

In this section, we start with two numerical examples on solving saddle point systems by the preconditioned MINRES method with the AMLI preconditioner proposed in this work. Then, we present more numerical experiments for studying the numerical behavior of the MhFE approximations, where we use the AMLI preconditioner presented by Kraus in [99] together with the preconditioned MINRES method as solver.

Numerical experiments on a linear AMLI preconditioned MINRES method

We present numerical results on solving the optimal control problem (4.1)-(4.2), i.e.,

$$\min_{y,u} \mathcal{J}(y,u) := \frac{1}{2} \int_0^T \int_{\Omega} [y(\mathbf{x},t) - y_d(\mathbf{x},t)]^2 d\mathbf{x} dt + \frac{\lambda}{2} \int_0^T \int_{\Omega} [u(\mathbf{x},t)]^2 d\mathbf{x} dt$$

subject to the parabolic time-periodic BVP (3.1)-(3.3), which has been presented in Chapter 4, by the AMLI preconditioned MINRES method using the block-diagonal preconditioners \mathcal{P} together with the linear AMLI solver proposed in Chapter 5. The MhFE discretization of the variational problem leads to the linear saddle point systems (4.11) and (4.12), e.g.,

$$\begin{pmatrix} M_h & 0 & -K_{h,\nu} & k\omega M_{h,\sigma} \\ 0 & M_h & -k\omega M_{h,\sigma} & -K_{h,\nu} \\ -K_{h,\nu} & -k\omega M_{h,\sigma} & -\lambda^{-1}M_h & 0 \\ k\omega M_{h,\sigma} & -K_{h,\nu} & 0 & -\lambda^{-1}M_h \end{pmatrix} \begin{pmatrix} \underline{y}_k^c \\ \underline{y}_k^s \\ \underline{p}_k^c \\ \underline{p}_k^s \end{pmatrix} = \begin{pmatrix} \underline{y}_{dk}^c \\ \underline{y}_{dk}^s \\ 0 \\ 0 \end{pmatrix}$$

for $k = 1, 2, \dots, N$. The numerical experiments are designed for the following settings:

1. the desired state is periodic and analytic in time,
2. the desired state is a characteristic function in space and time and there are jumps in ν and σ .

In both settings, the desired states are not time-harmonic and unreachable, and, hence, their Fourier coefficients have to be computed for different modes k in order to expand the desired states into Fourier series. Again, the coarsest mesh has 3×3 elements, where one element consists of two triangular finite elements. In the first example, the material coefficients are $\nu = \sigma = 1$, whereas jumping coefficients are considered in the second example. Table 7.9 and Table 7.10 present the number of MINRES iteration steps for different values of the cost parameter λ computing the solution for one mode $k = 1$ stopping all iterations after reducing the initial residual by a factor of 10^{-6} together with 15 inner AMLI iterations in each MINRES iteration step. The results for all other modes k are similar and can be computed in parallel. In these two examples, we use stabilization polynomials up to the degree $v = 5$ and the parameter choice $\alpha = 0.9$. The tables include numerical results computed on a 729×729 grid resulting in a linear system with more than 2.000.000 unknowns.

Table 7.9: Number of MINRES iterations using the linear AMLI preconditioned MINRES method for different values of λ (Example 1).

grid / λ	10^{-8}	10^{-6}	10^{-4}	10^{-2}	1	10^2	10^4	10^6	10^8
27×27	20	20	18	12	8	6	6	6	6
81×81	22	18	16	12	8	6	6	6	6
243×243	22	20	16	12	8	6	6	6	8
729×729	22	20	18	12	8	6	6	6	8

In **Example 1**, we consider a time-periodic and time-analytic, but not time-harmonic, desired state

$$y_d(\mathbf{x}, t) = e^t \sin(t) \left((3 + 4\pi^4) \sin^2(t) - 6 \cos^2(t) - 6 \sin(t) \cos(t) \right) \sin(x_1 \pi) \sin(x_2 \pi)$$

with $T = 2\pi$ and $\omega = 1$. Table 7.9 presents the number of MINRES iteration steps on grids of different mesh size varying the parameter λ . The numerical results accompany our theoretical results having a robust solver of optimal complexity. For instance, the parameter setting $\lambda = 1$ yields 8 iterations in 0.06, 0.48, 4.81 and 47.04 seconds on the 27×27 , 81×81 , 243×243 and 729×729 grids, respectively.

In **Example 2**, we consider a desired state

$$y_d(\mathbf{x}, t) = \chi_{[\frac{1}{4}, \frac{3}{4}]}(t) \chi_{[\frac{1}{3}, \frac{2}{3}]^2}(\mathbf{x})$$

with $T = 1$ and $\omega = 2\pi$, which is a characteristic function in space and time. In addition, we allow jumps in the values of the material coefficients ν and σ . More precisely, $\nu = 10^{-2}$ and $\sigma = 1$ on subdomain $\Omega_1 = (0, 1) \times (0, \frac{1}{3})$ and $\nu = 1$ and $\sigma = 10^2$ on $\Omega_2 = \Omega \setminus \overline{\Omega_1}$. Table 7.10 presents the number of MINRES iteration steps of this example for different values of λ and on grids of different mesh size. The computational times are similarly behaving as in Example 1. For instance, the parameter setting $\lambda = 10^4$ yields 14 iterations in 0.09, 0.80, 8.01 and 78.38 seconds on the 27×27 , 81×81 , 243×243 and 729×729 grids, respectively.

Table 7.10: Number of MINRES iterations using the linear AMLI preconditioned MINRES method for different values of λ (Example 2).

grid / λ	10^{-8}	10^{-6}	10^{-4}	10^{-2}	1	10^2	10^4	10^6	10^8
27×27	12	17	14	16	14	14	14	16	18
81×81	13	18	16	16	13	12	14	16	20
243×243	17	18	16	14	12	12	14	16	20
729×729	15	20	14	12	10	12	14	16	18

In summary, the numerical results confirm our theoretical findings, demonstrating the robustness and optimal complexity of the linear AMLI preconditioned MINRES solver. This theoretical statement is proven in Chapter 5.

Numerical experiments on a nonlinear AMLI preconditioned MINRES method

In the following five examples, we use the AMLI preconditioner proposed by Kraus in [99] for an inexact realization of our block-diagonal preconditioner in the MINRES method. We consider again the optimal control problem (4.1)-(4.2), which leads, after the MhFE discretization, to the saddle point systems (4.11) and (4.12). We present numerical results for the following very general settings:

3. the desired state is time-harmonic,
4. the desired state is periodic and analytic in time, but not time-harmonic,
5. the desired state is analytic in time, but not time-periodic,
6. the desired state is a characteristic function in space and time, and
7. the desired state is a characteristic function in space and time, but in addition, there are jumps in the coefficients ν and σ .

We mention here that the desired state is unreachable and not time-harmonic in the last four examples. Therefore, we have to compute their Fourier coefficients for different modes k in order to expand the desired states into Fourier series. The material coefficients are supposed to be piecewise constant on Ω . In Examples 3 - 6, $\sigma = \nu = 1$, whereas jumping material coefficients are considered in Example 7. Example 6 has been borrowed from [1], but with homogeneous Dirichlet boundary conditions instead of Robin conditions. We mention that, in all tables where the number of MINRES iterations is presented, the iteration was stopped after reducing the initial residual by a factor of 10^{-6} . For the inexact version of the preconditioned MINRES method, we have used the AMLI preconditioner according to [99] in each MINRES iteration step. In Example 3, we use the AMLI preconditioner with 4 inner iterations, whereas we only use 2 inner AMLI iteration steps in the latter four examples. The computations for the figures of all examples were obtained on a 64×64 grid. The results on grids of smaller mesh sizes were very similar.

In **Example 3**, we consider a time-harmonic desired state given by the formula

$$y_d(\mathbf{x}, t) = 2(4 - x_1x_2(x_2 - 1) + x_1^2x_2(x_2 - 1))(\cos(\omega t) + \sin(\omega t)).$$

For this example, we present numerical results for both the exact preconditioner \mathcal{P} in connection with a direct solver for the preconditioning equations with the system matrix D from \mathcal{P} and an inexact preconditioner according to [99]. In order to study the robustness of our preconditioners, we have performed numerical experiments for several parameter settings. In particular, we have varied the values of the parameters ω and λ .

Table 7.11: Number of MINRES iterations for different values of ω and λ on a 64×64 grid using the EXACT version of the preconditioner \mathcal{P} (Example 3).

λ / ω	10^{-8}	10^{-6}	10^{-4}	10^{-2}	1	10^2	10^4	10^6	10^8
10^{-8}	18	18	18	18	18	24	12	4	4
10^{-6}	18	18	18	18	18	20	12	4	4
10^{-4}	14	14	14	14	14	14	12	4	4
10^{-2}	8	8	8	8	10	14	12	4	4
1	6	6	6	6	10	14	12	4	4
10^2	4	4	4	6	10	14	12	4	4
10^4	4	4	4	4	10	14	12	4	4
10^6	4	4	4	4	10	14	12	4	4
10^8	2	2	4	4	10	14	12	4	4

Table 7.11 presents the number of MINRES iterations using the exact preconditioner. The theoretical bound for reducing the residual of the MINRES method by a factor of 10^{-6} lies at 24 iterations for the exact version. The numerical results for the inexact version are presented in Table 7.12 and Table 7.13, where we have used the AMLI preconditioner with 4 inner iterations according to Kraus [99]. In Table 7.11 and Table 7.12, we have computed the solutions on a 64×64 grid. Table 7.13 presents the numerical results using the AMLI preconditioner on a 512×512 grid where more than 1.000.000 unknowns are involved.

Again, the numerical experiments impressively confirm our theoretical results. Both the exact and the inexact versions lead to a parameter independent bound for the convergence rate respectively iteration numbers of the corresponding preconditioned MINRES solver. However, we observe some dependence of the preconditioned MINRES iterations on parameters λ and ω if we vary them in the large range from 10^{-8} to 10^8 . This behavior is not surprising since the parameters λ and ω change the system matrix and the preconditioner at the same time. This becomes clear for the limit cases where λ and ω ($k\omega$) tend to ∞ or 0. Tables 7.12 and 7.13 show the AMLI preconditioner is not affected nor by mesh refinement neither by parameter variations. The comparison of the latter two

Table 7.12: Number of MINRES iterations for different values of ω and λ on a 64×64 grid using the INEXACT version of the preconditioner \mathcal{P} with 4 inner AMLI iteration steps (Example 3).

