Robust Multigrid Methods for Parameter Dependent Problems

DISSERTATION

zur Erlangung des akademischen Grades
"Doktor der Technischen Wissenschaften"

eingereicht von
Dipl.-Ing. Joachim Schöberl

angefertigt am Institut für Analysis und Numerik,
Arbeitsgruppe Numerische Mathematik und Optimierung
der Technisch-Naturwissenschaftlichen Fakultät
der Johannes Kepler Universität Linz

o. Univ.-Prof. Dr. U. Langer	1. Gutachter
o. Univ.-Prof. Dr. D. Braess	2. Gutachter

Linz, im Juni 1999

For Elsa

Abstract

This thesis is concerned with the construction and analysis of robust multigrid preconditioners for parameter dependent problems in primal variables. The developed framework is applied to the specific examples of nearly incompressibility, the Timoshenko beam, and the Reissner Mindlin plate. The suggested multigrid components are simple to implement.

The work is based on two theories. On one side, the theory of mixed finite element methods is essential for the analysis of parameter dependent problems. The material is presented in a rather self-contained way. The point of departure is the primal form, where the stability conditions are formulated. The obtained discretization schemes are known as reduced integration methods. We point out the essence of a Fortin interpolation operator.

The other base contains the additive Schwarz theory and multigrid theory. These techniques provide a framework for the construction and analysis of efficient preconditioners. We will collect the concepts of one-level and two-level subspace splitting, approximation property and smoothing property by means of elliptic problems without parameters. The formulation is such that it carries over to parameter dependent problems.

The central point of the thesis are the combination of both theories. We will start with one-level preconditioners for parameter dependent problems. We see that block smoothers capturing base functions of the kernel are robust with respect to the small parameter. The analysis uses function splitting with a partition of unity, and interpolation with the Fortin operator. We proceed with two level methods. A trivial prolongation by embedding is not uniformly bounded with respect to the parameter dependent energy norm. The idea for the construction of robust prolongation operators is that coarse grid kernel functions must be lifted to fine grid kernel functions. This can be obtained by adjusting proper degrees of freedom locally. In the considered applications, the implementation consists of solving local sub-problems of the assembled stiffness matrix. The prolongation operator is an approximative right inverse to the Fortin operator.

The components developed for the two-level method can be used in a multigrid algorithm. The analysis requires two norms, for which the approximation property and the smoothing property are verified. One is the parameter dependent energy norm, the other one combines three terms, namely improved convergence in a weaker norm of the primal variable, stability in primal energy, as well as stability in average for the dual variable. The approximation property can be proven under abstract assumptions. The smoothing property has to be checked individually for the considered examples. Even under strongest realistic regularity assumptions, we have to apply interpolation norms. We point out some relations to the smoother by Braess and Sarazin.

Finally, numerical experiments are presented. They agree with the analysis for the W-cycle and variable V-cycle. Additionally, they show optimal and robust convergence of V-cycle methods.

Zusammenfassung

Diese Arbeit beschäftigt sich mit der Kontruktion und Analyse von robusten Vorkonditionierern für parameterabhängige Probleme in primalen Variablen. Die entwickelte Methodik wird auf fast inkompressible Materialien, den Timoshenko Balken und die Reissner Mindlin Platte angewandt. Die vorgeschlagenen Multigrid Komponenten sind einfach zu realisieren.

Die Arbeit beruht auf zwei Theorien. Das eine Standbein ist die gemischte Finite Elemente Methode, die für die Analyse von parameterabhängigen Problemen fundamental ist. Der Stoff wird weitgehend selbstenthaltened dargestellt. Der Ausgangspunkt ist das primale Problem, für das die Stabilitätsbedingungen formuliert werden. Die erhaltene Diskretisierung ist als Methode mit reduzierter Integration bekannt. Wir unterstreichen die Bedeutung des Interpolationsoperatos nach Fortin.

Das zweite Standbein umfaßt Additiv Schwarz Methoden und die Multigrid Theorie. Diese Techniken bilden die Methodik zur Konstruktion und Analyse von effizienten Vorkonditionereren. Wir stellen die Konzepte von Ein- und Zweigittermethoden, und von Approximations- und Glättungseigenschaft an Hand von elliptischen Problemen ohne Parameter dar. Die gewählte Formulierung läßt sich auf parameterabhängige Probleme übertragen.

Der Kern der vorliegenden Arbeit ist die Verbindung beider Theorien. Wir beginnen mit Eingitter Vorkonditionierern für parameterabhängige Probleme. Wir sehen, daß Blockglätter, die Basisfunktionen für den Kern umfassen, robust bezüglich des Parameters arbeiten. Die Analyse verwendet Zerlegung der Eins und den Interpolationsoperator nach Fortin. Als nächsten Schritt nehmen wir ein Grobgittersystem hinzu. Eine trivial Prolongation ist nicht gleichmäßig stetig bezüglich des kleinen Parameters. Die Idee einer robusten Prolongation besteht darin, daß Grobgitter-Kernfunktionen auf Feingitter-Kernfunktionen abgebildet werden. Das kann durch lokale Anpassung geeigneter Freiheitsgrade erreicht werden. Bei den betrachteten Beispielen kann die Realisierung durch Lösen von Teilproblemen der assemblierten Steifigkeitsmatrix geschehen. Der Prolongationsoperator ist eine approximative Rechtsinverse des Fortin-Operators.

Die Komponenten des Zweigitterverfahrens können in einem Mehrgitteralgorithmus eingesetzt werden. Die Analyse benötigt zwei Normen, für welche die Approximationseigenschaft und die Glättungseigenschaft gezeigt werden. Eine ist die parameterabhängige Energienorm, die andere ist eine Kombination von drei Termen, nämlich einer verbesserten Approximation der primalen Variable in einer schwächeren Norm, die Stabilität in der primalen Energie, und Stabilität der gemittelten dualen Variablen. Die Approximationseigenschaft kann unter abstrakten Voraussetzungen gezeigt werden. Die Glättungseigenschaft wird für die spezifischen Beispiele überprüft. Auch unter stärksten realistischen Regularitätsannahmen müssen Interpolationsnormen verwendet werden. Es werden Verbindungen zu dem Glätter von Braess und Sarazin hergestellt.

Numerische Beispiele bestätigen die theoretischen Ergebnisse für die robusten und optimalen Ratenabschätzungen beim W-Zykus und variablen V-Zyklus. Weiters zeigen sie auch entsprechende Konvergenzraten für den V-Zyklus.

Acknowledgements

First of all, I want to express my gratitude to my supervisor Professor U. Langer. He introduced me to the theory, and guided me to that interesting topics of research. Discussions with him have always been fruitful to bring structure into vague thoughts. I want to thank Prof. D. Braess (Bochum), Doz. W. Zulehner (Linz), Dr. Ch. Wieners (Stuttgart) and Dr. A. Tamme (Linz) for a lot of discussions on that topics over the last three years. I want to thank Prof. S. V. Nepomnyaschikh (Novosibirsk) for telling me much about splitting methods. I want to thank Prof. R. Stenberg (Innsbruck) for showing me how to go on with plate elements.

I want to acknowledge the Austrian Science Foundation FWF for financial support under the project grants P 10643-TEC and P 11215-TEC.

I want to thank the crews from the department of Numerical Analysis and Optimization as well as from the SFB 013 for the scientific and social environment.

I want to thank Elsa, my parents, Sabine, my grand-parents, Else and Josef for their support in general and during the time for preparing this work. I apologize for the moments when my presence has been only physical.

Finally, I want to thank everybody (they are too many to list) who pushed me to finish this work.

Contents

Chapter 1

Introduction

State of the Art

Many engineering problems are mathematically modeled by elliptic boundary value problems. The finite element (fe) method is certainly the most frequently applied tool for the numerical approximation of these models. The computer power available today enables very accurate computer simulations by using a million and more elements. The large problem size requires numerical algorithms of optimal, this means linear, time complexity. The challenge is to design efficient iterative methods for the solution of the high dimensional linear systems. Multigrid methods meet these requirements.

There is a lot of literature available about the fe method and multigrid methods (see the textbooks [BS94], [Bra97] for two more recent presentations of the theory of finite elements and an introduction to multigrid methods, as well as many references therein). The monograph [Hac85] gives a wide overview about the theory of multigrid methods. The textbook [Bra93] presents the more recent results related to multi-level methods. See also [Gri94], [Osw94].

The accuracy of fe methods depends on the regularity properties of the underlying boundary value problem, and of the approximation quality of the fe space. For a lot of problems the standard (i.e. conforming) approach works well and gives optimal error estimates of the form

$$rel.err. \leq c\, N^{-\alpha}$$

Here, $rel.err.$ is the relative error of the computed finite element approximation. The constant c has a moderate size, N is the number of unknowns. The power α depends on the spatial dimension, the polynomial degree of the fe space, and the adaption of the mesh.

For many realistic problems, the constant c is really large, and many unknowns have to be spent, until the error drops below 50%. This misbehavior is called *locking effect*. This topic is discussed in many textbooks of finite element analysis in mechanical engineering, see e.g. [Bat96] for a lot of practical examples.

The first step is to detect the source of locking. Often, a small parameter ε of the equation can be specified. Then the family of problems depending on $\varepsilon \in (0, 1)$ may behave like

$$rel.err \leq c\,\varepsilon^{-\beta}\,N^{-\alpha}, \tag{1.1}$$

where c is a constant of moderate size, and β is a positive constant.

One typical examples of locking in solid mechanics is nearly incompressibility locking (also known as volume locking or Poisson locking). Here, the source of the problem stems from the material parameters. If the Poisson ration ν, which is allowed to take values in $(0, 1/2)$, is close to $1/2$, the material behaves (nearly) incompressible. The error estimate (1.1) applies with $\varepsilon = 1/2 - \nu$. Nearly incompressible subregions occur e.g. in non-linear problems from elasto-plasticity [KL84]. We will consider this type of locking in this thesis.

An other source of locking may be the shape of the geometry. Problems from elasticity on thin domains are usually described by models from structural mechanics, see [DS96], [Ber96], [CTV95], [Cia90]. A hierarchy of plate models is derived in [Sch89], [BL91], [AAFM99], and in [Sch96] with a posteriori control.

The lowest order models for flat geometries are the Bernoulli beam model and the Kirchhoff plate model. These are fourth order equations, which can be well approximated by proper (conforming or non-conforming) fe methods. The next models in the asymptotical expansion are the Timoshenko beam model, and the Reissner Mindlin plate model, which are systems of second order equations. These systems depend explicitely on the thickness t of the structure. A standard fe approximation of these models leads to so called *shear locking* (see Section 2.1.3 for numerical examples). We will investigate these two models in this thesis. The more involved shell models cover the behavior of curved thin structures. Already the lowest order model of Kroiter may suffer from *membrane locking* [Pit92]. In [CB98] a systematic classification of shell structures is given. [Bis99] applies the EAS concept for shell problems. We will not consider shell models in the present work.

The considered examples and (many more) belong to the class of *parameter dependent problems* in the sense of Arnold [Arn81]. They lead to variational problems: Find $u \in V$ such that

$$A^{\varepsilon}(u, v) = f(v) \qquad \forall\, v \in V,$$

where the bilinear form $A^{(\varepsilon)}(.,.)$ has the special structure

$$A^{\varepsilon}(u, v) = a(u, v) + \frac{1}{\varepsilon} c(\Lambda u, \Lambda v).$$

The form $a(.,.)$ is symmetric, non-negative and continuous on V, the form $c(.,.)$ is symmetric, elliptic and continuous on a second Hilbert space Q. The continuous operator $\Lambda : V \to Q$ has the non-trivial kernel V_0. The bilinear form $A^1(.,.)$ is assumed to be elliptic on V. Then the form $A^{\varepsilon}(.,.)$ is uniformly elliptic on V, but the continuity constant depends on ε. In general, this class of problems leads to a priori estimates of the form (1.1) for conforming finite element methods (see [Arn81], [CP94], [BS92b], [BS92a], and [Bra97]).

The standard approach for a robust discretization is to define the dual variable

$$p := \frac{1}{\varepsilon}\Lambda u \qquad \in Q,$$

and pass to the equivalent mixed problem: Find $u \in V$, $p \in Q$ such that

$$\begin{array}{rcrcll} a(u,v) & + & c(\Lambda v, p) & = & f(v) & \quad \forall\, v \in V, \\ c(\Lambda u, p) & - & \varepsilon\, c(p,q) & = & 0 & \quad \forall\, q \in Q. \end{array}$$

The concept of mixed variational problems go back to [Bre74]. A deep source of collected information is [BF91], see also [RT91], [GR86]. For proper norms on V and Q, the mixed problem is continuous and stable, and thus provides an isomorphism between $V \times Q$ and $V^* \times Q^*$. The mixed form is discretized by the choice of fe subspaces V_h and Q_h. The discretization has to fulfill two conditions. On one hand side, the discrete LBB condition condition must hold, on the other hand side, a discrete ellipticity condition must be guaranteed. Then the discrete problem is stable as well, and the theory of Babuška and Aziz [BA72] provides robust convergence. In [AB93] a general *stabilization trick* was applied to make the form $a(.,.)$ elliptic itself, and thus the construction of stable fe spaces becomes much simpler. In [Bra96], a unified analysis of mixed problems with penalty is developed.

Often, the dual fe variable p_h can be eliminated locally, and one returns to a positive definite fe problem: Find $u_h \in V_h$ such that

$$a(u_h, v_h) + \frac{1}{\varepsilon} c(\Lambda_h u_h, \Lambda_h v_h) = f(v_h) \qquad \forall\, v_h \in V_h.$$

This approach of was systematic analyzed by [MH78], see also [Sin78], and [JP82] for delicate questions of unstable elements. It is known as *reduced integration technique*. The operator Λ_h is a softening of the original one, Λ. An alternative approach is the enhanced assumed strain (EAS) concept by Simo and Rifai (see [SR90], [AR93]). Often, the method is equivalent to a mixed method ([Bra98]).

From the viewpoint of computation, the problem in primal variables is preferable against the mixed form, because only one field of variables is used and thus the system matrix is considerable smaller and more sparse. It is positive definite, such that the conjugate gradient method can be applied. Last but not least, the version in primal variables is simpler to implement.

A widely used discretization scheme for plate and shell models is the MITC family of elements introduced in [BBF89] and further analyzed by [BFS91], [PB92]. In [AF89] a robust non-conforming linear element is analyzed. A simple but efficient alternative based on stabilized finite element techniques was introduced by Arnold and Brezzi [AB93]. Several versions of stabilization were investigated in [BL97] and [Lov96]. Chapelle and Stenberg [CS98] consider a special choice of stabilization. They establish improved regularity estimates and apply them for duality techniques. This work has been the bases to understand the multigrid method for Reissner Mindlin plates. Pitkäranta and Suri [PS96] give an overview of robust methods. Lyly and Stenberg [LS99] give an overview of robust stabilized methods.

The problem of bad finite element approximation was solved by the equivalent mixed form. But still, the small parameter ε spoils the condition number of the arising system matrix $\underline{A}^{(\varepsilon)}$. No robust preconditioner have been available for this class of problem. That was one of the reasons to use the mixed form, and apply a solver for the saddle point system of equations (see [AHU58], [Axe79], [FG83], [BP88], [LQ86], [LQ87], [Que89], [BWY90], [Pei91], [RW92], [VL96], [Kla98], [BPV97], [Zul98a], and more). The main goal of the present work is to develop and analyze robust and efficient preconditioners for the ill conditioned matrix $\underline{A}^{(\varepsilon)}$.

We will collect well-known results for the construction and analysis of preconditioners based on subspace correction methods. Because the special structure of reduced integration schemes lead to different bilinear forms on each level, we are only interested in methods applicable to non-nested forms.