λ / ω	10^{-8}	10^{-6}	10^{-4}	10^{-2}	1	10^2	10^4	10^6	10^8
10^{-8}	18	18	18	18	18	24	12	4	4
10^{-6}	18	18	18	18	18	20	12	4	4
10^{-4}	14	14	14	14	14	14	12	4	4
10^{-2}	8	8	8	8	10	14	12	4	4
1	6	6	6	6	10	14	12	4	4
10^2	4	4	4	6	10	14	12	4	4
10^4	4	4	4	4	10	14	12	4	4
10^6	4	4	4	4	10	14	12	4	4
10^8	2	2	4	4	10	14	12	4	4

Table 7.13: Number of MINRES iterations for different values of ω and λ on a 512×512 grid using the INEXACT version of the preconditioner \mathcal{P} with 4 inner AMLI iteration steps (Example 3).

λ / ω	10^{-8}	10^{-6}	10^{-4}	10^{-2}	1	10^2	10^4	10^6	10^8
10^{-8}	23	23	23	23	23	26	14	10	4
10^{-6}	20	20	20	20	20	22	12	10	4
10^{-4}	14	14	14	14	14	14	12	10	4
10^{-2}	8	8	8	8	10	14	12	10	4
1	6	6	6	6	10	14	12	10	4
10^2	4	4	4	6	10	14	12	10	4
10^4	4	4	4	4	10	14	12	10	4
10^6	4	4	4	4	10	14	12	10	4
10^8	4	4	4	4	10	14	12	10	4

tables with Table 7.11 show that the AMLI preconditioned MINRES solver is almost as good as the exact version with respect to the iteration numbers and, in contrast to the exact version, has optimal complexity.

In **Example 4**, we consider the same desired state as in Example 1, which is time-periodic and analytic, but not time-harmonic, i.e.,

$$y_d(\mathbf{x}, t) = e^t \sin(t) \left((3 + 4\pi^4) \sin^2(t) - 6 \cos^2(t) - 6 \sin(t) \cos(t) \right) \sin(x_1\pi) \sin(x_2\pi),$$

where $T = 2\pi/\omega$ with $\omega = 1$. The Fourier coefficients of the Fourier series expansion of the desired state y_d in time can be computed analytically. We truncate the Fourier series at an index N and approximate the Fourier coefficients by finite element functions. Finally, we solve the systems (4.11) and (4.12) for all $0 \leq k \leq N$. Figure 7.1 illustrates the fast convergence of the MhFE approximations $y_{Nh}(0.5, 0.5, t)$ to the exact solution $y(0.5, 0.5, t) = e^t \sin(t)^3$ at the spatial coordinates $(0.5, 0.5)$ for $\lambda = 1$ and $t \in [0, T]$ for increasing N . Already y_{5h} provides a very good approximation to the exact solution y . In Figure 7.2 (left), we illustrate the exact desired state y_d and the MhFE approximations y_{Nh} to the state y for different values of λ , more precisely, for $\lambda \in \{1, 10^{-2}, 10^{-4}, 10^{-6}\}$, as functions of time in $[0, T]$ at the spatial coordinates $\mathbf{x} = (0.5, 0.5)$ and for $N = 5$. Figure 7.2 (right) presents the corresponding controls u_{Nh} .

Table 7.14 presents the iteration numbers, and Table 7.15 the computational times for the mode $k = 1$ obtained on grids of different mesh sizes. We present the iteration numbers and the computational

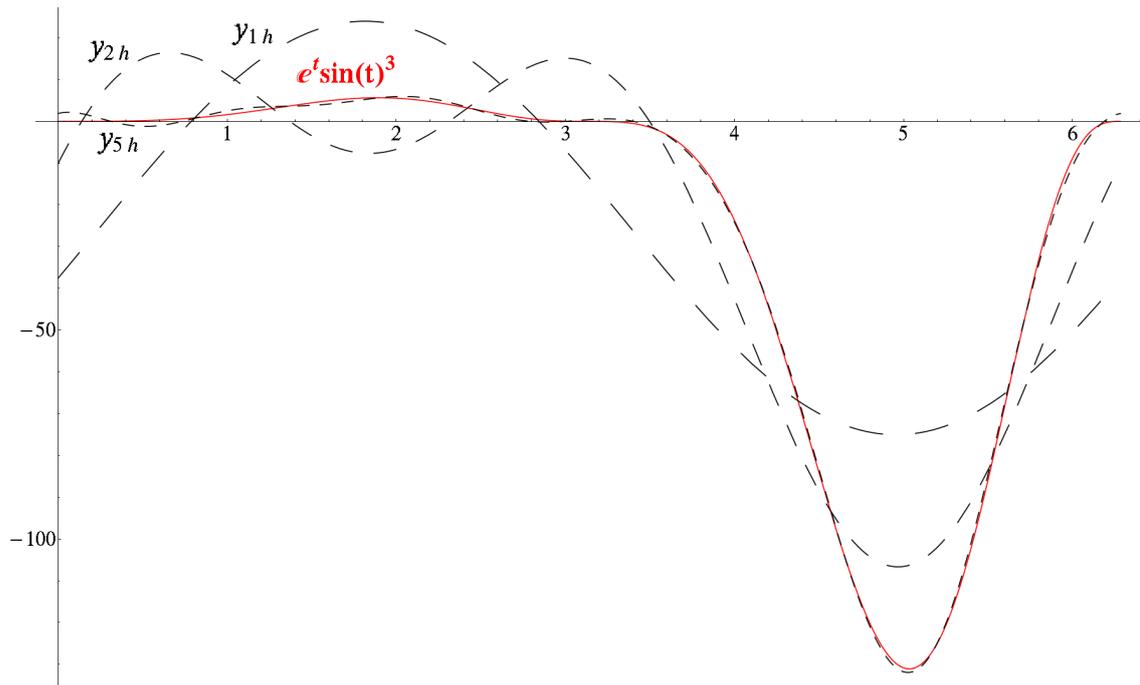


Figure 7.1: The exact state y (red) and its MhFE approximations y_{N_h} for $N = 1, 2, 5$ as functions of time in $[0, 2\pi]$ at the spatial coordinates $(0.5, 0.5)$, and for $\lambda = \omega = \sigma = \nu = 1$ (Example 4).

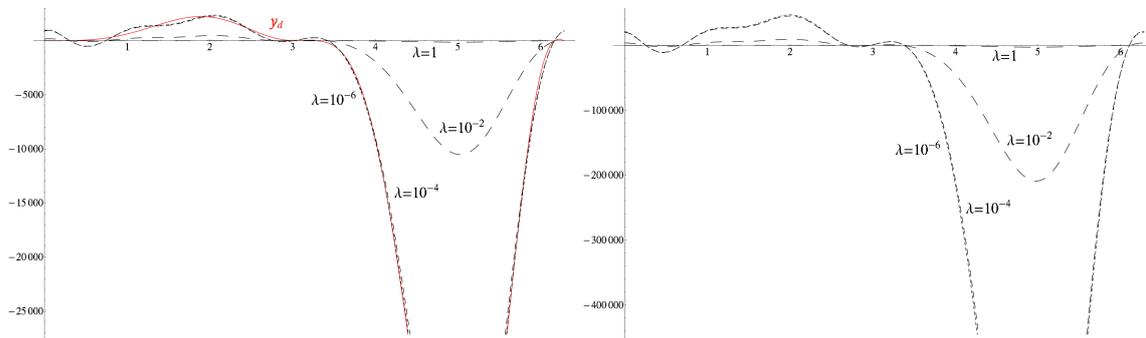


Figure 7.2: The desired state y_d (red) and the MhFE approximations y_{5h} to the state y (**left**) and the MhFE approximations u_{5h} to the control u (**right**) for $\lambda = 1, 10^{-2}, 10^{-4}, 10^{-6}$ as functions of time in $[0, 2\pi]$ at the spatial coordinates $(0.5, 0.5)$, and for $\omega = \sigma = \nu = 1$ (Example 4).

times only for $k = 1$ because the computations of all modes up to the truncation index N can be done totally in parallel and lead to similar results.

Altogether, Figure 7.1 and Figure 7.2 as well as Table 7.14 and Table 7.15 confirm that the MhFEM is a very efficient approach for solving time-periodic problems. In the following, we will consider an example, where the desired state is not time-periodic anymore.