There are two underlying principles. One is the splitting of finite element function into subspace functions by means of local operators. This principle does not require problem regularity, but is restricted to the two level case.

The other principle is Hackbusch's multigrid theory based on the *smoothing property* and the *approximation property*. It uses problem regularity and proves methods of optimal complexity for W-cycle [Hac82], [BD81], [BD85], and variable V-cycle [BPX91] multigrid methods. Optimal convergence of V-cycle schemes is proved in [BH83] if full regularity is available, and in [BP93] for less then full regularity. Up to the knowledge of the author, optimal V-cycle analysis is available only for nested bilinear forms.

The framework to design preconditioners based on splitting of finite element functions is the additive Schwarz method. The method of [Sch69] was adapted by [Nep86], [Lio88], [DW90], [Zha91] to domain decomposition preconditioning. [SBG96] is a recent textbook.

We will mainly focus on additive Schwarz methods, but the multiplicative counterpart is understood as well from the work of [BPWX91]. The framework of multilevel Schwarz methods combines both concepts ([Yse86], [Xu92], [Zha92]).

Multigrid methods applied directly to the system in mixed variables have been provided by [Ver84], [Wit89a], [Hua90], [BB90], several papers by S. Brenner unified in [Bre96], [BS97]. The papers mainly differ by the type of smoother. In [Wie99], the smoother from Braess and Sarazin is adapted for the problem in primal variables

More specific, we borrow techniques for the analysis of multigrid preconditioners for fourth order problems [PB87], [Bre89], and, closely related, Stokes' equations with divergence free basis functions [Bre90], [Tur94]. The design of our smoothers shares ideas of [EW92], [VW92], [AFW97b] and [Hip99] for the construction of multigrid methods for $H(\text{div})$ and $H(\text{rot})$. These methods can be used to precondition the blocks arising from mixed finite element systems, e.g. for Reissner Mindlin problems [AFW97a].

An alternative approach to multigrid schemes is the AMLI method by Axelsson [Axe96]. It is based only on the strengthened Cauchy Schwarz inequality. In [KM99] an AMLI algorithm for nearly incompressible materials is analyzed.

In the present work we do not consider the parameter dependent problem of anisotropy. We refer to [Wit89b], [Ste93] [Neu98] for anisotropies on aligned meshes. The method of [Pad97] is proved to provide optimal and robust preconditioners for non-aligned anisotropies as well.

Overview

The emphasis of this thesis is to present techniques for the construction and analysis of multigrid methods for parameter dependent problems in primal variables. The obtained preconditioners are the first ones of optimal time complexity which are robust with respect to the mesh size and the small parameter. For the analysis of the multigrid preconditioners some results from mixed finite element theory had to be adapted.

The construction of the presented preconditioners was presented in [Sch99c]. It contains also the one-level and two-level analysis for nearly incompressible materials and for the Timoshenko beam. [Sch99b] contains the multigrid analysis for the case of nearly incompressible materials. [Sch99a] contains the abstract multigrid analysis, and the application to the Reissner Mindlin plate.

The work is organized as follows. Chapter 2 contains an introduction into selected topics of finite element methods. The theory of parameter dependent problems is presented as far as needed in a self-contained manner. The examples of Timoshenko beam, nearly incompressible materials and Reissner Mindlin are formulated. Chapter 3 gives an introduction to subspace correction methods. The principle of one-level, two-level and multigrid methods are explained for problems without parameters. Chapter 4 contains the new results for preconditioning of parameter dependent problems. Chapter 5 presents numerical results.

In particular, this work contributes to the following approaches and results.

- In Section 2.3 the formulation, stabilization and discretization of parameter dependent problems is presented in a unified manner. A relaxed version of Fortin's criterion, a sufficient condition for stable mixed finite element discretization is formulated (Theorem 2.12).

- Section 2.4.3 collects the results of [AB93] and [CS98] for the stabilization and discretization of the Reissner Mindlin plate model. By means of the relaxed version of Fortin's criterion, the scheme with mesh dependent stabilization terms fits into the abstract framework. We obtain first stability of the infinite dimensional problem and then perform a stable discretization. This result may be useful for a posteriori error estimates. We modify the norm for the duality trick, such that problem regularity, the approximation inequality and the inverse inequality fit together.

- Section 4.1 explains the construction of local preconditioners, which are robust with respect to the parameter, but depends on the mesh size. The idea of block smoothers is borrowed from [VW92] and [AFW97b], but the environment is different.

- Section 4.2 extends the local preconditioner by a coarse grid system to obtain condition numbers robust in the parameter and mesh size. Since the forms are non-nested, grid transfer operators are necessary. To be robust, the operator has to map coarse grid kernel functions to fine grid kernel functions, essentially.

- Sections 4.3 formulates multigrid methods, which can be performed equivalently in primal variables or in mixed variables. This is the basis for the multigrid analysis. The norms to prove the approximation property and the smoothing property are defined. The norms differ from the norms usually used for multigrid methods in mixed variables. Our norms are related stronger to the primal problem.

- Section 4.4 proves the approximation property under abstract assumptions.

- The smoother of Braess and Sarazin [BS97] was adapted by Wieners [Wie99] to the formulation in primal variables. Section 4.5 compares the analysis by Wieners to the present work. A corresponding smoother for the Reissner Mindlin plate model is suggested.

- In Section 4.6 the smoothing property for the local smoothers suggested is proved for different applications. The technique uses additive Schwarz methods and operator interpolation in Hilbert spaces.

- Finally, by combining the approximation property and the smoothing property of the local smoothers, robust solvers of optimal arithmetic complexity are obtained.

Notation

The symbol $a \preceq b$ means that there exists a constant c independent of a and b, as well as ε, and the discretization parameter $\hbar$ defined below such that $a \leq c\, b$. We write $a \succeq b$ for $b \preceq a$, and $a \simeq b$ for $a \preceq b$ and $b \preceq a$.

When $(V, \|.\|_V, (.,.)_V)$ is a Hilbert space, and $A(.,.)$ is a continuous bilinear form on $V \times V$, we will associate to it the operator $A : V \to V$ defined by

$$(Au, v)_V = A(u, v) \qquad \forall\, u, v, \in V.$$

If $A(.,.)$ is symmetric and non-negative, it defines the energy semi-norm

$$\|u\|_A := A(u, u)^{1/2}.$$

Throughout this work, the following symbols keep their meaning:

ε	small parameter
V	Hilbert space of primal variable $(u, v, w \in V)$
Q	Hilbert space of dual variable $(p, q, r \in Q)$
X	$X = V \times Q$
$a(.,.)$	symmetric and non-negative bilinear form on V
$c(.,.)$	symmetric and elliptic bilinear form on Q
Λ	linear operator $V \to Q$
V_0	kernel of Λ
X_0	$\{(u,p) \in X : \Lambda u = \varepsilon\, p\}$
$A^\varepsilon(.,.)$	primal form $A^\varepsilon(u,v) = a(u,v) + \varepsilon^{-1} c(\Lambda u, \Lambda v)$
$B^\varepsilon(.,.)$	mixed form $B^\varepsilon((u,p),(v,q)) = a(u,v) + c(\Lambda v, p) + c(\Lambda u, q) - \varepsilon\, c(p,q)$
C	preconditioner (with varying indices)
D	simple preconditioner used for smoother

Chapter 2

Finite Element Theory

In this chapter a brief introduction to selected topics of the finite element method is given. We will start with a conforming discretization method for the Poisson equation and list the steps of the analysis. We proceed with the saddle point problem of Stokes'. These examples can be found in most finite element textbooks. Then we will see that a conforming method behaves bad for the beam model of Timoshenko (see e.g. [Bra97]).

We will introduce the toolkit of Sobolev spaces, the technique of Hilbert space interpolation, finite element interpolation and partition of unity techniques. Then we will proceed with parameter dependent problems. We will pass to mixed finite element methods. The extended toolkit involves saddle point problems, stabilization techniques, and the Fortin interpolation operator. We will end up with elliptic, but non-conforming finite element methods.

2.1 Finite Element Basics

2.1.1 A positive definite problem

First, we demonstrate by means of the most elementary partial differential equation the method of finite elements. We consider the Poisson equation with homogeneous Dirichlet boundary conditions

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega \end{aligned} \tag{2.1}$$

on the domain $\Omega \subset \mathbb{R}^2$. The finite element method requires the weak form of the equation. We seek for the solution u in the Hilbert function space $V = H_0^1(\Omega)$. The solution has to fulfill the variational problem

$$\int_\Omega \nabla u \, \nabla v \, \mathrm{dx} = \int_\Omega f v \, \mathrm{dx} \qquad \forall\; v \in V. \tag{2.2}$$

We define the symmetric bilinear form $A(.,.) : V \times V \to \mathbb{R}$ and the linear form $f(.) : V \to \mathbb{R}$ as

$$\begin{aligned} A(u,v) &:= \int_\Omega \nabla u \, \nabla v \, \mathrm{dx}, \\ f(v) &:= \int_\Omega f \, v \, dx. \end{aligned}$$

Then we can rewrite the problem abstract as: Find $u \in V$ such that

$$A(u,v) = f(v) \qquad \forall v \in V. \tag{2.3}$$

The analysis of symmetric and and elliptic problems is based on the fact that the energy norm $\|u\|_A := A(u,u)^{1/2}$ is an equivalent norm on V, i.e. there exists constants $c_1 > 0$ and $c_2 > 0$ such that

$$c_1 \|u\|_V \leq \|u\|_A \leq c_2 \|u\|_V \qquad \forall \, u \in V. \tag{2.4}$$

By means of the Lemma of Lax and Milgram there exists a unique solution $u \in V$ of the variational problem (2.3). The finite element method is an approach to approximate the (unknown) solution by something computable. We have to chose a finite dimensional space V_h where we will find the finite element solution u_h. In the standard (= conforming) approach the space V_h is a subspace of V, and the problem is the reduction of the variational problem (2.3) to: Find $u_h \in V_h$ such that

$$A(u_h, v_h) = f(v_h) \qquad \forall v_h \in V_h. \tag{2.5}$$

We have to analyze how good u_h approximates u. The first step is the question of stability. Variational problems with elliptic forms are always stable. The finite element solution is the best approximating function in the finite element space with respect to energy norm. Using norm equivalence (2.4), we obtain the estimate

$$\|u - u_h\|_V \leq \frac{c_2}{c_1} \inf_{v_h \in V_h} \|u - v_h\|_V \tag{2.6}$$

We estimated the *discretization error* of the variational problem by the *best approximation error* to the true solution u. The second step is the estimation of the approximation error. This question can be decided by properties of the space V_h and additional properties of the true solution concerning smoothness. Define a family of spaces $\{V_\hbar\}_{\hbar>0}$ characterized by the positive parameter $\hbar \to 0$. We assume that the solution is not only in V, but also in the dense subspace V^+ with the stronger norm $\|.\|_{V^+}$. We need a result from approximation theory of the form

$$\inf_{v_h \in V_h} \|u - v_h\|_V \leq \delta(\hbar) \, \|u\|_{V^+}. \tag{2.7}$$

The function $\delta(.) : \mathbb{R}^+ \to \mathbb{R}^+$ has to fulfill

$$\delta(\hbar) \to 0 \qquad \text{as} \qquad \hbar \to 0. \tag{2.8}$$

The third step in error analysis is to check that $u \in V^+$ is not only an assumption, but is indeed true. This part needs properties of the infinite dimensional problem and of the function space V^+. These are regularity theorems, which require properties of the domain (convex, cone-conditions, smooth boundary, etc.), the right hand side as well as the boundary conditions. We will cite some results of the form later.

Knowing that we have a good discretization method for the problem, we can prepare it for computing. Therefore, we have to chose a basis $(\varphi_i)_{i=1}^N$ for the finite element space V_h, with $N = \dim V_h$. By means of the basis we can rewrite problem (2.5) equivalently as the linear system

$$\underline{A}\,\underline{u} = \underline{f}. \tag{2.9}$$

The system matrix $\underline{A} \in \mathbb{R}^{N\times N}$, the load vector $\underline{f} \in \mathbb{R}^N$ and the solution vector $\underline{u} = (u_i)_{i=1}^N$ are defined by

$$\begin{aligned} \underline{A} &= (A(\varphi_i, \varphi_j))_{i,j=1}^N, \\ \underline{f} &= (f(\varphi_i))_{i=1}^N, \\ u_h &= \sum_{i=1}^N u_i\,\varphi_i. \end{aligned}$$

In principal, algorithms for the solution of linear systems are not new, but special properties of linear systems from finite element methods make it more exiting. First, the systems can be very large ($N = 10^6$ or more), and secondly, there are only a few non-zero elements per row in the matrix. This gives the chance for algorithms of optimal arithmetic complexity $O(N)$. Indeed, multigrid methods are such methods. We will give an overview of available modern fast methods in Chapter 3 and will analyze and apply these methods for the rest of the monograph.

2.1.2 A saddle point problem

The analysis of parameter dependent problems is strongly related to mixed variational problems. The best known example is Stokes' equation. The strong form of the equation reads as: Find the velocity u and the pressure p such that

$$\begin{array}{rcrcll} -\Delta u & - & \nabla p & = & f & \text{in } \Omega, \\ \operatorname{div} u & & & = & 0 & \text{in } \Omega, \\ & & u & = & 0 & \text{on } \partial\Omega. \end{array} \tag{2.10}$$

The weak form requires the Hilbert spaces $V = [H_0^1(\Omega)]^2$ and $Q = L_2/\mathbb{R}$. We search for $u \in V$ and $p \in Q$ such that

$$\begin{array}{lclcll} \int_\Omega \nabla u\, \nabla v \,\mathrm{dx} & + & \int_\Omega \operatorname{div} v\, p \,\mathrm{dx} & = & \int_\Omega f\, v \,\mathrm{dx} & \forall\, v \in V, \\ \int_\Omega \operatorname{div} u\, q \,\mathrm{dx} & & & = & 0 & \forall\, q \in Q. \end{array}$$

For an abstract notation we define the bilinear forms $a(.,.) : V \times V \to \mathbb{R}$ and $b(.,.) : V \times Q \to \mathbb{R}$, and the linear form $f(.) : V \to \mathbb{R}$ by

$$\begin{aligned} a(u,v) &:= \int_\Omega \nabla u \, \nabla v \, \mathrm{dx}, \\ b(u,q) &:= \int_\Omega \operatorname{div} u \, q \, \mathrm{dx}, \\ f(v) &:= \int_\Omega f \, v \, \mathrm{dx}. \end{aligned}$$

The mixed variational problem in abstract form reads as: Find $u \in V$ and $p \in Q$ such that

$$\begin{array}{llll} a(u,v) + b(v,p) & = & f(v) & \forall\, v \in V, \\ b(u,q) & = & 0 & \forall\, q \in Q. \end{array} \tag{2.11}$$

For a more compact notation we define the product space $X = V \times Q$ with the product norm $\|(u,p)\|_X^2 = \|u\|_V^2 + \|p\|_Q^2$ and the block bilinear form $B(.,.) : X \times X \to \mathbb{R}$ by

$$B((u,p),(v,q)) := a(u,v) + b(v,p) + b(u,q).$$

By adding both rows of the variational equation (2.11), the problem can be rewritten as: Find $(u,p) \in X$ such that

$$B((u,p),(v,q)) = f(v) \qquad \forall\, (v,q) \in X. \tag{2.12}$$

The condition (2.4) of equivalent norms is the essential property of positive definite problems. The corresponding extensions to more general problems are the conditions by Babuška and Aziz [BA72], namely stability

$$\sup_{(u,p)\in X} \frac{B((u,p),(v,q))}{\|(u,p)\|_X} \geq c_1 \|(v,q)\|_X \qquad \forall\, (v,q) \in X \tag{2.13}$$

and continuity

$$B((u,p),(v,q)) \leq c_2 \|(u,p)\|_X \|(v,q)\|_X \qquad \forall\, (u,p),(v,q) \in X. \tag{2.14}$$

Indeed, Babuška and Aziz can handle non-symmetric problems by taking care of the different kernels, but we will stay in the class of symmetric forms. If conditions (2.13) and (2.14) are fulfilled, then problem (2.12) has a unique solution. We define the kernel V_0 of $b(.,.)$ by

$$V_0 := \{v \in V : b(v,q) = 0 \;\forall\, q \in Q\}.$$

The theory of Brezzi [Bre74] gives sharp conditions onto the bilinear forms $a(.,.)$ and $b(.,.)$ to obtain stability (2.13). Namely, there must hold V_0-ellipticity of $a(.,.)$, i.e.

$$a(u,u) \geq c_a \|u\|_V^2 \qquad \forall\, u \in V_0, \tag{2.15}$$

and the famous LBB condition

$$\sup_{u \in V} \frac{b(u,q)}{\|u\|_V} \geq c_b \, \|q\|_Q \qquad \forall \, q \in Q. \tag{2.16}$$

The LBB condition for Stokes' problem is not trivial, see [DL76]. Clearly, continuity of $B(.,.)$ follows by continuity of $a(.,.)$ and $b(.,.)$.

The problem is called saddle point problem, because the solution $(u,p) \in V \times Q$ is a saddle point of the Lagrangian functional

$$L(v,q) := \frac{1}{2} a(v,v) + b(v,q) - f(v).$$

The component u is also a solution of the constrained minimization problem

$$\min_{v \in V_0} \frac{1}{2} a(v,v) - f(v).$$

The fe discretization needs finite element spaces for both variables, namely $V_h \subset V$ and $Q_h \subset Q$, and $X_h = V_h \times Q_h$. The fe problem is the restriction of (2.12): Find $(u_h, p_h) \in X_h$ such that

$$B((u_h,p_h),(v_h,q_h)) \qquad \forall \, (v_h,q_h) \in X_h.$$

But unlike to positive definite problems, additional conditions are required. Namely, discrete counterparts to the kernel ellipticity (2.15) and to the LBB condition (2.16) must be verified for the fe spaces V_h and Q_h. A lot of theory for the construction of mixed finite element spaces is available [BF91].

2.1.3 A parameter dependent problem

In Section 2.1.1 we formulated the basic principle of the finite element method for positive definite problems. Now, we consider examples for which the standard approach fails.

Let us consider the linear elasticity problem on the domain

$$\Omega = (0,1) \times (-t/2, t/2),$$

where $t \in (0,1)$ is a small parameter. The part of the boundary with Dirichlet boundary conditions is

$$\Gamma_D = \{0\} \times (-t/2, t/2),$$

see Figure 2.1.

Let V be the function space

$$V = \{ v \in [H^1(\Omega)]^2 : v = 0 \text{ on } \Gamma_D \}.$$

We consider the problem of linear elasticity: Find $u \in V$ such that

$$\int e(u) : D : e(v) \, \mathrm{dx} = \int f v \, \mathrm{dx}. \tag{2.17}$$

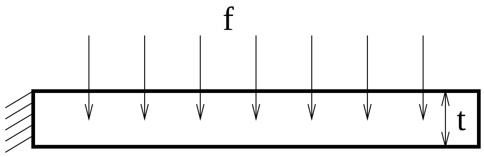
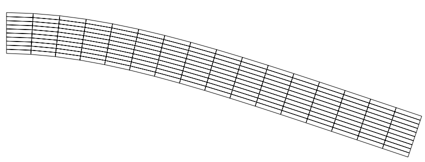


Figure 2.1: Thin beam with transverse loading

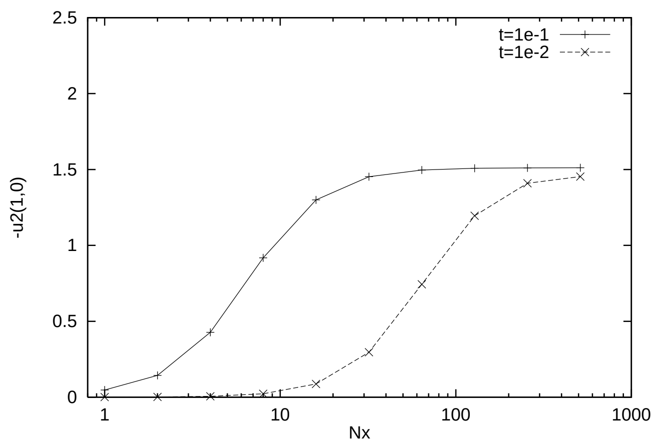
The operator $e(v)$ is the strain operator

$$e(v) = 0.5(\nabla v + (\nabla v)^T),$$

and D is the fourth order tensor of elastic coefficients. We consider the specific problem of uniform load $f = t^2(0, -1)$, and material parameters $E = 1$, $\nu = 0.2$. We perform finite element computations with bilinear rectangular elements. We use a fixed number of 10 elements in vertical direction, and vary the number of elements in horizontal directions. The solution on a mesh with 16×10 elements is plotted below, the thickness t is 0.1:



The vertical displacement $u_2(1, 0)$ at the right end of the beam is compared for varying thickness and number of elements in horizontal direction:



We observe that the results are totally useless until the mesh size in longitudinal direction is of the same dimension as the thickness.

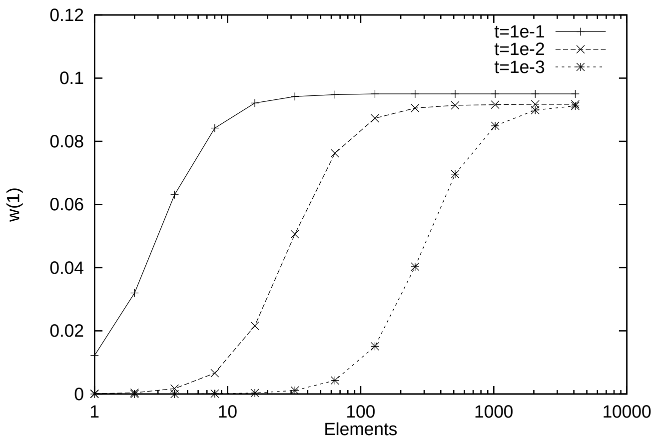
Usually, elasticity problems on thin domains are described by beam or plate models. The beam model of Timoshenko uses the infinite dimensional Ansatz space

$$\{u \in V : (u_1, u_2) = (x_2\beta(x_1), w(x_1)) \text{ with } \beta, w \in H^1((0,1))\}$$

described by the vertical deflection w and the rotation β. In principal, Galerkin projection of 2.17 to the space $\widetilde{V}$ gives the beam model of Timoshenko. By integration in vertical direction one obtains the variational problem: Find $(w, \beta) \in \widetilde{V} \subset [H^1((0,1))]^2$ such that

$$\int_0^1 \{\beta'\eta' + t^{-2}(w' - \beta)(v' - \eta)\}\,\mathrm{dx} = \int_0^1 f_2 v\,\mathrm{dx}. \qquad \forall\,(v,\eta) \in \widetilde{V}. \tag{2.18}$$

We neglected the coefficients in front of the differential terms. There are two reasons. First, they are not essential for the following analysis, and, secondly, the coefficients obtained by Galerkin projection are not asymptotically correct for $t \to 0$. This is tried to compensate by the so called shear correction factor found in any engineering work. We perform finite element simulations with piecewise linear elements for equation (2.18). The results are similar to the investigation above:



The locking effect was not removed by the use of the beam model, but the difficulties are more transparent. The energy of the model consists of two parts, the bending energy

$$\int (\beta')^2\,\mathrm{dx}$$

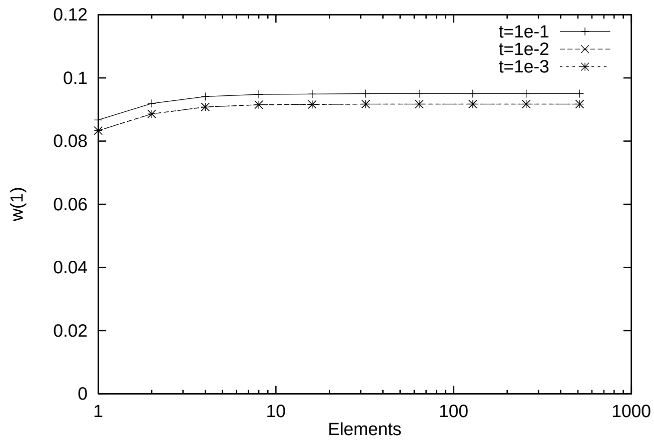
and the shear energy

$$t^{-2}\int (w' - \beta)^2\,\mathrm{dx}.$$

The problem is that for some functions the second term is by the factor t^{-2} larger, while for other functions the second term vanishes. There is the old technique of selective reduced integration (see [MH78], [Arn81]) to remove this problem. Instead of an exact formula of numerical integration, one of lower order is used. Often, this is equivalent to averaging some term of the bilinear form. We indicate element-wise averaging on the finite element mesh by $\bar{\cdot}^h$. We replace the form of (2.18) by

$$\int_0^1 \{\beta'\eta' + t^{-2}\overline{(w'-\beta)}^h\,\overline{(v'-\eta)}^h\}\,\mathrm{dx} = \int_0^1 f_2 v\,\mathrm{dx} \tag{2.19}$$

and repeat the simulations. The obtained results are:



By this trick, the locking effect is removed. We will discuss the relation to mixed methods in Section 2.3 in detail. We mention that also by increasing the polynomial degree of the finite element space the locking effect can be removed ([SBS95] discusses shear locking for plate models).

2.2 Approximation Results

2.2.1 Sobolev spaces

The natural spaces for variational problems are Sobolev spaces [Ada76]. In this monograph we are interested in linear problems and thus the subset of Hilbert spaces is enough.

The domain Ω is a bounded open set from $\mathbb{R}^d$ with $d = 2$ or $d = 3$. We want to avoid technical difficulties and assume that Ω is polygonal. The boundary $\partial\Omega$ is denoted by Γ, and $\Gamma_D \subset \Gamma$ is a closed subset.

The inner product $(.,.)_{0,\omega}$ and the associated norm $\|u\|_{0,\omega} := (u,u)_0^{1/2}$ are the of $L_2(\omega)$. We will use $\|.\|_0 := \|.\|_{0,\Omega}$. We will write the same symbol for vector and tensor valued functions. If γ is a manifold of lower dimension, then $\|.\|_{0,\gamma}$ is the corresponding L_2 norm on γ.

Let $\omega \subset \Omega$ be an open set with Lipschitz continuous boundary. We define the space C^∞ of infinitely differentiable functions on ω. The subspaces C_0^∞ and $C_{0,D}^\infty$ consist of functions which function values and all derivatives vanish on Γ and Γ_D, respectively. Indeed, we allow different sets Γ_D for different derivatives as well as for different components of vector valued functions.

For $k \in \mathbb{N}$ we define recursively the norms

$$\|u\|_{k,\omega} := \Big(\|u\|^2_{k-1,\omega} + \sum_{i=1,\dots,d} \Big\| \frac{\partial u}{\partial x_i} \Big\|^2_{k-1,\omega} \Big)^{1/2}.$$

We set $|u|_{0,\omega} := \|u\|_{0,\omega}$ and define for $k \in \mathbb{N}$ the semi-norms

$$|u|_{k,\omega} := \Big(\sum_{i=1,\dots,d} \Big| \frac{\partial u}{\partial x_i} \Big|^2_{k-1,\omega} \Big)^{1/2}.$$

The Sobolev spaces $H^k(\omega)$, $H^k_0(\omega)$, and $H^k_{0,D}(\omega)$ are the closures of $C^\infty(\omega)$, $C^\infty_0(\omega)$, and $C^\infty_{0,D}(\omega)$, respectively, with respect to the norms $\|u\|_{k,\omega}$. Again, for $\omega = \Omega$ we will skip the domain. For $k \in \mathbb{N}$ we define the the Sobolev spaces $H^{-k}(\omega)$ as dual spaces to $H^k_0(\omega)$.

Theorem 2.1. *Let $k_0, k_1 \in \mathbb{N}$ with $k_0 < k_1$. Then there holds $H^{k_1} \subset H^{k_0}$ with compact embedding.*

2.2.2 Hilbert space interpolation

Interpolation techniques are simple but powerful tools for finite element analysis. We will use the real method of interpolation going back to [LP64]. We refer also to [BS94]. In [Bra93] the special case of Hilbert spaces is discussed in more details.

Let $(V_0, (.,.)_0)$ and $(V_1, (.,.)_1)$ be two Hilbert spaces with compact embedding $V_1 \subset V_0$. The goal is to define a scale of spaces in between. For $t > 0$ and $u \in V_0$ define the K-functional as

$$K(t,u) := \inf_{u=u_0+u_1} (\|u_0\|^2_0 + t^2 \|u_1\|^2_1)^{1/2}. \tag{2.20}$$

The interpolation norm $\|.\|_\alpha$ for $0 < \alpha < 1$ is defined to be

$$\|u\|_\alpha := \|u\|_{[V_0,V_1]_\alpha} := \left(\int_0^\infty t^{-1-2\alpha} K(t,u)^2 \, \mathrm{dt} \right)^{1/2}. \tag{2.21}$$

The norm fulfills the parallelogram law, and thus it leads to a Hilbert space. The space $V_\alpha = [V_0, V_1]_\alpha$ is the closure of V_1 with respect to $\|.\|_\alpha$.

Theorem 2.2. *Let $V_1 \subset V_0$ and $W_1 \subset W_0$ be two pairs of compactly embedded Hilbert spaces. Let T be an linear operator mapping V_0 into W_0 as well as V_1 into W_1 with norm bounds*

$$\begin{aligned} \|Tv\|_{W_0} &\leq c_0 \|v\|_{V_0} \quad \forall\, v \in V_0, \\ \|Tv\|_{W_1} &\leq c_1 \|v\|_{V_1} \quad \forall\, v \in V_1. \end{aligned}$$

Then for $\alpha \in (0,1)$, T maps $[V_0, V_1]_\alpha$ into $[W_0, W_1]_\alpha$ with bound

$$\|Tv\|_{[W_0,W_1]_\alpha} \leq c_0^{1-\alpha} c_1^\alpha \|v\|_{[W_0,W_1]_\alpha} \quad \forall\, v \in [V_0, V_1]_\alpha. \tag{2.22}$$

An important application is interpolation of Sobolev spaces. Take for $k \in \mathbb{N}_0$ the spaces $V_0 = H^k(\omega)$ and $V_1 = H^{k+1}(\omega)$. The embedding is compact, so one can apply interpolation to define $H^{k+\alpha}(\omega) := [H^k(\omega), H^{k+1}(\omega)]_\alpha$ and $H_0^{k+\alpha}(\omega) := [H_0^k(\omega), H_0^{k+1}(\omega)]_\alpha$. Sobolev spaces of negative, non-integral parameters $-s$ are defined as dual spaces of $H_0^s(\omega)$.

Theorem 2.3. *Let Ω be a domain with Lipschitz-continuous boundary. Let $m_0, m_1, m \in \mathbb{N}_0$ with $m_0 < m < m_1$. Set α such that $m_0 = (1-\alpha)m_0 + \alpha m_1$. Then the norms of the spaces*

- *$H^m(\omega)$ and $[H^{m_0}(\omega), H^{m_1}(\omega)]_\alpha$*
- *$H_0^m(\omega)$ and $[H_0^{m_0}(\omega), H_0^{m_1}(\omega)]_\alpha$*

are equivalent, respectively.