In **Example 5**, we choose the time-analytic desired state

$$y_d(\mathbf{x}, t) = e^t(-2 \cos(t) + \sin(t) + 4\pi^4 \sin(t)) \sin(x_1\pi) \sin(x_2\pi),$$

which is obviously not time-periodic. We set again $\omega = 1$. Hence, the time period $T = 2\pi/\omega$ is equal to 2π . For this example, we can compute the Fourier coefficients of the desired state again analytically. In Figure 7.3, we present the exact state y and its MhFE approximations y_{N_h} with

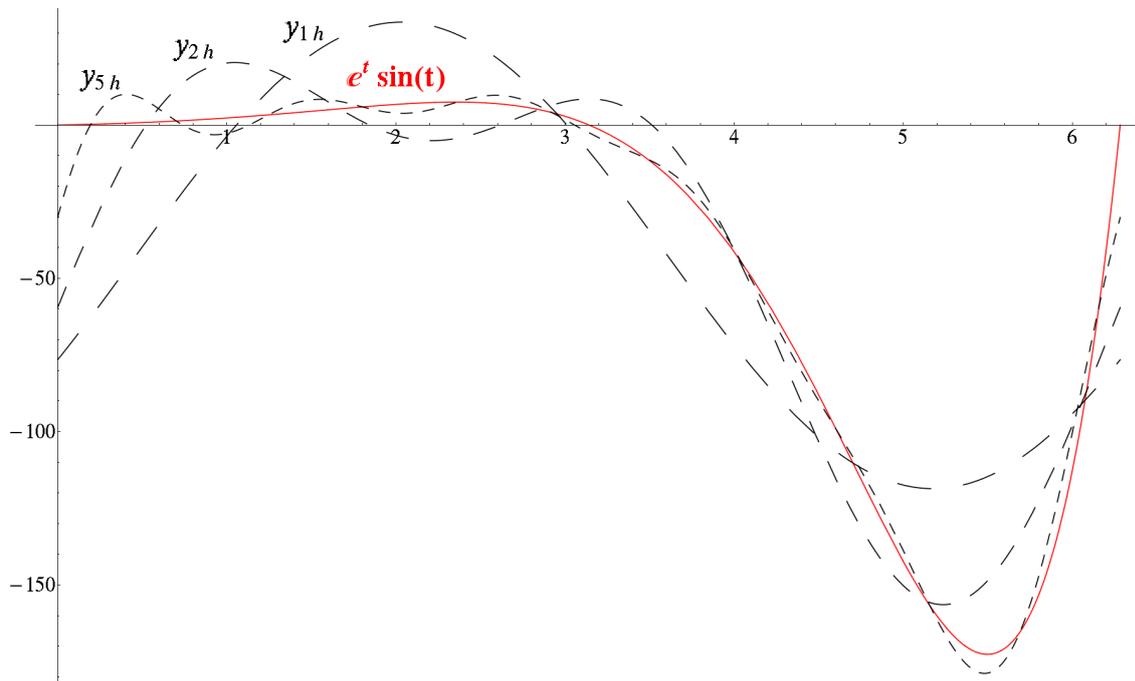
Table 7.14: Number of MINRES iterations for different λ on different grids (Example 4).

grid / λ	10^{-8}	10^{-6}	10^{-4}	10^{-2}	1	10^2	10^4	10^6	10^8
64×64	19	16	14	12	12	12	12	12	12
128×128	18	16	14	12	14	13	14	14	14
256×256	18	16	14	13	16	17	15	15	15
512×512	18	17	14	16	18	21	41	35	33

Table 7.15: The CPU times in seconds for different values of λ and on different grids (Example 4).

grid / λ	10^{-8}	10^{-6}	10^{-4}	10^{-2}	1	10^2	10^4	10^6	10^8
64×64	0.19	0.16	0.14	0.12	0.12	0.12	0.13	0.12	0.13
128×128	0.88	0.80	0.71	0.59	0.69	0.64	0.70	0.70	0.70
256×256	4.02	3.59	3.15	2.93	3.61	3.80	3.37	3.38	3.37
512×512	17.20	16.29	13.51	15.38	17.34	20.02	38.53	33.08	31.16

$N = 1, 2, 5$ as functions of time in $[0, T]$ at the spatial coordinates $(0.5, 0.5)$ and for the parameter choice $\lambda = 1$. In this case, the exact state is given by $y(0.5, 0.5, t) = e^t \sin(t)$. As for the previous

Figure 7.3: The exact state y (red) and its MhFE approximations y_{Nh} with $N = 1, 2, 5$ as functions of time in $[0, 2\pi]$ at the spatial coordinates $(0.5, 0.5)$, and for $\lambda = \sigma = \nu = 1$ (Example 5).

example, we illustrate in Figure 7.4 again how the MhFE approximations y_{5h} of the state y approach the desired state y_d as λ goes to zero. We obtain almost the same results for the number of MINRES iterations and for the computational times as have been shown for Example 4 in Tables 7.14 and 7.15,

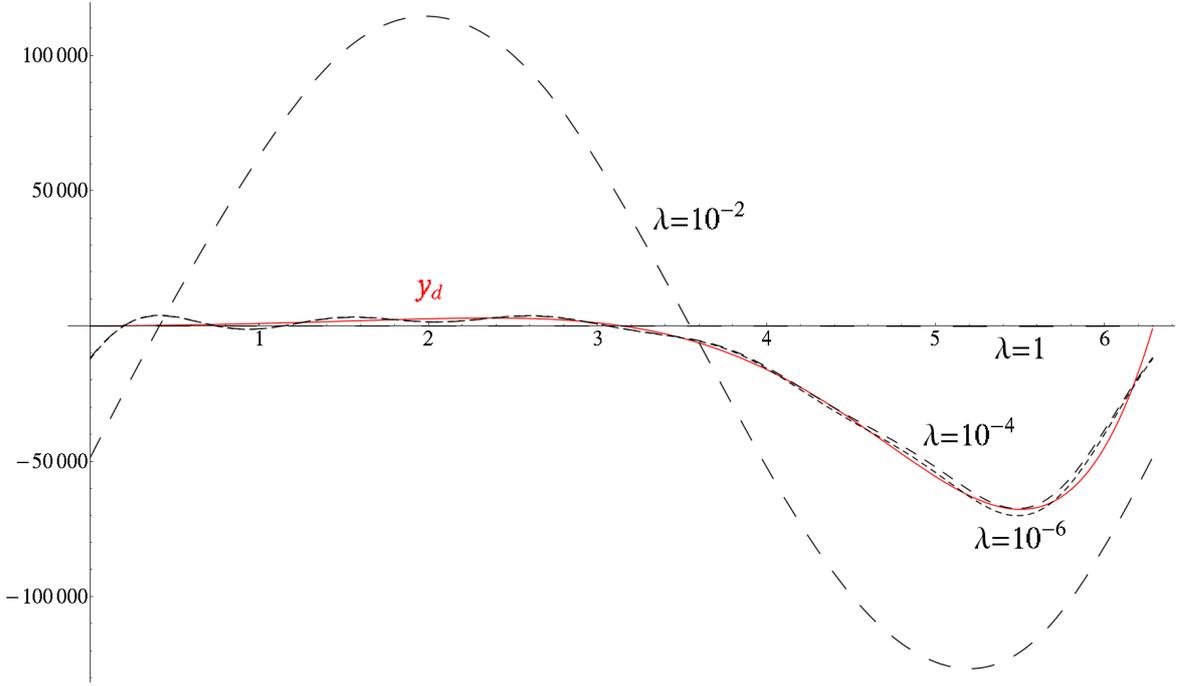


Figure 7.4: The desired state y_d (red) and the MhFE approximation y_{5h} to the state y for $\lambda = 1, 10^{-2}, 10^{-4}, 10^{-6}$ as functions of time in $[0, 2\pi]$ at the spatial coordinates $(0.5, 0.5)$, and for $\omega = \nu = \sigma = 1$ (Example 5).

i.e., the solver is not only robust with respect to λ and h , but also of optimal complexity.

In **Example 6**, we consider the desired state

$$y_d(\mathbf{x}, t) = \chi_{[\frac{1}{4}, \frac{3}{4}]}(t) \chi_{[\frac{1}{2}, 1]^2}(\mathbf{x}),$$

that is a characteristic function in space and time. The time period T is set to 1, hence $\omega = 2\pi$. We can again compute the Fourier coefficients of the desired state analytically. Indeed, we get

$$y_{dk}^c(\mathbf{x}) = \frac{2}{T} \int_0^T y_d(\mathbf{x}, t) \cos(k\omega t) dt = \chi_{[\frac{1}{2}, 1]^2}(\mathbf{x}) \frac{1}{k\pi} (-\sin(\frac{k\pi}{2}) + \sin(\frac{3k\pi}{2}))$$

and

$$y_{dk}^s(\mathbf{x}) = \frac{2}{T} \int_0^T y_d(\mathbf{x}, t) \sin(k\omega t) dt = \chi_{[\frac{1}{2}, 1]^2}(\mathbf{x}) \frac{2}{k\pi} \sin(\frac{k\pi}{2}) \sin(k\pi) = 0$$

for all $k \in \mathbb{N}$, and $y_{d0}^c(\mathbf{x}) = \chi_{[\frac{1}{2}, 1]^2}(\mathbf{x})$ in the case when $k = 0$. Figure 7.5 presents the MhFE approximations y_{Nh} to the state y for $N = 5$ as functions of time at the spatial coordinates $(0.5, 0.5)$ for different values of λ . In Figure 7.6, we illustrate the approximations u_{Nh} to the control u for $N = 5$ and $N = 11$, and for different spatial coordinates, more precisely, for $(0.25, 0.25)$, $(0.5, 0.5)$ and $(0.75, 0.75)$, where we set $\lambda = 0.01$ for all cases. Finally, we refer to Tables 7.16 and 7.17 where the number of MINRES iterations and the computational times in seconds are presented for different regularization parameters λ and mesh sizes h in comparison with the results for the last example.

In **Example 7**, we consider again a desired state which is a characteristic function in space and time, but in addition, we allow jumps in the values of the material coefficients ν and σ . More precisely, $\nu = 10^{-4}$ and $\sigma = 1$ on the subdomain $\Omega_1 = (0, 1) \times (0, \frac{1}{2})$, and $\nu = 10^4$ and $\sigma = 10^2$ on $\Omega_2 = \Omega \setminus \overline{\Omega}_1$.

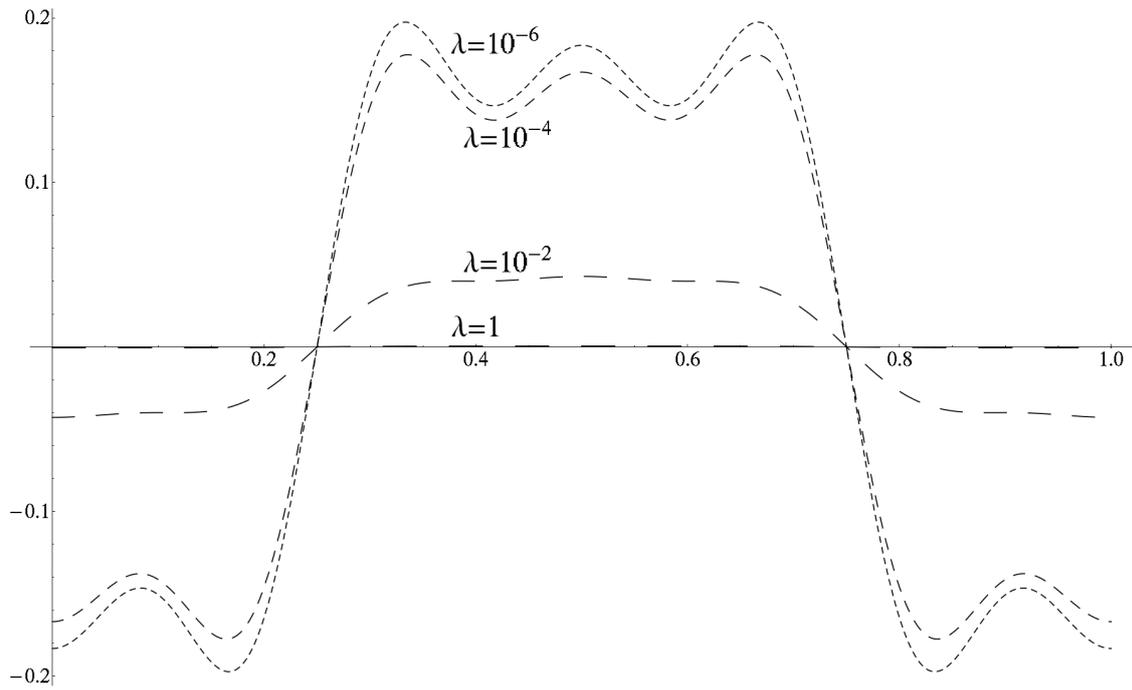


Figure 7.5: The MhFE approximations y_{5h} to the state y for $\lambda = 1, 10^{-2}, 10^{-4}, 10^{-6}$ as functions of time in $[0, 1]$ at the spatial coordinates $(0.5, 0.5)$, and for $\omega = 2\pi$, $\nu = \sigma = 1$ (Example 6).

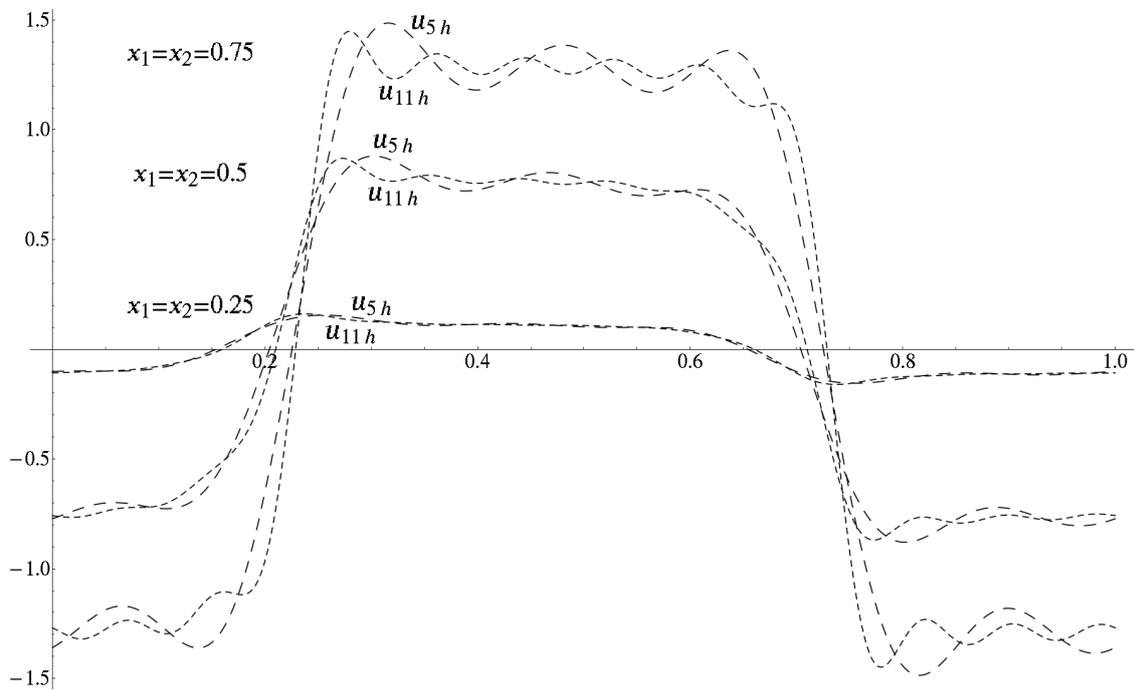


Figure 7.6: The MhFE approximations u_{5h} and u_{11h} at the spatial coordinates $(0.25, 0.25)$, $(0.5, 0.5)$ and $(0.75, 0.75)$ as functions of time in $[0, 1]$ for $\lambda = 0.01$, $\omega = 2\pi$, $\nu = \sigma = 1$ (Example 6).

The time period is set $T = 1$, hence $\omega = 2\pi$. We again vary the regularization parameter λ , and compute the solutions on grids of different mesh size. The desired state

$$y_d(\mathbf{x}, t) = \chi_{[\frac{1}{4}, \frac{3}{4}]}(t) \chi_{[\frac{1}{2}, 1]^2}(\mathbf{x})$$

is chosen as in the preceding example. We again expand the desired state in a Fourier series, where the Fourier coefficients can be computed analytically. We truncate then the Fourier series and approximate the Fourier coefficients by finite element functions. Finally, we solve the systems (4.11) and (4.12) for all $0 \leq k \leq N$ by our preconditioned MINRES iteration. The number of MINRES iterations and the corresponding computational times can be found in Table 7.16 and Table 7.17, respectively. We observe from these tables that our solver also remains robust with respect to both the regularization parameter λ and the mesh size h in case of large jumps in the values of the coefficient functions ν and σ . In addition to this, the comparison of the number of iterations and the corresponding computational times for Examples 6 and 7 show that the efficiency of our solver is not affected by large jumps in the coefficients ν and σ . This is a very important issue for many practical applications where we have usually large jumps in the values of material coefficients.

Table 7.16: Number of MINRES iterations for different values of λ on grids of different mesh size (**Example 6** / Example 7).

grid / λ	10^{-8}	10^{-4}	1	10^4	10^8
64×64	22 / 14	18 / 18	12 / 16	10 / 14	10 / 14
128×128	22 / 15	20 / 20	12 / 18	10 / 18	10 / 18
256×256	22 / 16	20 / 24	14 / 18	12 / 16	12 / 16
512×512	23 / 16	22 / 24	14 / 17	12 / 12	12 / 12

Table 7.17: The CPU times in seconds for different values of λ on grids of different mesh size (**Example 6** / Example 7).

grid / λ	10^{-8}	10^{-4}	1	10^4	10^8
64×64	0.2 / 0.1	0.2 / 0.2	0.1 / 0.2	0.1 / 0.1	0.1 / 0.1
128×128	1.1 / 0.7	1.0 / 0.9	0.6 / 0.8	0.5 / 0.8	0.5 / 0.8
256×256	4.8 / 3.3	4.4 / 4.7	3.1 / 3.7	2.6 / 3.3	2.7 / 3.3
512×512	21.5 / 14.0	20.6 / 20.9	13.2 / 15.1	11.3 / 10.7	11.3 / 10.7

Chapter 8

Conclusions and outlook

In this chapter, we want to summarize our results and give an outlook on some future work.

Conclusions

We have provided a complete numerical analysis of linear parabolic boundary value and optimal control problems in a time-periodic setting and their discretization by means of the multiharmonic finite element method (MhFEM). Moreover, we have developed new algebraic multilevel preconditioners for solving the discrete problems by the preconditioned minimal residual (MINRES) method.

The mathematical and numerical analysis includes an existence and uniqueness proof of the weak solution to a special variational setting of the parabolic time-periodic boundary value problem and the corresponding optimal control problem in Chapters 3 and 4, respectively. More precisely, we have introduced the space $H^{1,\frac{1}{2}}$ which has provided a suitable framework for deducing existence and uniqueness of the problems by proving inf-sup and sup-sup conditions such that the theorem of Babuška and Aziz can be applied.

We have intensely studied the MhFEM for solving parabolic time-periodic problems, where all – given and unknown – functions are approximated by truncated Fourier series and the Fourier coefficients by the finite element method. The MhFEM is a very powerful tool for solving linear time-periodic problems since it reduces a large time-dependent problem to a sequence of smaller time-independent ones that can completely be solved in parallel. More precisely, the large systems of linear algebraic equations fortunately decouple into smaller linear systems each of them defining the cosine and sine Fourier coefficients with respect to a single frequency. The resulting systems have a saddle point structure and can be solved by the preconditioned MINRES method. We have constructed block-diagonal preconditioners leading to robust and fast convergence rates for the MINRES method following the work by Zulehner in [187].

The diagonal blocks of the MINRES preconditioners are sums of stiffness and mass matrices. Since the finite element discretization of reaction-diffusion type problems with heterogeneous reaction and diffusion coefficients leads to such sums of stiffness and mass matrices, we have presented the construction of efficient preconditioners for these problems by the linear algebraic multilevel iteration (AMLI) method, which has been introduced in [14, 15]. Moreover, one of the main achievements of this work is not only the construction of efficient multilevel preconditioners but the derivation of a rigorous proof for the robustness and optimality of this AMLI method for heterogeneous reaction-diffusion type problems in two space dimensions, see Chapter 5. More precisely, we have verified the optimality conditions for linear AMLI preconditioners constructed in the framework of hierarchical splittings of lowest-order conforming finite element spaces for reaction-diffusion type problems. A new estimate of the constant γ in the strengthened Cauchy-Bunyakowski-Schwarz inequality has been presented for the mass matrix in case of a general m -refinement. Moreover, an additive preconditioner for the pivot blocks arising in the recursive two-by-two block factorization has been analyzed for the

case $m = 3$. The derived uniform condition number estimates together with the verification of the optimality conditions lead to robust and optimal linear AMLI methods for linear systems with sums of stiffness and mass matrices.

The numerical analysis has also involved full a priori and a posteriori discretization error estimates for the parabolic time-periodic boundary value and optimal control problems, including the analysis for the space $H^{1, \frac{1}{2}}$. We have analyzed the error coming from the approximation via truncated Fourier series as well as from the finite element approximation of the Fourier coefficients. The a posteriori error analysis, which can be found in Chapter 6, is based on the method presented in Repin [152], but we had to incorporate proper changes regarding the space $H^{1, \frac{1}{2}}$ and the special features of the MhFEM.