The proof for H^m can be found in [BS94], while H_0^m is proved in [Bra95].

2.2.3 The Bramble-Hilbert lemma

There exists different formulations of the Bramble-Hilbert lemma in the literature, see e.g. [Bra97], [BS94]. We formulate first an abstract theorem ([GR86], Theorem 2.1):

Theorem 2.4. *Let $\widetilde{V}$ be a Hilbert space with norm $\|.\|_V$. Let $\widetilde{V} \subset W$ be a compact embedding into the Hilbert space $(W, \|.\|_W)$. Let $\|.\|_A$ be a semi-norm an $\widetilde{V}$ with kernel V_{00}. Assume that the following norms are equivalent:*

$$\|u\|_V \simeq \|u\|_W + \|u\|_A \qquad \forall\, u \in \widetilde{V}.$$

Then the following is true:

i. The kernel V_{00} is of finite dimension. The semi-norm is equivalent to the norm on the factor space, i.e.

$$\|u\|_A \simeq \inf_{u_0 \in V_{00}} \|u - u_0\|_V \qquad \forall\, u \in \widetilde{V}.$$

ii. Let $\|.\|_B$ be a continuous semi-norm on $\widetilde{V}$ such that there hold for all $u \in \widetilde{V}$

$$\|u\|_B + \|u\|_A = 0 \quad \Rightarrow \quad u = 0.$$

Then there holds the equivalence of norms

$$\|u\|_B + \|u\|_A \simeq \|u\|_V \qquad \forall\, u \in V.$$

iii. Let $V \subset \widetilde{V}$ be a closed subspace such that

$$V \cap V_{00} = \{0\}.$$

Then there holds the equivalence of norms

$$\|u\|_A \simeq \|u\|_V \qquad \forall\, u \in V.$$

The lemma of Bramble-Hilbert is a corollary to the above theorem:

Lemma 2.5 (Bramble-Hilbert). *Let $k \in \mathbb{N}$, and Q be a Hilbert space. Let $L : H^k \to Q$ be a linear and continuous operator. Assume that L vanishes on the space $\mathcal{P}_{k-1}$ of polynomials up to order $k-1$. Then L is bounded by the semi-norm, i.e.*

$$\|Lu\|_Q \preceq |u|_k.$$

2.2.4 Finite element spaces

For computing one has to approximate elements in Sobolev spaces by elements in finite dimensional function spaces. Finite element spaces are such spaces. The material in this section is referred to [Cia78]. One divides the domain Ω into a finite set of simple subdomains T, called elements. Together they form the triangulation $\mathcal{T} = \{T\}$. For simplification we assume that Ω is a polygonal domain. Usual elements are simplicials, namely segments in 1D, triangles in 2D and tetrahedra in 3D, and elements build by tensor products from simplicial elements, namely quadrilaterals, hexahedra and prisms.

Each element T is interpreted as the image of the mapping $x^T(\xi)$ from a reference element $T^{(R)}$. We define some terms for the triangulation:

- A triangulation is conforming, iff the intersection of two different elements is either empty, or contains one common corner point, one common edge or one common face.

- A conforming triangulation is shape regular, iff for all elements the condition number of the Jacobian is bounded, i.e.

$$\left\|\frac{d\,x^T(\xi)}{d\xi}\right\| \left\|\left(\frac{d\,x^T(\xi)}{d\xi}\right)^{-1}\right\| \preceq 1 \qquad \forall\, T \in \mathcal{T},\ \forall\, \xi \in T. \tag{2.23}$$

- We define the local mesh size $h(x)$ for $x = x^T(\xi)$ as

$$h(x) = \left\|\frac{d\,x^T(\xi)}{d\xi}\right\|, \tag{2.24}$$

  and

$$h_T = \sup_{x \in T} h(x)$$

  A triangulation is quasi-uniform, iff it is shape regular and there exists one global $h > 0$ such that

$$h \preceq h(x) \preceq h \qquad \forall\, x \in \Omega \tag{2.25}$$

Next, the set of shape functions has to be defined on the reference element. We will need shape functions of full or partial polynomial type, denoted as $\mathcal{P}_k$ and $\mathcal{Q}_k$, respectively.

In addition, often so called bubble functions are necessary. There are element bubble functions which vanish at the boundary of the element, and face bubbles vanishing on all but one face of an element. For example, the lowest order element bubble function for a triangle is $\lambda_1\lambda_2\lambda_3$, and one of the face bubble functions of lowest degree is $\lambda_1\lambda_2$ (the λ_i are the barycentric coordinates).

Third, a set of unisolvent linear functionals are defined on the space of shape functions, called the degrees of freedom. These functionals are most often point evaluations, but also other possibilities like integrals over edges or faces are in use.

These three components define the finite element in the sense of Ciarlet [Cia78]. The global finite element spaces is constructed by transformation of the geometry, the shape functions, and linear functionals. The demand of well-definition of the linear functionals give the desired continuity properties of the global finite element space.

The finite element space will be denoted by V_h, where h is the global mesh size parameter for quasi-uniform triangulations, and just notation for more general conforming triangulations. It is spanned by a set of basis functions, i.e.

$$V_h = \operatorname{span}\{\varphi_i\}. \tag{2.26}$$

The global linear functionals $l_i(.)$ shall form a bi-orthogonal basis for V_h^*. Then a finite element function $u_h \in V_h$ can be represented by

$$u_h = \sum_{i=1}^{N} l_i(u_h)\varphi_i. \tag{2.27}$$

According to the continuity, the finite element space is a subspace of the Sobolev space of order s. Including also the cases $V_h \not\subset H^k$ we define the broken Sobolev norms as

$$\|v_h\|_{k,h} := \Big(\sum_{T\in\mathcal{T}} \|v_h\|_{k,T}^2\Big)^{1/2}, \tag{2.28}$$

for $k \in \mathbb{N}_0$, and for positive parameters by interpolation between these norms.

On finite element spaces on shape regular meshes we can estimate Sobolev norms of higher order by lower order ones. These are inverse inequalities. For $v_h \in V_h$ there holds

$$\|v_h\|_{k,T} \preceq h_T^{l-k}\,\|v_h\|_{l,T} \qquad k \geq l \geq 0. \tag{2.29}$$

A proof can be found in any finite element book, e.g. [Cia78].

2.2.5 Local interpolation operators

We will often have to approximate a function in a Sobolev space by some finite element function. Therefore we need a mapping $\underline{I_h} : H^s \to V_h$. The mapping should be local, i.e. $(\underline{I_h}u)|_T$ should depend on $u|_{\widetilde{T}}$ only, where $\widetilde{T}$ is close to and not much larger than T.

The approximation shall become better as the image norm gets weaker. The optimal approximation is

$$|u - \underline{I_h}u|_{k,T} \preceq h^{l-k}\,|u|_{l,\widetilde{T}} \qquad 0 \leq k \leq l \tag{2.30}$$

for proper integers k and l. The classical interpolation operators are nodal interpolation operators I_N. They are defined by extending the representation (2.27) to

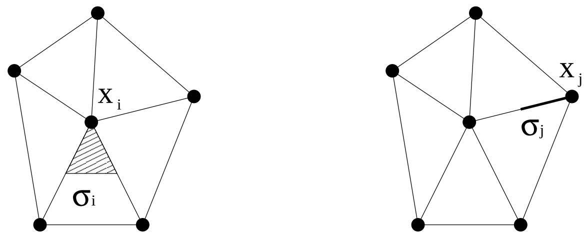
$$I_N u := \sum_{i=1}^{N} l_i(u)\varphi_i \qquad \forall\, u \in W \subset V. \tag{2.31}$$

The space W is such that the linear functionals $l_i(.)$ are well defined on W. By (2.27), the operator is a projection on V_h. The large disadvantage of the interpolation operator (2.31) is the necessary restriction to the sub-space W. If the linear functionals are point evaluation, then $W = H^s$ with $s > d/2$ is required. Thus the approximation estimate (2.30) can hold for $l > d/2$, only.

Other possibilities are interpolation operators of Clément [Clé75] and Scott-Zhang [SZ90] type. They are defined as follows. For each linear functional $l_i(.)$ define a set σ_i such that $l_i(u)$ depends on $u|_{\sigma_i}$, only. It can be a subset of non-zero measure, but also a manifold. Define the $L_2(\sigma_i)$-orthogonal projection Π_i^k onto $\mathcal{P}^k(\sigma_i)$. Then the interpolation operator is

$$I_{SZ} u = \sum_{i}^{N} l_i(\Pi_i^k u)\varphi_i. \tag{2.32}$$

This interpolation operator is a projection if $V_h|_{\sigma_i} \subset \mathcal{P}^k(\sigma_i)$. Two examples to define the sets σ_i for the point evaluation functional $l_i(.)$ in the node x_i are shown below. The set σ_i has non-trivial measure in $\mathbb{R}^2$, while σ_j is a one-dimensional manifold.



The Scott-Zhang projector is well defined for the Sobolev space L_2 iff all sets σ_i have non-zero measure in $\mathbb{R}^d$. Then the approximation inequality (2.30) holds for $l \geq 0$. If the sets are $d-1$ dimensional manifolds, then the operator is well defined on H^s with $s > 1/2$. Then inequality (2.30) holds for $l \geq 1$. The upper bound for l is one plus the minimum of the full polynomial degree of the fe space and of the local space used to construct Π_i^k.

2.2.6 Partition of unity

The partition of unity method is a useful tool for all part of finite element analysis ([BA72], [MB96]). We will use it for the analysis of multigrid and domain decomposition methods.

Let $\{\omega_i\}$ be a decomposition of Ω, i.e.

$$\Omega = \bigcup_{i=1}^{M} \omega_i.$$

We set

$$H_i = \text{diam}\, \omega_i.$$

Each set ω_i is increased to the set $\Omega_i \subset \Omega$ such that

$$\text{dist}\,\{\omega_i, \partial\Omega_i \setminus \partial\Omega\} \succeq H_i.$$

Then there exists functions $\psi_i \in C^\infty$ such that

$$0 \le \psi_i(x) \le 1, \qquad \sum_i \psi_i = 1,$$

and

$$\text{supp}\,\psi_i \subset \Omega_i$$

holds, and all derivatives are bounded by

$$\|\nabla^k \psi_i\|_{L_\infty} \preceq H_i^{-k}.$$

Lemma 2.6. *The multiplication*

$$(\psi_i u)(x) := \psi_i(x) u(x) \qquad \forall\, a.e.\, x \in \Omega$$

is well defined on Sobolev spaces $H^k, k \ge 0$. For $u \in H^k_{0,D}(\Omega)$ there holds the estimate

$$|\psi_i u|_k \preceq \|u\|_{k,\Omega_i} + H^{-k} \|u\|_{0,\Omega_i}$$

2.3 Parameter Dependent Problems

The material presented in this section is referred to [BF91].

2.3.1 Primal and mixed formulation

Let $(V, \|.\|_V, (.,.)_V)$ and $(Q, \|.\|_c, c(.,.))$ be Hilbert spaces. The typical cases are the Sobolev spaces $V = H^1(\Omega)$ and $Q = L_2(\Omega)$ on the domain $\Omega \subset \mathbb{R}^d$. We consider variational problems: Find $u \in V$ such that

$$A^\varepsilon(u, v) = f(v) \qquad \forall\, v \in V \tag{2.33}$$

with the symmetric bilinear form $A^{(\varepsilon)}(.,.) : V \times V \to \mathbb{R}$, and the continuous linear form $f(.) : V \to \mathbb{R}$. The bilinear form is assumed to have the special structure

$$A^\varepsilon(u, v) = a(u, v) + \varepsilon^{-1}\, c(\Lambda u, \Lambda v). \tag{2.34}$$

The bilinear form $a(.,.) : V \times V \to \mathbb{R}$ is assumed to be symmetric, continuous and non-negative. The bilinear from $c(.,.)$ was defined to be the inner product in Q. The operator $\Lambda : V \to Q$ is assumed to be continuous. We assume that ΛV, the range of Λ, is dense in

Q. It is not necessarily closed. The parameter $\varepsilon \in (0, 1]$ is typically small. The problem is of special interest, if

$$V_0 := \text{kern}\,\Lambda$$

is non-trivial. We will skip the superscript $^\varepsilon$ of $A^\varepsilon(.,.)$, later.

We assume that $A^1(u, u)^{1/2}$ is an equivalent norm on V, i.e.

$$\|u\|_V^2 \preceq a(u, u) + c(\Lambda u, \Lambda u) \preceq \|u\|_V^2. \tag{2.35}$$

Norm equivalence (2.35) implies that $A^\varepsilon(.,.)$ is elliptic

$$\|u\|_V^2 \preceq A^\varepsilon(u, u).$$

It is continuous with the parameter dependent estimate

$$A^\varepsilon(u, u) \preceq \varepsilon^{-1}\|u\|_V^2.$$

The theorem of Lax and Milgram ensures a unique solution of (2.33) and the robust stability estimate

$$\|u\|_V \preceq \|f\|_{V^*}$$

Let V_h be a finite element subspace of V. The conforming finite element discretization scheme is: Find $u_h \in V_h$ such that

$$A^\varepsilon(u_h, v_h) = f(v_h) \qquad \forall\, v_h \in V_h. \tag{2.36}$$

Using the norm estimates we get the parameter dependent a priori estimate

$$\|u - u_h\|_V \preceq \varepsilon^{-1/2} \inf_{v_h \in V_h} \|u - v_h\|_V. \tag{2.37}$$

On the other hand, we have the robust estimate in energy norm, i.e.

$$\|u - u_h\|_{A^\varepsilon} \leq \inf_{v_h \in V_h} \|u - v_h\|_{A^\varepsilon}. \tag{2.38}$$

For the problems we are interested in, the non-robust estimate in $\|.\|_V$ norm is more realistic than the robust estimate in $\|.\|_A$ norm. One explanation is that the approximation estimates for standard finite element spaces apply in norms related to V. In Section 2.1.3 we have seen that the estimate (2.37) is not only a theoretical lack, but also occurs in practical computations.