Although the main focus of this work was to develop the theoretical machinery for a rigorous numerical analysis of parabolic time-periodic problems, we also implemented the algorithms developed and confirmed our theoretical results by proper numerical experiments in Chapter 7.

Altogether this thesis presents the MhFEM as a very efficient approach for the discretization of linear parabolic time-periodic simulation and optimal control problems as well as provides optimal and robust solvers for this type of problems.

Outlook

- In this work, we have assumed that the Fourier coefficients of the given data, e.g., the source term f or the desired state y_d , can be computed exactly, but, in general, the Fourier coefficients of the data have to be computed numerically. Hence, the efficient numerical computation of the Fourier coefficients together with an error analysis is a matter of future work, cf. also [93], where it is briefly discussed that three-term recurrences can be used to evaluate the integrals appearing in the Fourier coefficients by a forward-backward recursion.
- In case of given time-analytic data, the error estimates can be definitely improved, which is also observed in our numerical experiments, cf. Remark 4.11 and also Remark 3.20. For given time-analytic data, we may expect exponential convergence with respect to the truncation parameter of the Fourier series.
- The incorporation of initial conditions instead of time-periodic ones is definitely an important topic for further investigations. Especially, in case of optimal control problems, this leads to many interesting questions since the initial condition for the forward problem turns into a final condition for the adjoint problem. Ideas for solving this problem are, for instance, to incorporate the initial condition somehow into the Fourier series approximation of our unknown functions or to use – instead of Fourier series – other spectral methods like Legendre or integrated Legendre polynomials for the approximation in time.
- In case of optimal control problems, the efficient treatment of inequality constraints for the control and the state using the MhFEM is a challenging topic. However, inequality constraints imposed on the Fourier coefficients of the state or the control can easily be included into the MhFE approach, although one loses the robustness with respect to the cost or regularization parameter when solving the optimality system by the preconditioned MINRES method, see [88]. The inclusion of inequality constraints imposed on the state or the control itself is much harder to handle. One technique to handle this problem is to include them as penalty term in the cost functional or to use barrier methods. However, this makes the optimality system nonlinear. Nonlinearities of this kind, but also nonlinearities arising from nonlinear partial differential equations as in the case of coefficients which depend on the solution, e.g., $\nu = \nu(\mathbf{x}, t, |\nabla y|)$ or $\nu = \nu(\mathbf{x}, t, y)$, lead to coupled nonlinear optimality systems. The Newton linearization results in linear systems where all modes are coupled. However, the block-diagonal preconditioners constructed for the linear case could be very useful for the efficient solution of the linear systems

arising at each step of the Newton method, see [27] for the solution of time-periodic eddy current problems. The primal-dual active set method may also help to handle nonlinearities arising from prescribing inequality constraints. This method is equivalent to a semi-smooth Newton method, see [80].

- Although the parallel implementation of computing the Fourier coefficients for different modes is straightforward, it should be put into practice, since the practical parallelization as well as a proper optimization of our algorithms should improve the efficiency of our solver.
- In the a posteriori error analysis of Chapter 6, we have obtained majorants for our parabolic time-periodic boundary value problem as well as for the optimality system and cost functional of the corresponding optimal control problem. Of course, the computation of co-called minorants for the parabolic time-periodic problems is a matter of future work.
- Another important topic is the practical implementation of the final bounds from the majorants, which we have obtained in the a posteriori error analysis of Chapter 6. Here, we can use techniques that are known from the elliptic case, see [153, 123]. The construction of an adaptive multiharmonic finite element method (AMhFEM) based on a posteriori error estimates is another challenging area for further research and should yield adaptivity in space and time.
- We have proved robustness and optimality of our linear AMLI method for heterogeneous reaction-diffusion type problems in two space dimensions. A rigorous proof for three space dimensions is another important issue for further investigations.
- In order to prove robustness and optimality of our AMLI method, we have assumed that the reaction and diffusion coefficients are constant on the coarsest mesh partitioning. Hence, the next step would be to prove robustness and optimality considering parameters which have jumps in their values on the finest mesh leading to highly heterogeneous problems which arise in multiscale analysis, see, e.g., [66].

Bibliography

- [1] D. Abbeloos, M. Diehl, M. Hinze, and S. Vandewalle. Nested multigrid methods for time-periodic, parabolic optimal control problems. *Comput. Vis. Sci.*, 14(1):27–38, 2011.
- [2] G. Acosta and R. G. Durán. An optimal Poincaré inequality in L^1 for convex domains. *Proceedings of the AMS*, 132(1):195–202, 2004.
- [3] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*. Pure and Applied Mathematics 140, Elsevier/Academic Press, second edition, 2008.
- [4] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics, Wiley-Interscience, New York, 2000.
- [5] S. G. Andrade and A. Borzi. Multigrid second-order accurate solution of parabolic control-constrained problems. *Comput. Optim. Appl.*, 51(2):835–866, 2012.
- [6] R. Andreev. Stability of sparse space-time finite element discretizations of linear parabolic evolution equations. *IMA J. Numer. Anal.*, 33(1):242–260, 2013.
- [7] R. Andreev. Stability of space-time Petrov-Galerkin discretizations for parabolic evolution equations, PhD thesis, ETH Zürich, 2012.
- [8] O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, 1994.
- [9] O. Axelsson and V. A. Barker. *Finite Element Solution of Boundary Value Problems: Theory and Computations*. Academic Press, Orlando, 1984.
- [10] O. Axelsson and R. Blaheta. Two simple derivations of universal bounds for the C.B.S. inequality constant. *Appl. Math.*, 49(1):57–72, 2004.
- [11] O. Axelsson and V. Eijkhout. The nested recursive two-level factorization method for nine-point difference matrices. *SIAM J. Sci. Stat. Comput.*, 12(6):1373–1400, 1991.
- [12] O. Axelsson and I. Gustafsson. Preconditioning and two-level multigrid methods of arbitrary degree of approximations. *Math. Comp.*, 40(161):219–242, 1983.
- [13] O. Axelsson and A. Padiy. On the additive version of the algebraic multilevel iteration method for anisotropic elliptic problems. *SIAM J. Sci. Comput.*, 20(5):1807–1830, 1999.
- [14] O. Axelsson and P. S. Vassilevski. Algebraic multilevel preconditioning methods I. *Numer. Math.*, 56(2–3):157–177, 1989.
- [15] O. Axelsson and P. S. Vassilevski. Algebraic multilevel preconditioning methods II. *SIAM J. Numer. Anal.*, 27(6):1569–1590, 1990.
- [16] O. Axelsson and P. S. Vassilevski. Variable-step multilevel preconditioning methods, I: Self-adjoint and positive definite elliptic problems. *Numer. Linear Algebra Appl.*, 1(1):75–101, 1994.

- [17] O. Axelsson and P. S. Vassilevski. *The AMLI Method: an Algebraic Multilevel Iteration Method for Positive Definite Sparse Matrices*. University of Nijmegen, Departement of Math., 1999.
- [18] I. Babuška. Error-bounds for finite element method. *Numer. Math.*, 16(4):322–333, 1971.
- [19] I. Babuška. The finite element method with Lagrangian multipliers. *Numer. Math.*, 20:179–192, 1973.
- [20] I. Babuška and A. K. Aziz. Survey lectures on the mathematical foundation of the finite element method. *The Mathematical Foundation of the Finite Element Method with Applications to Partial Differential Equations (A. K. Aziz ed.)*, Academic Press, New York, pages 1–359, 1972.
- [21] I. Babuška and T. Janik. The h-p version of the finite element method for parabolic equations. part I. the p-version in time. *Numer. Methods Partial Differential Equations*, 5(4):363–399, 1989.
- [22] I. Babuška and T. Janik. The h-p version of the finite element method for parabolic equations. II. the h-p version in time. *Numer. Methods Partial Differential Equations*, 6(4):343–369, 1990.
- [23] I. Babuška and W. C. Rheinboldt. A posteriori error estimates for the finite element method. *Int. J. Numer. Meth. Eng.*, 12(10):1597–1615, 1978.
- [24] F. Bachinger. Multigrid solvers for 3D multiharmonic nonlinear magnetic field computations, Master thesis, JKU Linz, 2003.
- [25] F. Bachinger, M. Kaltenbacher, and S. Reitzinger. An efficient solution strategy for the HBFEM method. *Proceedings of the IGTE '02 Symposium Graz, Austria*, pages 385–389, 2002.
- [26] F. Bachinger, U. Langer, and J. Schöberl. Numerical analysis of nonlinear multiharmonic eddy current problems. *Numer. Math.*, 100:593–616, 2005.
- [27] F. Bachinger, U. Langer, and J. Schöberl. Efficient solvers for nonlinear time-periodic eddy current problems. *Comput. Vis. Sci.*, 9(4):197–207, 2006.
- [28] R. E. Bank and A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comp.*, 44(170):283–301, 1985.
- [29] M. Bebendorf. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen*, 22(4):751–756, 2003.
- [30] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numer.*, 14:1–137, 2005.
- [31] J. Bergh and J. Löfström. *Interpolation spaces: An introduction*. Springer, Berlin-Heidelberg-New York, 1976.
- [32] C. Bernardi, M. Dauge, and Y. Maday. *Spectral methods for axisymmetric domains*. Series in Applied Mathematics 3, Gauthier-Villars & North-Holland, 1999.
- [33] S. Beuchler. AMLI preconditioner for the p-version of the FEM. *Numer. Linear Algebra Appl.*, 10(8):721–732, 2003.
- [34] J. Bewersdorff. *Galois Theory for Beginners: A Historical Perspective*. Student Mathematical Library 35, AMS, Providence, RI, 2006.
- [35] D. A. Bini and B. Iannazzo. A note on computing matrix geometric means. *Adv. Comput. Math.*, 35(2–4):175–192, 2011.
- [36] A. Borzi. Multigrid methods for parabolic distributed optimal control problems. *J. Comput. Appl. Math.*, 157(2):365–382, 2003.

- [37] A. Borzi and V. Schulz. Multigrid methods for PDE optimization. *SIAM Review*, 51(2):361–395, 2009.
- [38] A. Borzi and V. Schulz. *Computational Optimization of Systems Governed by Partial Differential Equations*. SIAM, Philadelphia, 2012.
- [39] P. Boyanova and S. Margenov. Robust multilevel preconditioning methods. *Efficient Preconditioned Solution Methods for Elliptic Partial Differential Equations (O. Axelsson, J. Karatson eds.)*, pages 3–22, 2011.
- [40] D. Braess. *Finite elements: Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, second edition, 2005.
- [41] D. Braess. *Finite Elemente: Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie*. Springer, Berlin-Heidelberg, 5. Auflage, 2013.
- [42] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Texts in Applied Mathematics 15, Springer, New York, second edition, 2002.
- [43] C. W. Brown. QEPCAD B: a program for computing with semi-algebraic sets using cads. *ACM SIGSAM Bulletin*, 37(4):97–108, 2003.
- [44] C. Canuto, Y. Maday, and A. Quarteroni. Analysis of the combined finite element and Fourier interpolation. *Numer. Math.*, 39(2):205–220, 1982.
- [45] P. L. Chebyshev. Sur les polynômes représentant le mieux les valeurs des fonctions fractionnaires élémentaires pour les valeurs de la variable contenues entre deu limites données. *Oeuvres de P. L. Chebyshev*, 2:669–678, 1899–1907.
- [46] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Studies in Mathematics and its Applications 4, North-Holland, Amsterdam, 1978. Republished by SIAM in 2002.
- [47] G. E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decomposition. *Automata theory and formal languages (Second GI Conf., Kaiserslautern, 1975), Lecture Notes in Comput. Sci.*, 33:134–183, 1975.
- [48] G. E. Collins and H. Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *J. Symb. Comput.*, 12(3):299–328, 1991.
- [49] D. M. Copeland and U. Langer. Domain decomposition solvers for nonlinear multiharmonic finite element equations. *J. Numer. Math.*, 18(3):157–175, 2010.
- [50] M. Costabel. Boundary integral operators for the heat equation. *Integral Equations Operator Theory*, 13(4):498–552, 1990.
- [51] R. Courant. Variational methods for the solution of problems of equilibrium and vibrations. *Bulletin of AMS*, 49(1):1–23, 1943.
- [52] H. de Gersen, H. V. Sande, and K. Hameyer. Strong coupled multiharmonic finite element simulation package. *COMPEL*, 20(2):535–546, 2001.
- [53] J. Douglas and T. Dupont. Galerkin methods for parabolic equations. *SIAM J. Numer. Anal.*, 7(4):575–626, 1970.
- [54] V. Eijkhout and P. Vassilevski. The role of the strengthened Cauchy-Buniakowskii-Schwarz inequality in multilevel methods. *SIAM Review*, 33(3):405–419, 1991.
- [55] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems I: A linear model problem. *SIAM J. Numer. Anal.*, 28(1):43–77, 1991.

- [56] K. Eriksson, C. Johnson, and V. Thomée. Time discretization of parabolic problems by the discontinuous Galerkin method. *Modél. Math. Anal. Numér.*, 19(4):611–643, 1985.
- [57] L. C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics 19, AMS, Providence, RI, 1998.
- [58] J. B. J. Fourier. Mémoire sur la propagation de la chaleur dans les corps solides. *Nouveau Bulletin des sciences par la Société philomatique de Paris*, 1(6):112–116, 1808. Présenté le 21 décembre 1807 à l’Institut national.
- [59] A. Gaevskaya, R. H. W. Hoppe, and S. Repin. A posteriori estimates for cost functionals of optimal control problems. *Numerical Mathematics and Advanced Applications, Proceedings of the ENUMATH 2005*, pages 308–316, 2006.
- [60] A. Gaevskaya, R. H. W. Hoppe, and S. Repin. Functional approach to a posteriori error estimation for elliptic optimal control problems with distributed control. *Journal of Mathematical Sciences*, 144(6):4535–4547, 2007.
- [61] A. V. Gaevskaya and S. I. Repin. A posteriori error estimates for approximate solutions of linear parabolic problems. *Differential Equations*, 41(7):970–983, 2005.
- [62] B. G. Galerkin. Beams and plates. Series in some questions of elastic equilibrium of beams and plates. *Vestnik Ingenerov*, 19:897–908, 1915. In Russian.
- [63] M. J. Gander and S. Vandewalle. Analysis of the parareal time-parallel time-integration method. *SIAM J. Sci. Comput.*, 29(2):556–578, 2007.
- [64] M. J. Gander and G. Wanner. From Euler, Ritz, and Galerkin to modern computing. *SIAM Review*, 54(4):627–666, 2012.
- [65] W. Gong, M. Hinze, and Z. J. Zhou. Space-time finite element approximation of parabolic optimal control problems. *J. Numer. Math.*, 20(2):111–145, 2012.
- [66] I. G. Graham, T. Y. Hou, O. Lakkis, and R. Scheichl. *Numerical Analysis of Multiscale Problems*. Lecture Notes in Computational Science and Engineering 83, Springer, Berlin-Heidelberg, 2012.
- [67] A. Greenbaum. *Iterative Methods for Solving Linear Systems*. Frontiers in Applied Mathematics 17, SIAM, Philadelphia, 1997.
- [68] C. Grossmann, H. G. Roos, and M. Stynes. *Numerical Treatment of Partial Differential Equations*. Springer, Berlin-Heidelberg, 2007.
- [69] M. Gunzburger and C. Trenchea. Analysis and discretization of an optimal control problem for the time-periodic MHD equations. *J. Math. Anal. Appl.*, 308(2):440–466, 2005.
- [70] M. Gunzburger and C. Trenchea. Optimal control of the time-periodic MHD equations. *Nonlinear Anal.*, 63(5–7):e1687–e1699, 2005.
- [71] M. D. Gunzburger and A. Kunoth. Space-time adaptive wavelet methods for optimal control problems constrained by parabolic evolution equations. *SIAM J. Control Optim.*, 49(3):1150–1170, 2011.
- [72] B. Gustafsson, H. O. Kreiss, and J. Olinger. *Time dependent problems and difference methods*. Pure and Applied Mathematics 67, Wiley-Interscience, New York, 1995.
- [73] J. Gyselinck, P. Dular, C. Geuzaine, and W. Legros. Harmonic-balance finite-element modeling of electromagnetic devices: a novel approach. *IEEE Trans. Magn.*, 38(2):521–524, 2002.

- [74] W. Hackbusch. Fast numerical solution of time-periodic parabolic problems by a multigrid method. *SIAM J. Sci. Stat. Comput.*, 2(2):198–206, 1981.
- [75] W. Hackbusch. Parabolic multigrid methods. *Computing Methods in Applied Sciences and Engineering VI (R. Glowinski, J. L. Lions eds.)*, North-Holland, Amsterdam, pages 189–197, 1984.
- [76] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer Series in Computational Mathematics 4, Springer, Berlin-Heidelberg, 1985.
- [77] B. Heinrich. The Fourier-finite-element method for Poisson’s equation in axisymmetric domains with edges. *SIAM J. Numer. Anal.*, 33(5):1885–1911, 1996.
- [78] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards*, 49(6):409–436, 1952.
- [79] E. Hille and R. S. Phillips. *Functional Analysis and Semi-Groups*. Colloquium Publications 31, AMS, Providence, RI, 1957.
- [80] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888, 2002.
- [81] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*. Mathematical Modelling: Theory and Applications 23, Springer, Berlin, 2009.
- [82] R. Hiptmair. Operator preconditioning. *Comput. Math. Appl.*, 52(5):699–706, 2006.
- [83] G. Horton and S. Vandewalle. A space-time multigrid method for parabolic partial differential equations. *SIAM J. Sci. Comput.*, 16(4):848–864, 1995.
- [84] M. Jung and U. Langer. *Methode der finiten Elemente für Ingenieure: Eine Einführung in die numerischen Grundlagen und Computersimulation*. Springer, Wiesbaden, second edition, 2013.
- [85] Y. Katznelson. *An Introduction to Harmonic Analysis*. Cambridge Mathematical Library, Cambridge University Press, Cambridge, third edition, 2004.
- [86] M. Kauers. How to use cylindrical algebraic decomposition. *Sém. Lothar. Combin.*, 65(B65a):1–16, 2011.
- [87] M. Kollmann. Efficient iterative solvers for saddle point systems arising in PDE-constrained optimization problems with inequality constraints, PhD thesis, JKU Linz, 2013.
- [88] M. Kollmann and M. Kolmbauer. A preconditioned MinRes solver for time-periodic parabolic optimal control problems. *Numer. Linear Algebra Appl.*, 20(5):761–784, 2013.
- [89] M. Kollmann, M. Kolmbauer, U. Langer, M. Wolfmayr, and W. Zulehner. A finite element solver for a multiharmonic parabolic optimal control problem. *Comput. Math. Appl.*, 65(3):469–486, 2013.
- [90] M. Kolmbauer. A robust FEM-BEM MinRes solver for distributed multiharmonic eddy current optimal control problems in unbounded domains. *ETNA*, 39:231–252, 2012.
- [91] M. Kolmbauer. Efficient solvers for multiharmonic eddy current optimal control problems with various constraints and their analysis. *IMA J. Numer. Anal.*, 33(3):1063–1094, 2013.
- [92] M. Kolmbauer. A multiharmonic solver for nonlinear parabolic problems, Master thesis, JKU Linz, 2009.
- [93] M. Kolmbauer. The multiharmonic finite element and boundary element method for simulation and control of eddy current problems, PhD thesis, JKU Linz, 2012.