The standard technique is to pass to a mixed formulation. We define the dual variable $p \in Q$ as

$$p := \varepsilon^{-1}\Lambda u. \tag{2.39}$$

We use p in the variational problem (2.33) and get

$$a(u, v) + c(\Lambda v, p) = f(v) \qquad \forall\, v \in V. \tag{2.40}$$

The weak form of (2.39) is

$$c(\Lambda u, q) - \varepsilon\, c(p, q) = 0 \qquad \forall\, q \in Q. \tag{2.41}$$

In mixed finite element publications, usually the bilinear form $b(.,.) : V \times Q \to \mathbb{R}$ defined by

$$b(u, p) := c(\Lambda u, p)$$

is used. Combining equations (2.40) and (2.41), we obtain the mixed variational problem: Find $u \in V$ and $p \in Q$ such that

$$B^\varepsilon((u, p), (v, q)) = f(v) \qquad \forall\, v \in V,\ \forall\, q \in Q \tag{2.42}$$

with the block bilinear form

$$B^\varepsilon((u, p), (v, q)) := a(u, v) + c(\Lambda u, q) + c(\Lambda v, p) - \varepsilon\, c(p, q). \tag{2.43}$$

The mixed bilinear form is well defined for the limit $\varepsilon = 0$. A solution (u, p) of (2.42) is in the space

$$X_0 = \{(u, p) \in X : \Lambda u = \varepsilon\, p\}\,. \tag{2.44}$$

This space will play an essential role for stabilization techniques as well as for iterative solvers. We define the norm

$$\|(u, p)\|_{V \times \varepsilon c} := (\|u\|_V^2 + \varepsilon \|p\|_c^2)^{1/2}$$

on the space $V \times Q$. It degenerates to a semi-norm for $\varepsilon = 0$. The bilinear form $B(.,.)$ is continuous with parameter dependent bounds for that norm, namely

$$\begin{aligned}
B((u, p), (v, q)) &= a(u, v) + c(\Lambda u, q) + c(\Lambda v, p) - \varepsilon\, c(p, q) \\
&\preceq a(u, u)^{1/2} a(v, v)^{1/2} + \|\Lambda u\|_c \|q\|_c + \|\Lambda v\|_c \|p\|_c + \varepsilon \|p\|_c \|q\|_c \\
&\preceq \left(a(u, u) + \|\Lambda u\|_c^2 + \|p\|_c^2 + \varepsilon\, \|p\|_c^2\right)^{1/2} \\
&\quad \times \left(a(v, v) + \|\Lambda v\|_c^2 + \|q\|_c^2 + \varepsilon\, \|q\|_c^2\right)^{1/2} \\
&\preceq \left(\|u\|_V^2 + \|p\|_c^2\right)^{1/2} \left(\|v\|_V^2 + \|q\|_c^2\right)^{1/2} \\
&\preceq \varepsilon^{-1} \left(\|u\|_V^2 + \varepsilon\, \|p\|_c^2\right)^{1/2} \left(\|v\|_V^2 + \varepsilon\, \|q\|_c^2\right)^{1/2} \\
&= \varepsilon^{-1} \|(u, p)\|_{V \times \varepsilon c} \|(v, q)\|_{V \times \varepsilon c}.
\end{aligned} \tag{2.45}$$

On the other hand, $B(.,.)$ provides a uniformly continuous mapping from $[V \times \varepsilon Q]^*$ into $V \times \varepsilon Q$. This is formulated in the following theorem:

Theorem 2.7. *Let $g(.) : Q \to \mathbb{R}$ bet a continuous linear form with norm $\|g\|_{c^*}$. Then the extended mixed problem: Find $(u, p) \in V \times Q$ such that*

$$B^\varepsilon((u, p), (v, q)) = f(v) + g(q) \qquad \forall\, (v, q) \in V \times Q \tag{2.46}$$

has a unique solution. There holds the a priori bound

$$\|u\|_V^2 + \varepsilon^{-1} \|\Lambda u\|_c^2 + \varepsilon \|p\|_c^2 \preceq \|f\|_{V^*}^2 + \varepsilon^{-1} \|g\|_{c^*}^2.$$

Proof. First, we construct a solution by means of the primal problem. Because $\Lambda : V \to Q$ is continuous, and $g(.)$ in Q^*, the functional $g(\Lambda.)$ is continuous on V. Let $u \in V$ be the solution of

$$a(u, v) + \varepsilon^{-1} c(\Lambda u, \Lambda v) = f(v) + \varepsilon^{-1} g(\Lambda v) \qquad \forall\, v \in V.$$

We use the ellipticity of $a(.,.) + c(\Lambda.,.\Lambda)$ to estimate

$$\begin{aligned}
\|u\|_V^2 + \varepsilon^{-1} \|\Lambda u\|_c^2 &\preceq a(u,u) + \varepsilon^{-1}\, c(\Lambda u, \Lambda u) \\
&= f(u) + \varepsilon^{-1} g(\Lambda u) \\
&\preceq \|f\|_{V^*} \|u\|_V + \varepsilon^{-1/2} \|g\|_{c^*}\, \varepsilon^{-1/2} \|\Lambda u\|_c \\
&\preceq \left(\|f\|_{V^*}^2 + \varepsilon^{-1} \|g\|_{c^*}^2 \right)^{1/2} \left(\|u\|_V^2 + \varepsilon^{-1} \|\Lambda u\|_c^2 \right)^{1/2}.
\end{aligned}$$

Dividing by $(\|u\|_V^2 + \varepsilon^{-1} \|\Lambda u\|_c^2)^{1/2}$ gives the bound for u. By Riesz' representation theorem we define $\tilde{g} \in Q$ such that

$$c(\tilde{g}, q) = g(q) \quad \forall\, q.$$

We set

$$p = \varepsilon^{-1} \left(\Lambda u - \tilde{g} \right).$$

It can be bounded by

$$\varepsilon\, \|p\|_c^2 \preceq \varepsilon^{-1} \|\Lambda u\|_c^2 + \varepsilon^{-1} \|g\|_{c^*}^2 \preceq \|f\|_{V^*}^2 + \varepsilon^{-1} \|g\|_{c^*}^2.$$

We verify, that (u, p) is a solution of (2.46), namely for all $(v, q) \in V \times Q$ there holds

$$\begin{aligned}
B^\varepsilon((u,p),(v,q)) &= a(u,v) + c(\Lambda u, q) + c(\Lambda v, p) - \varepsilon\, c(p, q) \\
&= a(u,v) + c(\Lambda u, q) + c(\Lambda v, \varepsilon^{-1}(\Lambda u - \tilde{g})) - \varepsilon c(\varepsilon^{-1}(\Lambda u - \tilde{g}), q) \\
&= a(u,v) + \varepsilon^{-1} c(\Lambda u, \Lambda v) - \varepsilon^{-1} c(\tilde{g}, \Lambda v) + c(\tilde{g}, q) \\
&= f(v) + g(q).
\end{aligned}$$

Finally, we proof that the solution is unique. Otherwise, a non-trivial solution (u, p) of the homogenous problem would satisfy

$$\begin{aligned}
0 &= B^\varepsilon((u,p),(u, \Lambda u - p)) \\
&= a(u,u) + c(\Lambda u, \Lambda u - p) + c(\Lambda u, p) - \varepsilon\, c(p, \Lambda u - p) \\
&= a(u,u) + c(\Lambda u, \Lambda u) + \varepsilon\, c(p,p) - \varepsilon\, c(p, \Lambda u) \\
&\geq a(u,u) + c(\Lambda u, \Lambda u) + \varepsilon\, c(p,p) - \varepsilon/2\, c(p,p) - \varepsilon/2\, c(\Lambda u, \Lambda u) \\
&\geq a(u,u) + 1/2\, c(\Lambda u, \Lambda u) + \varepsilon/2\, c(p,p) \\
&\succeq \|u\|_V^2 + \varepsilon\, \|p\|_c^2 > 0,
\end{aligned}$$

and the proof is complete. $\square$

Up to now, there is no profit of the mixed form. One reason of the problem is that Λ has not necessarily a closed range. Let us assume that $\|.\|_{Q,0}$ is a norm on Q such that

$$\|p\|_{Q,0} \simeq \sup_{v\in V} \frac{c(p,\Lambda v)}{\|v\|_V} \qquad \forall\, p \in Q. \tag{2.47}$$

It is always possible to define the norm $\|.\|_{Q,0}$ by the supremum above. In general, $\|.\|_{Q,0}$ is a weaker norm than $\|.\|_c$. Per definition, Λ has a closed range in

$$Q_0 := \overline{\Lambda V}^{\|.\|_{Q,0}}. \tag{2.48}$$

The bilinear form of the limit $\varepsilon = 0$ is continuous and stable on $V \times Q_0$:

Theorem 2.8 (Brezzi). *The bilinear form*

$$B^0((u,p),(v,q)) = a(u,v) + c(\Lambda u, q) + c(\Lambda v, p)$$

is continuous, i. e.

$$B^0((u,p),(v,q)) \preceq (\|u\|_V^2 + \|p\|_{Q,0}^2)^{1/2} (\|v\|_V^2 + \|q\|_{Q,0}^2)^{1/2}, \tag{2.49}$$

and stable, i. e.

$$\sup_{u\in V, p\in Q_0} \frac{B^0((u,p),(v,q))}{(\|u\|_V^2 + \|p\|_{Q,0}^2)^{1/2}} \succeq (\|v\|_V^2 + \|q\|_{Q,0}^2)^{1/2}, \tag{2.50}$$

on the space $V \times Q_0$.

The theorem is the classical theorem of Brezzi. See [BF91], Prop 1.3.

For the case $\varepsilon > 0$ we need a norm depending on the parameter ε. We define

$$\|p\|_Q := \|p\|_{Q,\varepsilon} := \left(\|p\|_{Q,0}^2 + \varepsilon \|p\|_c^2\right)^{1/2}. \tag{2.51}$$

This norm is equivalent to $\|.\|_c$ for fixed $\varepsilon > 0$, but not necessarily uniformly equivalent with respect to ε. We define the product space

$$X = V \times Q$$

with the norm

$$\|(u,p)\|_X = (\|u\|_V^2 + \|p\|_Q^2)^{1/2}.$$

The following theorem states that $B^\varepsilon(.,.)$ provides an uniform isomorphism form X to X^*:

Theorem 2.9. *Assume the equivalence of norms (2.35) is true. Let* $B^\varepsilon(.,.)$ *and* $\|.\|_X$ *be defined as above. Then the following is true:*

- *The bilinear form* $B^\varepsilon(.,.)$ *is uniformly continuous on* X, *i.e.*

$$B^\varepsilon((u,p),(v,q)) \preceq \|(u,p)\|_X \, \|(v,q)\|_X \qquad \forall\, (u,p) \in X,\ \forall\, (v,q) \in X. \tag{2.52}$$

- *The bilinear form $B^\varepsilon(.,.)$ is uniformly stable on X, i.e.*

$$\sup_{(u,p)\in X} \frac{B^\varepsilon((u,p),(v,q))}{\|(u,p)\|_X} \succeq \|(v,q)\|_X \qquad \forall\,(v,q)\in X. \tag{2.53}$$

Proof. The proof of the continuity follows estimate (2.45). But now, we have the improved estimate

$$c(\Lambda u, q) \leq \|u\|_V \sup_{v\in V} \frac{c(\Lambda v, q)}{\|v\|_V} = \|u\|_V\,\|q\|_{Q,0}$$

for the mixed term. Thus we get uniform continuity. Now, fix $(v,q)\in X$. By definition of the norm $\|.\|_{Q,0}$, there exists a $\tilde v \in V$ such that

$$\frac{c(\Lambda\tilde v, q)}{\|\tilde v\|_V} \succeq \|q\|_{Q,0}.$$

We are free to scale $\tilde v$ such that

$$\|\tilde v\|_v = \|q\|_{Q,0} \qquad \text{and} \qquad c(\Lambda\tilde v, q) \succeq \|q\|^2_{Q,0}.$$

Let (u,p) be the unique solution (by Theorem 2.7) of

$$B^\varepsilon((u,p),(w,r)) = (v,w)_V + c(\Lambda\tilde v, r) + \varepsilon\, c(q,r) \qquad \forall\,(w,r)\in X. \tag{2.54}$$

We will prove that (u,p) is feasible to verify (2.53). First, we see that

$$\begin{aligned} B^\varepsilon((u,p),(v,q)) &= (v,v)_V + c(\Lambda\tilde v, q) + \varepsilon\, c(q,q) \\ &\succeq \|v\|_V^2 + \|q\|^2_{Q,0} + \varepsilon\,\|q\|_c^2 \\ &= \|(v,q)\|_X^2. \end{aligned} \tag{2.55}$$

By the definition of $B^\varepsilon(.,.)$ and (2.54) there holds for all $(w,r)\in X$:

$$\begin{aligned} B^\varepsilon((u-\tilde v,p),(w,r)) &= B^\varepsilon((u,p),(w,r)) - B^\varepsilon((\tilde v,0),(w,r)) \\ &= (v,w)_V + c(\Lambda\tilde v, r) + \varepsilon\, c(q,r) - [a(\tilde v,w) + c(\Lambda\tilde v, r)] \\ &= (v,w)_V - a(\tilde v,w) + \varepsilon\, c(q,r). \end{aligned}$$

We apply Theorem 2.7 to bound

$$\|u-\tilde v\|_V^2 + \varepsilon\,\|p\|_c^2 \preceq \|v\|_V^2 + \|\tilde v\|_V^2 + \varepsilon\|q\|_c^2,$$

and further

$$\begin{aligned} \|u\|_V^2 + \varepsilon\,\|p\|_c^2 &\preceq \|\tilde v\|_V^2 + \|v\|_V^2 + \varepsilon\,\|q\|_c^2 \\ &\preceq \|v\|_V^2 + \|q\|^2_{Q,0} + \varepsilon\,\|q\|_c^2 \\ &= \|(v,q)\|_X^2. \end{aligned}$$

We are left to estimate $\|p\|_{Q,0}$. Applying test functions $(w,0)$ to (2.54) we get

$$B^\varepsilon((u,p),(w,0)) = a(u,w) + c(\Lambda w, p) = (w,v)_V \qquad \forall\, w \in V,$$

and thus

$$\begin{aligned} \|p\|_{Q,0} &= \sup_{w\in V} \frac{c(\Lambda w,p)}{\|w\|_V} = \sup_{w \in V} \frac{(w,v)_V - a(u,w)}{\|w\|_V} \\ &\preceq \|v\|_V + \|u\|_V \preceq \|(v,q)\|_X. \end{aligned}$$

We got $\|(u,p)\|_X \preceq \|(v,q)\|_X$. Combining with (2.55) we get

$$\frac{B^\varepsilon((u,p),(v,q))}{\|(u,q)\|_X} \succeq \frac{\|(v,q)\|_X^2}{\|(v,q)\|_X} = \|(v,q)\|_X,$$

and the proof is complete. □

There are alternative assumptions to provide an uniform isomorphism. The weakest conditions are formulated in [Bra96].

2.3.2 Stabilization techniques

An advantageous property is the ellipticity of the bilinear form $a(.,.)$, i.e.

$$a(u,u) \geq \|u\|_V^2 \qquad \forall\, u \in V. \tag{2.56}$$

This is not really an additional assumption. Due to (2.35) it can always be constructed. Let us split $c(.,.)$ into

$$\varepsilon^{-1} c(u,v) = \tilde{c}(u,v) + \varepsilon^{-1}\hat{c}(u,v)$$

such that $(Q, \|.\|_{\hat{c}})$ is a Hilbert space, and

$$\hat{a}(u,v) := a(u,v) + \tilde{c}(\Lambda u, \Lambda v)$$

is elliptic on V. One possibility is to set $\tilde{c}(.,.) = 1/2\, c(.,.)$. Then the bilinear form

$$\begin{aligned} A^\varepsilon(u,v) &= a(u,v) + \varepsilon^{-1} c(\Lambda u, \Lambda v) \\ &= a(u,v) + \tilde{c}(\Lambda u, \Lambda v) + \varepsilon^{-1}\hat{c}(\Lambda u, \Lambda v) \\ &= \hat{a}(u,v) + \varepsilon^{-1}\hat{c}(\Lambda u, \Lambda v) \end{aligned}$$

is a splitting with elliptic part $\hat{a}(.,.)$. We substitute $\hat{a}(.,.)$ by $a(.,.)$ and $\hat{c}(.,.)$ by $c(.,.)$, and return to a problem of the original structure. This approach can be found in [AB93]. There are many possibilities to construct the splitting. It may depend on the finite element mesh, as well as on the parameter.

An alternative approach (used e.g. in [CS98]) with the same result is the following. The solution $(u,p) \in X$ of the variational problem

$$B^\varepsilon((u,p),(v,q)) = f(v) \qquad \forall\, (v,q) \in X$$

is in the space X_0, i.e.

$$\Lambda u - \varepsilon\, p = 0.$$

Thus, (u,p) solves also the stabilized problem

$$\widetilde{B}^{\varepsilon}((u,p),(v,q)) = f(v) \qquad \forall\,(v,q) \in X$$

with the stabilized bilinear form

$$\widetilde{B}^{\varepsilon}((u,p),(v,q)) := B^{\varepsilon}((u,p),(v,q)) + \tilde{c}(\Lambda u - \varepsilon p, \Lambda v - \varepsilon q).$$

The forms $\hat{c}(.,.), \tilde{c}(.,.)$, and $\hat{a}$ are the same from above. One verifies that the bilinear form $\widetilde{B}^{\varepsilon}(.,.)$ has the alternative representation

$$\begin{aligned} \widetilde{B}((u,p),(v,q)) &= a(u,v) + \tilde{c}(\Lambda u, \Lambda v) + c(\Lambda u, q) - \varepsilon\, \tilde{c}(\Lambda u, q) \\ &+ c(\Lambda v, p) - \varepsilon\, \tilde{c}(\Lambda v, p) - \varepsilon\, c(p,q) + \varepsilon^2\, \tilde{c}(p,q) \\ &= \hat{a}(u,v) + \hat{c}(\Lambda u, q) + \hat{c}(\Lambda v, p) - \varepsilon\, \hat{c}(p,q). \end{aligned}$$

After renaming $\hat{a}(.,.)$ to $a(.,.)$ and $\hat{c}(.,.)$ to $c(.,.)$ we are back at the original structure.

An other version of stabilization is analyzed in . It makes the discrete stability condition hold for any finite element pairing. But the technique cannot be applied in primal variables. For relations between stabilization methods and bubble functions see [BBF+92], [Hug95], [BFHR97].

2.3.3 Non-conforming and mixed discretization techniques

We have seen already in Section 2.1.3 that conforming discretization methods for parameter dependent problems may deteriorate as the parameter gets small. The key to construct robust discretization schemes is the relation to mixed finite element methods. Let

$$V_h \subset V \qquad \text{and} \qquad Q_h \subset Q$$

be finite element spaces build on the shape regular triangulation $\mathcal{T}_h$ of the domain Ω. We set

$$X_h = V_h \times Q_h.$$

The idea of the discretization is the reduction of the operator Λ. We define the projection $P_h^c : Q \to Q_h$ by

$$c(P_h^c p, q_h) = c(p, q_h) \qquad \forall\, p \in Q,\ \forall\, q_h \in Q_h.$$

The reduced operator $\Lambda_h : V \to Q_h$ is defined as

$$\Lambda_h := P_h^c \Lambda.$$

We define $A_h^{\varepsilon}(.,.) : V \times V \to \mathbb{R}$ as

$$A_h^{\varepsilon}(u,v) := a(u,v) + \varepsilon^{-1} c(\Lambda_h u, \Lambda_h v). \tag{2.57}$$

The non-conforming finite element problem is: Find $u_h \in V_h$ such that

$$A_h^\varepsilon(u_h, v_h) = f(v_h) \qquad \forall\, v_h \in V_h. \tag{2.58}$$

We assume that A_h^1 is stable, i.e.

$$\|u_h\|_V^2 \preceq a(u_h, u_h) + c_h(\Lambda_h u_h, \Lambda_h u_h) \qquad \forall\, u_h \in V_h. \tag{2.59}$$

This is in general an additional assumption. It is harder to fulfill as the space V_h gets larger and Q_h gets smaller. If the bilinear form $a(.,.)$ is elliptic on V, then (2.59) is fulfilled trivially, and the space V_h may become arbitrarily large. This might be the reason to apply stabilization techniques.

We define the mixed form

$$B_h^\varepsilon((u_h, p_h), (v_h, q_h)) = a(u_h, v_h) + c(\Lambda_h u_h, q_h) + c(\Lambda_h v_h, p_h) - \varepsilon c(p_h, q_h)$$

and formulate the equivalent mixed finite element problem: Find $(u_h, p_h) \in X_h$ such that

$$B_h^\varepsilon((u_h, p_h), (v_h, q_h)) = f(v_h) \qquad \forall\, (v_h, q_h) \in X_h.$$

The variational definition of Λ_h gives

$$c(\Lambda_h u_h, q_h) = c(P_h^c \Lambda u_h, q_h) = c(\Lambda u_h, q_h).$$

Thus $B_h^\varepsilon(.,.)$ coincides with $B^\varepsilon(.,.)$ on the finite element space X_h, and the mixed finite element problem with the original form $B^\varepsilon(.,.)$: Find $(u_h, p_h) \in X_h$ such that

$$B^\varepsilon((u_h, p_h), (v_h, q_h)) = f(v_h) \qquad \forall\, (v_h, q_h) \in X_h$$

is equivalent to the non-conforming primal problem (2.58). Theorem 2.7 can be applied to prove stability and continuity of $B^\varepsilon(.,.)$ on the space $(X_h, \|.\|_{V\times\varepsilon c})$, but with bounds depending on ε.

The norm $\|.\|_X$ provides an uniform estimates for stability and continuity of $B^\varepsilon(.,.)$ on X. We need an additional assumption to obtain an uniform stability estimates on the subspace X_h.

Theorem 2.10 (Fortin's criterion). *Assume there exists an operator*

$$I_h^F : V \to V_h \tag{2.60}$$

which is continuous

$$\|I_h^F\|_V \preceq 1 \tag{2.61}$$

and fulfills the property

$$c(\Lambda I_h^F u, q_h) = c(\Lambda u, q_h) \qquad \forall\, u \in V,\ \forall\, q_h \in Q_h. \tag{2.62}$$

Then the bilinear form $B^\varepsilon(.,.)$ is stable on X_h, i.e.

$$\sup_{(u_h,p_h)\in X_h} \frac{B^\varepsilon((u_h, p_h), (v_h, q_h))}{\|(u_h, p_h)\|_X} \succeq \|(v_h, q_h)\|_X \qquad \forall\, (v_h, q_h) \in X_h. \tag{2.63}$$

Remark 2.11. *We can consider the parameter dependent problem as penalty approximation of a constrained minimization problem. Then property (2.62) means, that the constraints $\Lambda u = 0$ should be preserved by the Fortin operator in mean.*

Proof. In analogy to (2.47), let $\|.\|_{Q_h,0}$ be a norm such that

$$\|p_h\|_{Q_h,0} \simeq \sup_{v_h \in V_h} \frac{c(p_h, \Lambda v_h)}{\|v_h\|_V} \qquad \forall\, p \in Q_h. \tag{2.64}$$

We define the parameter dependent norm

$$\|p_h\|_{Q_h} := \|p\|_{Q_h,\varepsilon} := \left(\|p_h\|^2_{Q_h,0} + \varepsilon \|p_h\|_c^2\right)^{1/2}. \tag{2.65}$$

By Theorem 2.9, $B^\varepsilon(.,.)$ is stable and continuous on X_h with respect to the norm $\|.\|_{V\times Q_h}$. In general, the norm $\|.\|_{Q_h}$ is weaker then $\|.\|_Q$, because the supremum is taken over a smaller space. But, if there exists a Fortin operator I_h^F, the norms are equivalent. To prove, we fix a $p_h \in Q_h$. Let $u \in V$ such that

$$\frac{c(\Lambda u, p_h)}{\|u\|_V} \succeq \sup_{v \in V} \frac{c(\Lambda v, p_h)}{\|v\|_V} = \|p_h\|_{Q,0}.$$

Then, we can estimate

$$\begin{aligned} \|p_h\|_{Q_h,0} &= \sup_{v_h \in V_h} \frac{c(\Lambda v_h, p_h)}{\|v_h\|_V} \geq \frac{c(\Lambda I_h^F u, p_h)}{\|I_h^F u\|_V} \\ &= \frac{c(\Lambda u, p_h)}{\|I_h^F u\|_V} \succeq \frac{c(\Lambda u, p_h)}{\|u\|_V} \succeq \|p_h\|_{Q,0}. \end{aligned}$$

Thus, the norms $\|.\|_{V\times Q_h}$ and $\|.\|_{V\times Q}$ are equivalent, and $B^\varepsilon(.,.)$ is stable on X_h with respect to the norm $\|.\|_X$. ☐

Fortin's criterion is based on the limit problem with $\varepsilon = 0$. In Theorem 2.10 we proved equivalence of $\|.\|_{Q,0}$ and $\|.\|_{Q_h,0}$ to conclude that $\|.\|_Q$ and $\|.\|_{Q_h}$ are equivalent. We will consider problems, where the norms depend on the parameter. Then, the later norms can be equivalent, but the former are not. The term $\varepsilon\|.\|_c^2$ may have a stabilizing effect. A relaxed version of Fortin's criterion can handle that case:

Theorem 2.12 (Relaxed Fortin's criterion). *Assume, we can split Q_h into*

$$Q_h = Q_{h,0} + Q_{h,1}$$

with $Q_{h,0} \subset Q_h$ and $Q_{h,1} \subset Q_h$. The splitting is assumed to be stable in the norm $\|.\|_c$, i.e. for all $p_h \in Q_h$ there exists $p_0 \in Q_{h,0}$ and $p_1 \in Q_{h,1}$ such that

$$p_h = p_0 + p_1 \qquad \textit{and} \qquad \|p_0\|_c + \|p_1\|_c \preceq \|p_h\|_c. \tag{2.66}$$

We assume for $p_0 \in Q_{h,0}$ there holds the estimate

$$\|p_0\|_Q \preceq \varepsilon^{1/2} \|p_0\|_c. \tag{2.67}$$

Assume that the Fortin operator I_h^F of Theorem 2.10 fulfills the relaxed condition

$$c(\Lambda I_h^F u, q_1) = c(\Lambda u, q_1) \qquad \forall\, u \in V,\ \forall\, q_1 \in Q_{h,1} \tag{2.68}$$

instead of (2.62).

Then the form $B^\varepsilon(.,.)$ is stable on X_h with respect to the norm $\|.\|_X$.

Proof. Again, we have to prove that

$$\|p_h\|_Q \preceq \|p_h\|_{Q_h}$$

holds for $p_h \in Q_h$. Let $p_h = p_0 + p_1$ such that (2.66) holds. Then there holds

$$\begin{aligned} \|p_h\|_Q &\leq \|p_0\|_Q + \|p_1\|_Q \\ &\simeq \|p_0\|_Q + \sup_{v \in V} \frac{c(p_1, \Lambda\, v)}{\|v\|_V} + \varepsilon^{1/2} \|p_1\|_c. \end{aligned}$$

We use the Fortin operator like in the previous theorem to reduce the supremum to the finite element space:

$$\begin{aligned} \|p_h\|_Q &\preceq \|p_0\|_Q + \sup_{v_h \in V_h} \frac{c(p_1, \Lambda\, v_h)}{\|v_h\|_V} + \varepsilon^{1/2} \|p_1\|_c \\ &= \|p_0\|_Q + \|p_1\|_{Q_h,0} + \varepsilon^{1/2} \|p_1\|_c. \end{aligned}$$

We proceed with the triangle inequality applied to $p_1 = p - p_0$, the estimate $\|.\|_{Q_h,0} \preceq \|.\|_{Q,0} \leq \|.\|_Q$, assumption (2.67) and assumption (2.66):

$$\begin{aligned} \|p_h\|_Q &\preceq \|p_0\|_Q + \|p\|_{Q_h,0} + \|p_0\|_{Q_h,0} + \varepsilon^{1/2} \|p_1\|_c \\ &\preceq \|p_0\|_Q + \|p\|_{Q_h,0} + \varepsilon^{1/2} \|p_1\|_c \\ &\preceq \varepsilon \|p_0\|_c + \|p\|_{Q_h,0} + \varepsilon^{1/2} \|p_1\|_c \\ &\preceq \|p\|_{Q_h,0} + \varepsilon^{1/2} \|p\|_c \\ &= \|p\|_{Q_h}. \end{aligned}$$

The theorem is proven. □

A direct application of the Fortin operator I_h^F is as continuous interpolation operator for the parameter dependent problem in primal variables:

Theorem 2.13 (Robust interpolation operator). *Let I_h^F be an operator as in Theorem 2.12. Then it is bounded uniformly with respect to the energy norms*

$$\|I_h^F u\|_{A_h} \preceq \|u\|_A \qquad \forall\, u \in V. \tag{2.69}$$

Proof. We fix $u \in V$, and estimate the terms of

$$\|I_h^F u\|_{A_h}^2 = \|I_h^F u\|_a^2 + \varepsilon^{-1} \|\Lambda_h I_h^F u\|_c^2.$$

The first one follows by continuity of I_h^F, i.e.

$$\|I_h^F u\|_a^2 \preceq \|I_h^F u\|_V^2 \preceq \|u\|_V^2 \preceq \|u\|_A^2.$$

To estimate the second term, we use the $\|.\|_c$-stable decomposition of Q_h:

$$\begin{aligned} \|\Lambda_h I_h^F u\|_c^2 &= \sup_{p_h \in Q_h} \frac{c(\Lambda_h I_h^F u, p_h)^2}{\|p_h\|_c^2} \\ &\preceq \sup_{\substack{q_0 \in Q_{h,0} \\ q_1 \in Q_{h,1}}} \frac{c(\Lambda_h I_h^F u, p_0 + p_1)^2}{\|p_0\|_c^2 + \|p_1\|_c^2}. \end{aligned}$$

The supremum is taken at the solution of the variational problem

$$c(p_0, q_0) + c(p_1, q_1) = c(\Lambda_h I_h^F u, q_0 + q_1) \qquad \forall\, q_0 \in Q_{h,0},\ q_1 \in Q_{h,1},$$

and evaluates to

$$c(p_0, p_0) + c(p_1, p_1).$$

The variational problem consists of two decoupled problems. The same variational problems and the same value are obtained by two decoupled supremes:

$$c(p_0, p_0) + c(p_1, p_1) = \sup_{q_0 \in Q_{h,0}} \frac{c(\Lambda_h I_h^F u, q_0)^2}{\|q_0\|_c^2} + \sup_{q_1 \in Q_{h,1}} \frac{c(\Lambda_h I_h^F u, q_1)^2}{\|q_1\|_c^2}. \tag{2.70}$$

To estimate the first term, we use $c(\Lambda_h v, q_h) \leq \|v\|_V \|q\|_Q$ and assumption (2.67):

$$\sup_{q_0 \in Q_{h,0}} \frac{c(\Lambda_h I_h^F u, q_0)^2}{\|q_0\|_c^2} \leq \sup_{q_0 \in Q_{h,0}} \frac{\|I_h^F u\|_V^2 \|q_0\|_Q^2}{\|q_0\|_c^2} \leq \varepsilon \, \|u\|_V^2.$$

To estimate the second term, we use (2.68) and Cauchy-Schwarz for $c(.,.)$:

$$\sup_{q_1 \in Q_{h,1}} \frac{c(\Lambda_h I_h^F u, q_1)^2}{\|q_1\|_c^2} = \sup_{q_1 \in Q_{h,1}} \frac{c(\Lambda_h u, q_1)^2}{\|q_1\|_c^2} \leq \|\Lambda_h u\|_c^2 \|\Lambda u\|_c^2.$$

The combination gives

$$\varepsilon^{-1} \|\Lambda_h I_h^F u\|_c^2 \preceq \|u\|_A^2,$$

and the theorem is proven. ☐

Assume stability of $B^\varepsilon(.,.)$ on X_h with respect to the norm $\|.\|_X$. Then the finite element approximation is close to best approximation, i.e.

$$\|(u - u_h, p - p_h)\|_X \preceq \inf_{(v_h, q_h)} \|(u - v_h, p - q_h)\|_X.$$

To obtain a rate of convergence, we have to leave the natural space X. Assume, there exist spaces

$$V^+ \subset V \subset V^- \qquad \text{and} \qquad Q^+ \subset Q$$

with corresponding norms $\|.\|_{V^+}$, $\|.\|_{V^-}$, and $\|.\|_{Q^+}$. The embedding is assumed to be dense and continuous. The standard case is $V^+ = H^2(\Omega)$, $V^- = L_2(\Omega)$, and $Q^+ = H^1(\Omega)$. Let I_h^V and I_h^Q be interpolation operators fulfilling the approximation and continuity properties

$$\hbar^{-1}\|u - I_h^V u\|_{V^-} + \|u - I_h^V u\|_V \preceq \hbar\|u\|_{V^+} \tag{2.71}$$

$$\hbar^{-1}\|u - I_h^V u\|_{V^-} + \|I_h^V u\|_V \preceq \|u\|_V \tag{2.72}$$

$$\|I_h^V u\|_{V^-} \preceq \|u\|_{V^-}, \tag{2.73}$$

and

$$\|p - I_h^Q p\|_Q \preceq \hbar\|p\|_{Q^+} \tag{2.74}$$

for proper u and p. The parameter $\hbar$ depends on the finite element space. Usually, it will be the mesh size $\hbar = h$. If we consider problems of less then full elliptic regularity, we set $\hbar = h^\alpha$ with $\alpha \in (0, 1)$.