- [94] M. Kolmbauer and U. Langer. A frequency-robust solver for the time-harmonic eddy current problem. *Scientific Computing in Electrical Engineering SCEE 2010 (B. Michielsen, J. R. Poirier eds.), Mathematics in Industry 2012, Springer*, pages 97–105, 2012.
- [95] M. Kolmbauer and U. Langer. A robust preconditioned MinRes solver for distributed time-periodic eddy current optimal control problems. *SIAM J. Sci. Comput.*, 34(6):B785–B809, 2012.
- [96] M. Kolmbauer and U. Langer. Efficient solvers for some classes of time-periodic eddy current optimal control problems. *Numerical Solution of Partial Differential Equations: Theory, Algorithms, and Their Applications*, 45:203–216, 2013.
- [97] V. G. Korneev. Iterative methods of solving systems of equations of the finite element method. *USSR Computational Mathematics and Mathematical Physics*, 17(5):109–129, 1977.
- [98] J. Kraus. An algebraic preconditioning method for M-matrices: Linear versus non-linear multilevel iteration. *Numer. Linear Algebra Appl.*, 9(8):599–618, 2002.
- [99] J. Kraus. Additive Schur complement approximation and application to multilevel preconditioning. *SIAM J. Sci. Comput.*, 34(6):A2872–A2895, 2012.
- [100] J. Kraus, M. Lybery, and S. Margenov. Robust multilevel methods for quadratic finite element anisotropic elliptic problems. *Numer. Linear Algebra Appl.*, 2013. (to appear).
- [101] J. Kraus and S. Margenov. *Robust Algebraic Multilevel Methods and Algorithms*. Radon Series on Computational and Applied Mathematics 5, Walter de Gruyter, Berlin, 2009.
- [102] J. Kraus, P. Vassilevski, and L. Zikatanov. Polynomial of best uniform approximation to $1/x$ and smoothing in two-level methods. *Comput. Methods Appl. Math.*, 12(4):448–468, 2012.
- [103] J. Kraus and M. Wolfmayr. On the robustness and optimality of algebraic multilevel methods for reaction-diffusion type problems. *Comput. Vis. Sci.*, 2014. (to appear).
- [104] J. K. Kraus and S. K. Tomar. Algebraic multilevel iteration method for lowest order Raviart–Thomas space and applications. *Int. J. Numer. Meth. Eng.*, 86(10):1175–1196, 2011.
- [105] W. Krendl, V. Simoncini, and W. Zulehner. Stability estimates and structural spectral properties of saddle point problems. *Numer. Math.*, 124(1):183–213, 2013.
- [106] V. N. Kutrunov. Polynomial of best uniform approximation in an iterative method of solving systems of linear algebraic equations. *Siberian Mathematical Journal*, 33(1):48–53, 1992.
- [107] Y. A. Kuznetsov. Efficient iterative solvers for elliptic finite element problems on nonmatching grids. *Russ. J. Numer. Anal. Math. Model.*, 10(3):187–211, 1995.
- [108] O. A. Ladyzhenskaya. *The Boundary Value Problems of Mathematical Physics*. Nauka, Moscow, 1973. In Russian. Translated in *Appl. Math. Sci.* 49, Springer, 1985.
- [109] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural’ceva. *Linear and Quasilinear Equations of Parabolic Type*. AMS, Providence, RI, 1968.
- [110] J. Lang. *Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems. Theory, Algorithm, and Applications*. Lecture Notes in Computational Sciences and Engineering 16, Springer, Berlin, 2000.
- [111] U. Langer and M. Wolfmayr. Multiharmonic finite element solvers for time-periodic parabolic optimal control problems. *Proc. Appl. Math. Mech. (PAMM)*, 12(1):687–688, 2012.

- [112] U. Langer and M. Wolfram. Multiharmonic finite element analysis of a time-periodic parabolic optimal control problem. *J. Numer. Math.*, 21(4):265–300, 2013.
- [113] J. D. Lawson and Y. Lim. The geometric mean, matrices, metrics, and more. *Amer. Math. Monthly*, 108(9):797–812, 2001.
- [114] P. D. Lax and A. N. Milgram. Parabolic equations. *Contributions to the theory of partial differential equations, Ann. of Math. 33, Princeton University Press*, pages 167–190, 1954.
- [115] R. D. Lazarov and S. D. Margenov. CBS constants for multilevel splitting of graph-Laplacian and application to preconditioning of discontinuous Galerkin systems. *Journal of Complexity*, 23(4):498–515, 2007.
- [116] G. M. Lieberman. Time-periodic solutions of linear parabolic differential equations. *Comm. Partial Differential Equations*, 24(3–4):631–663, 1999.
- [117] J. L. Lions. *Optimal Control on Systems Governed by Partial Differential Equations*. Springer, Berlin-Heidelberg-New York, 1971.
- [118] J. L. Lions, Y. Maday, and G. Turinici. A “parareal” in time discretization of PDE’s. *Comptes Rendus de l’Académie des Sciences Series I Mathematics*, 332(7):661–668, 2001.
- [119] J. L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications*. Travaux et Recherches Mathématiques 1, Dunod, Paris, 1968. In French.
- [120] J. L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications*. Translated from French by P. Kenneth, *Die Grundlehren der mathematischen Wissenschaften*, Band 181, Springer, New York-Heidelberg, 1972.
- [121] Y. Maday, M. K. Riahi, and J. Salomon. Parareal in time intermediate targets methods for optimal control problems. *Control and Optimization with PDE Constraints*, 164:79–92, 2013.
- [122] J. F. Maitre and F. Musy. The contraction number of a class of two-level methods: An exact evaluation for some finite element subspaces and model problems. *Multigrid Methods (W. Hackbusch, U. Trottenberg eds.), Lect. Notes Math. 960*, pages 535–544, 1982.
- [123] O. Mali, P. Neittaanmäki, and S. Repin. *Accuracy Verification Methods. Theory and Algorithms*. Computational Methods in Applied Sciences 32, Springer, Netherlands, 2014.
- [124] K. A. Mardal and R. Winther. Preconditioning discretizations of systems of partial differential equations. *Numer. Linear Algebra Appl.*, 18(1):1–40, 2011.
- [125] D. Meidner and B. Vexler. Adaptive space-time finite element methods for parabolic optimization problems. *SIAM J. Control Optim.*, 46(1):116–142, 2007.
- [126] C. Mense and R. Nabben. On algebraic multilevel methods for non-symmetric systems - convergence results. *ETNA*, 30:323–345, 2008.
- [127] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comput.*, 21(6):1969–1972, 2000.
- [128] A. I. Nazarov and S. I. Repin. Exact constants in Poincaré type inequalities for functions with zero mean boundary traces. *arXiv:1211.2224*, 2012.
- [129] I. Neitzel and F. Tröltzsch. On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints. *ESAIM: COCV*, 15(2):426–453, 2009.
- [130] M. Neumüller. *Space-Time Methods: Fast Solvers and Applications*. Monographic Series TU Graz. Computation in Engineering and Science 20, PhD thesis, 2013.

- [131] M. Neumüller and O. Steinbach. Refinement of flexible space-time finite element meshes and discontinuous Galerkin methods. *Comput. Visual. Sci.*, 14(5):189–205, 2011.
- [132] J. Nečas. *Les Méthodes Directes en Théorie des Équations Elliptiques*. Masson, Paris, 1967.
- [133] Y. Notay and P. S. Vassilevski. Recursive Krylov-based multigrid cycles. *Numer. Linear Algebra Appl.*, 15(5):473–487, 2008.
- [134] J. T. Oden and J. N. Reddy. *An Introduction to the Mathematical Theory of Finite Elements*. Wiley-Interscience, New York, 1976. Republished by Courier Dover Publications in 2012.
- [135] H. Okochi. On the existence of periodic solutions to nonlinear abstract parabolic equations. *J. Math. Soc. Japan*, 40(3):541–553, 1988.
- [136] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12(4):617–629, 1975.
- [137] C. V. Pao. Numerical methods for time-periodic solutions of nonlinear parabolic boundary value problems. *SIAM J. Numer. Anal.*, 39(2):647–667, 2001.
- [138] G. Paoli, O. Bíró, and G. Buchgraber. Complex representation in nonlinear time harmonic eddy current problems. *IEEE Trans. Magn.*, 34(5):2625–2628, 1998.
- [139] L. E. Payne and H. F. Weinberger. An optimal Poincaré inequality for convex domains. *Arch. Rat. Mech. Anal.*, 5(1):286–292, 1960.
- [140] J. W. Pearson, M. Stoll, and A. J. Wathen. Regularization-robust preconditioners for time-dependent PDE-constrained optimization problems. *SIAM. J. Matrix Anal. & Appl.*, 33(4):1126–1152, 2012.
- [141] J. W. Pearson and A. J. Wathen. A new approximation of the Schur complement in preconditioners for PDE-constrained optimization. *Numer. Linear Algebra Appl.*, 19(5):816–829, 2012.
- [142] C. Pechstein. *Finite and Boundary Element Tearing and Interconnecting Solvers for Multi-scale Problems*. Lecture Notes in Computational Science and Engineering 90, Springer, Berlin-Heidelberg-New York, 2013.
- [143] C. Pechstein and R. Scheichl. Weighted Poincaré inequalities. *IMA J. Numer. Anal.*, 33(2):652–686, 2013.
- [144] D. Petz. Means of positive matrices: Geometry and a conjecture. *Annales Mathematicae et Informaticae*, 32:129–139, 2005.
- [145] V. Pillwein and S. Takacs. A local Fourier convergence analysis of a multigrid method using symbolic computation. *J. Symb. Comput.*, 63:1–20, 2014.
- [146] A. Potschka, M. S. Mommer, J. P. Schlöder, and H. G. Bock. Newton-Picard-based preconditioning for linear-quadratic optimization problems with time-periodic parabolic PDE constraints. *SIAM J. Sci. Comput.*, 34(2):A1214–A1239, 2012.
- [147] W. Pusz and S. L. Woronowicz. Functional calculus for sesquilinear forms and the purification map. *Rep. Math. Phys.*, 8(2):159–170, 1975.
- [148] P. A. Raviart and J. M. Thomas. A mixed finite element method for 2-nd order elliptic problems. *Mathematical Aspects of Finite Element Methods, Lect. Notes Math. 606*, pages 292–315, 1977.
- [149] J. N. Reddy. Existence and uniqueness of solutions to a stationary finite element model of the biharmonic equation. *Comput. Math. Appl.*, 3(2):135–147, 1977.