The dual norm $\|.\|_{(V^-)^*}$ is defined as usual

$$\|f\|_{(V^-)^*} := \sup_{v \in V^-} \frac{f(v)}{\|v\|_{V^-}},$$

and the dual space is

$$(V^-)^* = \{f \in V^* : \|f\|_{(V^-)^*} < \infty\}.$$

We have proved for $f \in V^*$ the solution is in $(u, p) \in V \times Q$. Now we assume that a more regular right hand side $f \in (V^-)^*$ provides a more regular solution $(u, p) \in V^+ \times Q^+$, and there holds the regularity estimate

$$\|u\|_{V^+} + \|p\|_{Q^+} \preceq \|f\|_{(V^-)^*}. \tag{2.75}$$

We finish the section about discretization by collecting the following theorems:

Theorem 2.14. *Let $A^\varepsilon(.,.)$, $B^\varepsilon(.,.)$ and $A_h^\varepsilon(.,.)$ be as defined above. Assume there holds equivalence of norms (2.35), the discrete counterpart (2.59), and there exists a Fortin operator as in Theorem 2.12. Then the discretization error is bounded by the approximation error uniformly in $\varepsilon \in (0, 1)$:*

$$\|(u - u_h, p - p_h)\|_{X,h} \preceq \inf_{(v_h, q_h)} \|(u - v_h, p - q_h)\|_{X,h}. \tag{2.76}$$

If, in addition, regularity is available, we obtain convergence rate estimates

Theorem 2.15. *Let the assumptions of Theorem 2.14 be fulfilled. Additionally, assume that the regularity estimate (2.75) is true. Let $I_h^X = (I_h^V, I_h^Q)$ be an interpolation operator fulfilling (2.71) and (2.74). Then there holds the a priori estimate*

$$\hbar^{-1}\|u - u_h\|_{V^-} + \|u - u_h\|_{V,h} + \|p - p_h\|_{Q,h} \preceq \hbar\, \|f\|_{(V^-)^*}. \tag{2.77}$$

2.4 Some Examples

We consider some parameter dependent problems fitting into the abstract framework.

2.4.1 The Timoshenko beam

In Section 2.1.3 we shortly considered the beam model of Timoshenko. Now, we apply the abstract machinery to this example. The domain Ω is the interval $(0,1)$, and the space $\widetilde{V}$ is $[H^1(\Omega)]^2$. We define the bilinear form

$$A^\varepsilon((w,\beta),(v,\eta)) = (\beta',\eta')_0 + t^{-2}(w'-\beta, v'-\eta)_0. \tag{2.78}$$

The small parameter is the square of the thickness

$$\varepsilon = t^2.$$

The space for the dual variable is $Q = L_2(\Omega)$. With the definitions

$$\begin{aligned} a((w,\beta),(v,\eta)) &:= (\beta',\eta')_0, \\ c(p,q) &:= (p,q)_0, \\ \Lambda(w,\beta) &:= w'-\beta, \end{aligned}$$

the bilinear form has the structure (2.34).

The kernel of the semi norm $\|.\|_A$ consists of the functions

$$V_{00} = \{(w,\beta) = (a+bx, b) : a,b \in \mathbb{R}\}.$$

We obtain the closed space $V \subset \widetilde{V}$ be posing enough boundary conditions such that

$$V_{00} \cap V = \{0\}.$$

We check the ellipticity of $A^1(.,.) = a(.,.) + c(\Lambda., \Lambda.)$ on V. First, we verify

$$\begin{aligned} \|(w,\beta)\|_V^2 &= \|w'\|_0^2 + \|w\|_0^2 + \|\beta'\|_0^2 + \|\beta\|_0^2 \\ &\le 2\,\|w'-\beta\|_0^2 + 2\,\|\beta\|_0^2 + \|w\|_0^2 + \|\beta'\|_0^2 + \|\beta\|_0^2 \\ &\preceq \|(w,\beta)\|_{A^1}^2 + \|(w,\beta)\|_0^2. \end{aligned}$$

Theorem 2.4 proves that the norms $\|.\|_V$ and $\|.\|_{A^1}$ are equivalent on V.

Following the abstract recipe, the dual variable is

$$p = \varepsilon^{-1}\Lambda(w,\beta) = t^{-2}(w'-\beta).$$

The bilinear form $B^\varepsilon(.,.)$ evaluates to

$$B^\varepsilon((w,\beta,p),(v,\eta,q)) = (\beta',\eta')_0 + (w'-\beta,q)_0 + (v'-\eta,p)_0 - t^2(p,q)_0.$$

We define the norms on the dual space and on the product space

$$\begin{aligned} \|p\|_Q &:= \|p\|_0, \\ \|(w,\beta,p)\|_X &:= (\|(w,\beta)\|_1^2 + \|p\|_Q^2)^{1/2}. \end{aligned}$$

We have to prove norm equivalence (2.51), i. e.

$$\|p\|_0^2 \simeq \sup_{(w,\beta)\in V} \frac{(p, w'-\beta)_0^2}{\|(w,\beta)\|_1^2} + \varepsilon\|p\|_0^2 \qquad \forall\, q \in Q = L_2.$$

Clearly, The left hand side dominates the right hand side. To verify the other direction, we have to check the LBB condition. We will prove it for boundary conditions leading to the smallest primal space V.

Theorem 2.16. *Let Ω be the interval $(0,L)$ with $L \preceq 1$. Then there holds the LBB condition*

$$\sup_{(w,\beta)\in[H_0^1(\Omega)]^2} \frac{(w'-\beta,p)_0}{\|w\|_1 + \|\beta\|_1} \succeq L\|p\|_0 \qquad \forall\, p \in L_2(\Omega). \tag{2.79}$$

On the reduced space $L_2^0 = \{q \in L_2 : \int q = 0\}$, there holds the LBB condition uniformly in L, even on the reduced space $H_0^1 \times \{0\}$:

$$\sup_{w\in H_0^1(\Omega)} \frac{(w',p)_0}{\|w\|_1} \succeq \|p\|_0 \qquad \forall\, p \in L_2^0(\Omega). \tag{2.80}$$

Proof. We fix a $p \in L_2$. Compute

$$\begin{aligned} \beta(x) &= \frac{\int_\Omega p\,d\xi}{\int_\Omega \xi(L-\xi)\,d\xi}\, x(L-x), \\ w(x) &= \int_0^x p(\xi) + \beta(\xi)\, d\xi. \end{aligned}$$

It fulfills $\int_0^L \beta\,d\xi = \int_0^L p\,d\xi$. There holds

$$\|\beta\|_1 \preceq L^{-1}\,\|\beta\|_0 \preceq L^{-1}\,\|p\|_0,$$

and $w \in H_0^1$ with $w' = p + \beta$ such that

$$\|w\|_1 \preceq \|p+\beta\|_0 \preceq \|p\|_0.$$

This choice gives the LBB condition, namely

$$\frac{(w'-\beta,p)_0}{\|w\|_1 + \|\beta\|_1} = \frac{(p,p)_0}{\|w\|_1 + \|\beta\|_1} \succeq \frac{\|p\|_0^2}{\|p\|_0 + L^{-1}\|p\|_0} \succeq L\,\|p\|_0.$$

If $p \in L_2^0$, then the choice gives $\beta = 0$, and no factor L enters the estimates. ∎

Next, we choose finite element subspaces for V and Q, namely piecewise linear elements for the primal and piecewise constants for the dual variables, i.e.

$$\begin{aligned} V_h &= \{v_h \in V : v_h|_T \in [\mathcal{P}^1]^2\}, \\ Q_h &= \{q_h \in Q : q_h|_T \in \mathcal{P}^0\}. \end{aligned}$$

This leads to the discrete operator

$$\Lambda_h(w_h, \beta_h) = \overline{(w_h' - \beta_h)}^h.$$

The non-conforming bilinear form $A_h^\varepsilon(.,.)$ is

$$A_h^\varepsilon((w_h, \beta_h), (v_h, \eta_h)) = (\beta_h', \eta_h')_0 + t^{-2}(\overline{w_h' - \beta_h}^h, \overline{v_h' - \eta_h}^h)_0.$$

Since the part $a(.,.)$ was not elliptic on V, we have to check ellipticity of $A_h^1(.,.)$ on V_h. We use that $\overline{w_h'}^h = w_h'$, and estimate

$$\begin{aligned} \|(w_h, \beta_h)\|_V^2 &\preceq \|(w_h, \beta_h)\|_{A^1}^2 = \|\beta_h'\|_0^2 + \|w_h' - \beta_h\|_0^2 \\ &= \|\beta_h'\|_0^2 + \|\overline{w_h' - \beta_h}^h + \beta_h - \overline{\beta_h}^h\|_0^2 \preceq \|\beta_h'\|_0^2 + \|\overline{w_h' - \beta_h}^h\|_0^2 + \|\beta_h - \overline{\beta_h}^h\|_0^2 \\ &\preceq \|\beta_h'\|_0^2 + \|\overline{w_h' - \beta_h}^h\|_0^2 + h\,\|\beta_h'\|_0^2 \preceq \|(w_h, \beta_h)\|_{A_h^1}^2. \end{aligned}$$

We will construct the Fortin operator I_h^F by

$$I_h^F = I_h^{F,1} + I_h^{F,2}(I - I_h^{F,1}). \tag{2.81}$$

The operator I_h^1 is the nodal interpolation operator fulfilling optimal approximation estimates

$$\|I_h^1(w,\beta)\|_1 + h^{-1}\|(w,\beta) - I_h^1(w,\beta)\|_0 \preceq \|(w,\beta)\|_1.$$

The operator $I_h^{F,2}$ fulfilling the constraints is constructed as follows. Let $\mathcal{M}_h = \{M\}$ be a macro triangulation by combining two by two elements. We set $M = (T_1, T_2)$. Then

$$(w_h, \beta_h) := I_h^{F,2}(w, \beta) \tag{2.82}$$

if

$$\begin{gathered} w_h = \beta_h = 0 \qquad \text{on } \partial M \; \forall\, M \in \mathcal{M}, \\ \int_M \beta_h \,\mathrm{dx} = \int_M (\beta - w')\,\mathrm{dx} \\ \int_T w_h' \,\mathrm{dx} = \int_T (w' - \beta + \beta_h)\,\mathrm{dx}. \end{gathered}$$

One checks that the construction is possible, and simple scaling arguments give

$$h^{-1}\,\|w_h\|_0 + \|\beta_h\|_0 \preceq \|w\|_1 + \|\beta\|_0. \tag{2.83}$$

We see that both variables scale differently. Especially, the estimate $\|\beta_h\|_1 \preceq h^{-1}\,\|\beta_h\|_0 \preceq h^{-1}\|w\|_1$ seems troublesome. Fortunately, the nodal interpolation operator fulfills additionally

$$\int_M (w - I_h^{F,1} w)' \,\mathrm{dx} = 0.$$

Thus, for

$$(w_h, \beta_h) := I_h^{F,2}(I - I_h^{F,1}), (w, \beta)$$

β_h does not depend on w. This leads to the improved estimate $\|\beta_h\|_0 \preceq \|\beta\|_0$, and one gets continuity of $I_h^{F,2}(I - I_h^{F,1})$:

$$\begin{aligned}
\|(w_h, \beta_h)\|_1 &\preceq h^{-1}\|w_h\|_0 + h^{-1}\|\beta_h\|_0 \\
&\preceq h^{-1}\|(I - I_h^{F,1})w\|_0 + h^{-1}\|(I - I_h^{F,1})\beta\|_0 \\
&\preceq \|(w, \beta)\|_1.
\end{aligned}$$

Thus also the Fortin operator I_h^F is continuous with respect to the norm $\|.\|_V$. An additional property is that the operator maps $(w, 0)$ to $(w_h, 0)$. Thus, there hold corresponding stability conditions for the finite element space V_h as proved in Theorem 2.16 for V.

There is full regularity available. For the extended problem: Find $(w, \beta) \in V$ such that

$$A^{\varepsilon}((w, \beta), (v, \eta)) = (f, v)_0 + (\delta, \eta)_0 \qquad \forall\, (w, \eta) \in V$$

there holds

$$\|w\|_2 + \|\beta\|_2 \preceq \|f\|_0 + \|\delta\|_0.$$

A proof is found in [Arn81].