- [150] S. Repin. A posteriori error estimates for approximate solutions to variational problems with strongly convex functionals. *Problems of Mathematical Analysis*, 17:199–226, 1997. In Russian. Translated in *J. Math. Sci.* 97(4):4311–4328, 1999.
- [151] S. Repin. A posteriori error estimation for variational problems with uniformly convex functionals. *Math. Comput.*, 69(230):481–500, 2000.
- [152] S. Repin. Estimates of deviation from exact solutions of initial-boundary value problems for the heat equation. *Rend. Mat. Acc. Lincei*, 13(2):121–133, 2002.
- [153] S. Repin. *A Posteriori Estimates for Partial Differential Equations*. Radon Series on Computational and Applied Mathematics 4, Walter de Gruyter, Berlin, 2008.
- [154] S. Repin, S. Sauter, and A. Smolianski. Two-sided a posteriori error estimates for mixed formulations of elliptic problems. *SIAM J. Numer. Anal.*, 45(3):928–945, 2007.
- [155] W. Ritz. Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. *J. Reine Angew. Math.*, 135:1–61, 1908.
- [156] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, second edition, 2003.
- [157] J. Schöberl and W. Zulehner. Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems. *SIAM J. Matrix Anal. Appl.*, 29(3):752–773, 2007.
- [158] C. Schwab and R. Stevenson. Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comp.*, 78(267):1293–1318, 2009.
- [159] A. Seidl and T. Sturm. A generic projection operator for partial cylindrical algebraic decomposition. *Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation (R. Sendra ed.)*, pages 240–247, 2003.
- [160] D. Sheen, I. H. Sloan, and V. Thomée. A parallel method for time discretization of parabolic equations based on laplace transformation and quadrature. *IMA J. Numer. Anal.*, 23(2):269–299, 2003.
- [161] O. Steinbach. *Numerische Näherungsverfahren für elliptische Randwertprobleme: Finite Elemente und Randelemente*. B. G. Teubner, Stuttgart-Leipzig-Wiesbaden, 2003.
- [162] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems: Finite and Boundary Elements*. Springer, New York, 2008.
- [163] M. Steuerwalt. The existence, computation, and number of solutions of periodic parabolic problems. *SIAM J. Numer. Anal.*, 16(3):402–420, 1979.
- [164] A. Strzeboński. Solving systems of strict polynomial inequalities. *J. Symb. Comput.*, 29(3):471–480, 2000.
- [165] S. Takacs and V. Pillwein. Smoothing analysis of an all-at-once multigrid approach for optimal control problems using symbolic computation. *Numerical and Symbolic Scientific Computing (U. Langer, P. Paule eds.)*, 1:175–191, 2012.
- [166] L. Tartar. *An Introduction to Sobolev Spaces and Interpolation Spaces*. Springer, Berlin-Heidelberg, 2007.
- [167] V. Thomée. *Galerkin finite element methods for parabolic problems*. Springer Series in Computational Mathematics 25, Springer, Berlin-Heidelberg, second edition, 2006.

- [168] A. Toselli and O. Widlund. *Domain Decomposition Methods - Algorithms and Theory*. Springer Series in Computational Mathematics 34, Springer, Berlin, 2005.
- [169] F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications*. Graduate Studies in Mathematics 112, AMS, Providence, RI, 2010.
- [170] F. Tröltzsch and I. Yousept. PDE-constrained optimization of time-dependent 3D electromagnetic induction heating by alternating voltages. *ESAIM: M2AN*, 46(04):709–729, 2012.
- [171] S. Vandewalle and R. Piessens. Efficient parallel algorithms for solving initial-boundary value and time-periodic parabolic partial differential equations. *SIAM J. Sci. Stat. Comput.*, 13(6):1330–1346, 1992.
- [172] S. Vandewalle and R. Piessens. On dynamic iteration methods for solving time-periodic differential equations. *SIAM J. Numer. Anal.*, 30(1):286–303, 1993.
- [173] P. S. Vassilevski. *Multilevel Block Factorization Preconditioners: Matrix-based Analysis and Algorithms for Solving Finite Element Equations*. Springer, New York, 2008.
- [174] A. Veiser and R. Verfürth. Poincaré constants for finite element stars. *IMA J. Numer. Anal.*, 32(1):30–47, 2012.
- [175] O. Vejjoda. *Partial differential equations: time-periodic solutions*. Martinus Nijhoff Publishers, The Hague, 1982. (in collaboration with L. Herrmann et al.).
- [176] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Stuttgart, 1996.
- [177] J. Wloka. *Partial Differential Equations*. Cambridge University Press, Cambridge, 1987.
- [178] D. Womble. A time-stepping algorithm for parallel computers. *SIAM J. Sci. Statist. Comput.*, 11(5):824–837, 1990.
- [179] S. Yamada and K. Bessho. Harmonic field calculation by the combination of finite element analysis and harmonic balance method. *IEEE Trans. Magn.*, 24(6):2588–2590, 1988.
- [180] I. Yousept. Finite element analysis of an optimal control problem in the coefficients of time-harmonic eddy current equations. *J. Optim. Theory Appl.*, 154(3):879–903, 2012.
- [181] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/A: Linear Monotone Operators*. Springer, New York, 1990.
- [182] E. Zeidler. *Nonlinear Functional Analysis and its Applications II/B: Nonlinear Monotone Operators*. Springer, New York, 1990.
- [183] O. C. Zienkiewicz and J. Z. Zhu. A simple error estimator and adaptive procedure for practical engineering analysis. *Int. J. Numer. Meth. Eng.*, 24(2):337–357, 1987.
- [184] M. Zlámal. On the finite element method. *Numer. Math.*, 12(5):394–409, 1968.
- [185] M. Zlámal. Curved elements in the finite element method. I. *SIAM J. Numer. Anal.*, 10(1):229–240, 1973.
- [186] W. Zulehner. *Numerische Mathematik: Eine Einführung anhand von Differentialgleichungsproblemen Band 1: Stationäre Probleme*. Birkhäuser Verlag. Mathematik Kompakt, 2008.
- [187] W. Zulehner. Nonstandard norms and robust estimates for saddle point problems. *SIAM J. Matrix Anal. Appl.*, 32(2):536–560, 2011.
- [188] W. Zulehner. *Numerische Mathematik: Eine Einführung anhand von Differentialgleichungsproblemen Band 2: Instationäre Probleme*. Birkhäuser Verlag. Mathematik Kompakt, 2011.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig 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. Die vorliegende Dissertation ist mit dem elektronisch übermittelten Textdokument identisch.

Linz, Februar 2014

Monika Wolfmayr

Curriculum Vitae

Name: Monika Wolfmayr (née Kowalska)

Nationality: Polish

Date of Birth: April 19, 1986

Place of Birth: Warsaw, Poland

Education:

1992–1996	Volksschule Preinsbacher Straße (elementary school), Amstetten, Austria
1996–2004	Bundesgymnasium (secondary comprehensive school), Amstetten
2004–2008	Studies in Technical Mathematics (Bachelor), JKU Linz
2008–2010	Studies in Industrial Mathematics (Master), JKU Linz
since 2010	Studies in Engineering Sciences (PhD), Technical Mathematics, JKU Linz PhD student at the Institute of Computational Mathematics and then at the Doctoral Program “Computational Mathematics”, project DK04

Selected Activities:

August 2008	22 nd ECMI Modelling Week, Eindhoven, Netherlands
February 2011	Projektwoche Angewandte Mathematik 2011, Austria, project leader
Mar–Apr 2011	3 rd Spring School on Analytical and Numerical Aspects of Evolution Equations, Essen, Germany
October 2011	ACSC Conference: Future Tools and Topics, Linz, Austria
Oct–Dec 2011	Special Semester on Multiscale Simulation and Analysis in Energy and the Environment, Linz, Austria
February 2012	Projektwoche Angewandte Mathematik 2012, Austria, project leader
May 2012	Austrian Numerical Analysis Day 2012, Vienna, Austria, talk
August 2012	European Multigrid Conference 2012, Schwetzingen Castle, Germany, talk

- August 2012 Symposium on Scientific Computing in Honour of Ulrich Langer,
Schwetzingen Castle, Germany
- August 2012 ESSIM 2012 Modelling Week, Dresden, Germany, instructor
- November 2012 Oberwolfach Seminar: Subspace Correction Methods,
Mathematisches Forschungsinstitut Oberwolfach, Germany
- March 2013 Workshop on Numerical Methods for Optimal Control and
Inverse Problems, Munich, Germany, talk
- April 2013 Austrian Numerical Analysis Day 2013, Graz, Austria, talk
- Apr–May 2013 Scientific stay at TU Berlin, Germany, talk
- May 2013 IFIP TC7.2 Workshop: Electromagnetics – Modelling, Simulation,
Control and Industrial Applications in Berlin at the WIAS
(Weierstraß – Institut für Angewandte Analysis und Stochastik),
Berlin, Germany, talk
- May 2013 Talk at the WIAS (Weierstraß – Institut für Angewandte Analysis
und Stochastik) invited by Prof. Hömberg, Berlin, Germany
- July 2013 6th Workshop on Analysis and Advanced Numerical Methods for
Partial Differential Equations, Strobl, Austria, talk
- September 2013 Domain Decomposition Methods for Optimization with PDE Constraints,
Monte Verità, Ascona, Switzerland, minisymposium talk

Publications:

- M. Kollmann, M. Kolmbauer, U. Langer, M. Wolfmayr, and W. Zulehner.
A finite element solver for a multiharmonic parabolic optimal control problem.
Comput. Math. Appl., 65(3):469-486, 2013, see [89].
- U. Langer and M. Wolfmayr.
Multiharmonic finite element solvers for time-periodic parabolic optimal control problems.
Proc. Appl. Math. Mech. (PAMM), 12(1):687-688, 2012, see [111].
- U. Langer and M. Wolfmayr.
Multiharmonic finite element analysis of a time-periodic parabolic optimal control problem.
J. Numer. Math., 21(4):265-300, 2013, see [112].
- J. Kraus and M. Wolfmayr.
On the robustness and optimality of algebraic multilevel methods for reaction-diffusion
type problems. *Computing and Visualization in Science*, 2014, (to appear), see [103].