2.4.2 Nearly incompressible materials

We consider the problem of linear elasticity. The problem is to find $u \in V := [H^1_{0,D}(\Omega)]^2$ such that

$$2\mu \int_\Omega e(u) : e(v) \,\mathrm{dx} + \lambda \int_\Omega \operatorname{div} u \operatorname{div} v \,\mathrm{dx} = \int_\Omega \tilde{f}^T v \,\mathrm{dx}, \tag{2.84}$$

with the positive constants λ and μ of Lamé, the strain operator $e(u) := 0.5(\nabla u + (\nabla u)^T)$, and the volume force $\tilde{f} \in [L_2(\Omega)]^2$. We are interested in the nearly incompressible case, i.e. the Poisson ration ν is close to 0.5. Then the parameter $\varepsilon := 2\mu/\lambda$ becomes small. After dividing by 2μ, the primal bilinear form $A^\varepsilon(.,.)$, the components $a(.,.)$ and $c(.,.)$, the operator Λ, and the mixed bilinear form are

$$\begin{aligned}
A^\varepsilon(u, v) &= (e(u), e(v))_0 + \varepsilon^{-1}(\operatorname{div} u, \operatorname{div} v)_0, \\
a(u, v) &= (e(u), e(v))_0 \\
c(p, q) &= (p, q)_0 \\
\Lambda &= \operatorname{div}, \\
B^\varepsilon((u, p), (v, q)) &= (e(u), e(v))_0 + (\operatorname{div} u, q)_0 + (\operatorname{div} v, p)_0 - \varepsilon\, (p, q)_0.
\end{aligned}$$

Due to Korn's inequality, and sufficient boundary conditions, $a(.,.)$ is continuous on V. We set

$$\|p\|_Q^2 = \|p\|_{L_2/\mathbb{R}}^2 + \varepsilon \|p\|_{L_2}^2$$

for pure Dirichlet boundary conditions, and $\|p\|_Q = \|.\|_{L_2}$ else. The LBB - condition of Stokes' problem

$$\sup_{u \in [H_0^1(\Omega)]^2} \frac{(\operatorname{div} u, p)_0}{\|u\|_1} \qquad \forall\, q \in L_2^0$$

is non-trivial, see [DL76]. On a convex polygonal domain and pure Dirichlet boundary conditions there holds the regularity pick up

$$\|u\|_2 + \|p\|_1 \preceq \|f\|_0$$

uniformly in the parameter ε, see [BS92c], and [KO76] for the limit problem of Stokes. The stability theory for finite element discretization of Stokes problem seems to be done [Ste84], [BF91]. A stable finite element pairing for Stokes is the combination $P_2 - P_0$. Thus we set

$$\begin{aligned} V_h &= \{v_h \in V : v_h|_T \in [\mathcal{P}^2]^2\}, \\ Q_h &= \{q_h \in Q : q_h|_T \in \mathcal{P}^0\}. \end{aligned}$$

The non-conforming bilinear form in primal variables evaluates to

$$A_h^\varepsilon(u, v) = (e(u), e(v))_0 + \varepsilon^{-1}(\overline{\operatorname{div} u}^h, \overline{\operatorname{div} v}^h)_0.$$

Because $a(.,.)$ is elliptic on V, the reduced form $A_h^\varepsilon(u,v)$ is elliptic as well, and the finite element always has a unique solution. The stability on X_h is verified by the construction of a Fortin operator adjusting edge bubbles (see [BF91], pp 211):

$$I_h^F = I_h^{F,1} + I_h^{F,2}(I - I_h^{F,1}).$$

The operator $I_h^{F,1}$ shall have full order of approximation, i.e.

$$h^{-1}\, \|(I - I_h^{F,1})u\|_0 + \|I_h^{F,1}\|_1 \preceq \|u\|_1.$$

Possible choices are Scott-Zhang operators with set $\sigma \subset T$, or only $\sigma \subset \partial T$. The other operator $I_h^{F,2} : V \to V_h$ is defined by

$$(I_h^{F,2} u)(x) = 0 \qquad \forall\, x \text{ vertex of } \mathcal{T}_h$$

and

$$\int_e I_h^{F,2} u \,\mathrm{ds} = \int_e u \,\mathrm{ds} \qquad \forall\, e \text{ edge of } \mathcal{T}_h.$$

It fulfills

$$\|I_h^{F,2} u\|_1 + h^{-1} \|I_h^{F,2} u\|_0 \preceq \|u\|_1 + h^{-1} \|u\|_0.$$

Combining both estimates, one obtains continuity of $I_h^{F,2}$ in energy. The construction is important to verify stability on subspaces. Other stable elements for problems of elasticity are found in [Ste88]. Robust and non-robust elements are found in [Fal91].

2.4.3 The Reissner Mindlin plate

We use the Reissner Mindlin plate model with the discretization of [CS98]. In the primal space $V = H_0^1(\Omega) \times [H_0^1]^2$ we search for the transverse displacement w and the vector of rotations β. The scaled problem is defined by

$$a^b(\beta,\eta) + \frac{k}{t^2}(\nabla w - \beta, \nabla v - \eta)_0 = (g,v)_0 \qquad \forall\,(v,\eta) \in V.$$

The bilinear form

$$a^b(\beta,\eta) := \frac{1}{6}\left[(e(\beta),e(\eta))_0 + \frac{\nu}{1-\nu}(\operatorname{div}\beta, \operatorname{div}\eta)_0\right]$$

is related to bending energy. The small parameter is the square of the plate thickness t. The transverse load is Gt^3g. Further, k is the shear correction factor, ν the Poisson ratio, G the shear modulus and $e(.)$ is the linear strain operator. We extend the linear functional $(g,.)_0$ onto V and set

$$f(v,\eta) = g(v) + \delta(\eta) \qquad \forall\,(v,\eta) \in V.$$

With the definitions

$$\begin{aligned} \Lambda &= \nabla w - \beta, \\ \varepsilon &= t^2/k \end{aligned}$$

the bilinear form $A^\varepsilon(.,.)$ has the structure of (2.34). The dual variable, the scaled shear force is

$$p = \varepsilon^{-1}\Lambda(w,\beta) = \frac{k}{t^2}(\nabla w - \beta).$$

The mixed form $B^\varepsilon(.,.)$ evaluates to

$$B^\varepsilon((w,\beta,p),(v,\eta,q)) = a^b(\beta,\eta) + (\nabla v - \eta, p)_0 + (\nabla w - \beta, q)_0 - \frac{t^2}{k}(p,q)_0.$$

In [CS98] results of [BR80] and [AF89] were extended to prove the following regularity theorem. Let $(g,\delta) \in L_2$. Then the solution (w,β,p) of the mixed problem

$$B^\varepsilon((w,\beta,p),(v,\eta,q)) = (g,v)_0 + (\delta,\eta)_0 \qquad \forall\,(v,\eta) \in V\ \forall\, p \in Q \tag{2.85}$$

has a representation

$$w = w_0 + w_r$$

such that

$$\|w_0\|_3 + t^{-1}\|w_r\|_2 + \|\beta_2\| + t\,\|p\|_1 + \|p\|_0 \preceq \|g\|_{-1} + t\,\|g\|_0 + \|\delta\|_0 \tag{2.86}$$

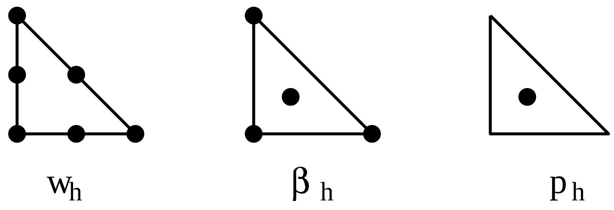


Figure 2.2: Finite elements for Reissner Mindlin

holds.We define the norms

$$\begin{aligned}\|(w,\beta)\|_{V^+}^2 = \|w\|_{V^+,w}^2 + \|\beta\|_{V^+,\beta}^2 &= \inf_{w=w_0+w_r}\left\{\|w_0\|_3^2 + t^{-2}\|w_r\|_2^2\right\} + \|\beta\|_2^2,\\ \|(w,\beta)\|_{V^-}^2 = \|w\|_{V^-,w}^2 + \|\beta\|_{V^-,\beta}^2 &= \inf_{w=w_0+w_r}\left\{\|w_0\|_1^2 + t^{-2}\|w_r\|_0^2\right\} + \|\beta\|_0^2,\\ \|p\|_{Q^+}^2 &= \|p\|_0^2 + t^2\|p\|_1^2.\end{aligned}$$

The dual norm to $\|.\|_{V^-}$ can be evaluated explicitely:

$$\begin{aligned}\|(g,\delta)\|_{(V^-)^*} &= \sup_{(v,\eta)\in V^-} \frac{g(v)+\delta(\eta)}{\left\{\inf\limits_{v=v_0+v_r}\|v_0\|_1^2 + t^{-2}\|v_r\|_0^2 + \|\beta\|_0^2\right\}^{1/2}}\\ &= \sup_{v_0,v_r,\beta} \frac{g(v_0)+g(v_r)+\delta(\eta)}{\{\|v_0\|_1^2 + t^{-2}\|v_r\|_0^2 + \|\beta\|_0^2\}^{1/2}}.\end{aligned}$$

The supremum is taken at the solution (w_0, w_r, β) of the variational problem

$$(w_0,v_0)_1 + t^{-2}(w_r,v_r)_0 + (\beta,\eta)_0 = g(v_0) + g(v_r) + \delta(\eta) \qquad \forall\, v_0, v_r, \eta$$

and is equal to $\{\|w_0\|_1^2 + t^{-2}\|w_r\|_0^2 + \|\beta\|_0^2\}^{1/2} = \{\|g\|_{-1}^2 + t^2\|g\|_0^2 + \|\delta\|_0^2\}^{1/2}$. Thus we computed

$$\|(g,\delta)\|_{(V^-)^*}^2 = \|g\|_{-1}^2 + t^2\|g\|_0^2 + \|\delta\|_0^2. \tag{2.87}$$

This is exactly the norm on the right hand side of (2.86). The estimate of both components of the splitting $w = w_0 + w_r$ give $(w,\beta) \in V^+$, and the estimate

$$\|(w,\beta)\|_{V^+} + \|p\|_{Q^+} \preceq \|(g,\delta)\|_{V^-}.$$

The discretization of [AB93] and [CS98] uses P_2 elements for the transverse displacement, P_1^+ for the rotations and P_0 for the shear, see Figure 2.2. The space P_1^+ are the piecewise linear functions enriched by the element bubbles $\lambda_1\lambda_2\lambda_3$ in barycentric coordinates. These spaces are non-nested due to the bubble functions. Following [CS98], the bilinear form $B^\varepsilon(.,.)$ is replaced by the stabilized form

$$\begin{aligned}\widetilde{B}((w,\beta,p),(v,\eta,q)) &= B^\varepsilon((w,\beta,p),(v,\eta,q)) \\ &+ \frac{\mu k}{t^2+h^2}\left(\nabla w-\beta-\frac{t^2}{k}p, \nabla v-\eta-\frac{t^2}{k}q\right)_0.\end{aligned}$$

In [CS98] a local mesh size is used. For only technical reasons, we restrict us to a global one. In [CS98] the stabilization parameter μ is assumed to be in $(0,1)$. The limit case $\mu=1$ can be allowed as well.

For the choice of piecewise constant elements for Q_h, the discrete shear operator can be computed locally and evaluates to

$$\Lambda_h(w_h,\beta_h)=\overline{(\nabla w_h-\beta_h)}^h,$$

where $\bar{\cdot}^h$ is the element-wise averaging operator. The reduction to the positive definite problem gives the form

$$\begin{aligned}A_h^\varepsilon((w_h,\beta_h),(v_h,\eta_h)) &= a^b(\beta_h,\eta_h)+\frac{k\mu}{h^2+t^2}(\nabla w_h-\beta_h,\nabla v_h-\eta_h)_0 \\ &+ \left(\frac{k}{t^2}-\frac{k\mu}{h^2+t^2}\right)(\overline{\nabla w_h-\beta_h}^h,\overline{\nabla v_h-\eta_h}^h)_0.\end{aligned} \tag{2.88}$$

By the stabilization trick, a part of the shear energy is assembled with full integration. In [AB93], a constant part of the shear energy is used for stabilization. By the mesh dependent stabilization, Chapelle and Stenberg could apply the duality trick. Also for multigrid methods, the mesh dependent versions seems to be advantageous, from the theoretical as well as practical point of view.

Chapelle and Stenberg use $\mu<1$. Thus, also for mesh sizes $h<t$ a reasonable part of the shear energy is reduced. By the choice of $\mu=1$ the non-conforming form converges to the original form $A^\varepsilon(.,.)$. Chapelle and Stenberg proved the discrete stability conditions to obtain optimal a priori bound. We will use the mixed machinery to obtain equivalent results. The choice $\mu=1$ is independent of technique.

We have to write the bilinear form (2.88) as

$$A_h^\varepsilon((w_h,\beta_h),(v_h,\delta_h))=a((w_h,\beta_h),(v_h,\delta_h))+\varepsilon^{-1}c(\Lambda_h(w_h,\beta_h),\Lambda_h(v_h,\delta_h)).$$

Thus, we are forced to set

$$\begin{aligned}a((w_h,\beta_h),(v_h,\delta_h)) &= a^b(\beta_h,\eta_h)+\frac{k\mu}{h^2+t^2}(\nabla w_h-\beta_h,\nabla v_h-\eta_h)_0 \\ c(p_h,q_h) &= \varepsilon\left(\frac{k}{t^2}-\frac{k\mu}{h^2+t^2}\right)(p_h,q_h)_0.\end{aligned}$$

We had set $\varepsilon=k/t^2$ already. Thus we obtain

$$\|p\|_c^2\simeq\left(1-\frac{\mu t^2}{h^2+t^2}\right)\|p\|_0^2.$$

Here is one difference between $\mu < 1$ and $\mu = 1$. For the first case, $\|.\|_c$ is equivalent to $\|.\|_0$ uniformly in h and t. For the second one, there holds

$$\|p\|_c^2 \simeq \min\{1, h^2/t^2\} \|p\|_0^2.$$

We define the norm $\|.\|_V$ such that $a(.,.)$ is elliptic and continuous, namely

$$\|(w,\beta)\|_V^2 := \|\beta\|_1^2 + \frac{1}{h^2+t^2} \|\nabla w - \beta\|_0^2. \tag{2.89}$$

The norm $\|.\|_Q$ is defined such that we obtain stability by definition, i.e.

$$\|p\|_Q^2 := \sup_{(w,\beta)\in V} \frac{c(\nabla w - \beta, p)^2}{\|(w,\beta)\|_V^2} + \varepsilon \, \|p\|_c^2.$$

The first part, the norm $\|.\|_{Q,0}$, can be estimated from above by

$$\begin{aligned} \|p\|_{Q,0} &\preceq \sup_{(w,\beta)\in V} \frac{c(\Lambda(w,\beta), q)}{\|(w,\beta)\|_V} \\ &= \sup_{w,\beta} \frac{\|\nabla w - \beta\|_c \, \|q\|_c}{\|\beta\|_1 + (h+t)^{-1} \|\nabla w - \beta\|_0} \\ &\preceq \sup_{w,\beta} \frac{\|\nabla w - \beta\|_0 \, \|q\|_c}{(h+t)^{-1} \|\nabla w - \beta\|_0} \\ &\preceq (h+t) \, \|q\|_c. \end{aligned} \tag{2.90}$$

We use the (relaxed) criterion of Fortin to prove stability of $B^\varepsilon(.,.)$ on on the finite element space X_h. We split

$$\Omega = \Omega_{h\leq t} \cup \Omega_{h>t}$$

such that

$$\Omega_{h\leq t} = \bigcup_{T\in\mathcal{T}:h_T<t} T \qquad \Omega_{h>t} = \Omega - \Omega_{h\leq t}.$$

Accordingly, we split

$$Q_h = Q_{h\leq t} + Q_{h>t}$$

such that

$$\begin{aligned} Q_{h\leq t} &= \{q_h \in Q_h : q_h = 0 \;\text{ on }\; \Omega_{h>t}\} \\ Q_{h>t} &= \{q_h \in Q_h : q_h = 0 \;\text{ on }\; \Omega_{h\leq t}\}. \end{aligned}$$

This splitting is defined for general shape regular triangulations. We restrict ourself to quasi-uniform meshes for a shorter notation, only. The space $Q_{h\leq t}$ belongs to elements, where the locking effect already disappeared. For $q_0 \in Q_{h\leq t}$ we have due to (2.90) the bound

$$\|q_0\|_Q \simeq \|q_0\|_{Q,0} + \varepsilon^{1/2} \, \|q_0\|_c \preceq (h+t) \, \|q_0\|_c + \varepsilon^{1/2} \|q_0\|_c \preceq \varepsilon^{1/2} \|q_0\|_c.$$

This is condition (2.67) of Theorem 2.12. We have to construct the Fortin operator such that $I_h^F : V \to V_h$ is continuous, and $(w_h, \beta_h) = I_h^F(w, \beta)$ fulfills

$$c(\nabla w_h - \beta_h, q_1) = c(\nabla w - \beta, q_1) \qquad \forall\, q_1 \in Q_{h>t}.$$

The space $Q_{h>t}$ consists of piecewise constant functions on elements with $h > t$. We build the operator by adjusting the bubbles of $\beta_h = (\beta_h^{(1)}, \beta_h^{(2)})$. Let b_T be the element bubble function on the triangle T. Then we define

$$\beta_h^{(i)} = \sum_{\substack{T\in\mathcal{T}\\ h_T>t}} \frac{(\partial_i w - \beta^{(i)}, 1)_T}{(b_T, 1)_T}\, b_T.$$

This local projection is L_2 stable

$$\|\beta_h\|_{0,\Omega_{h>t}} \preceq \|\nabla w - \beta\|_{0,\Omega_{h>t}}.$$

The function β_h depends on $(w, \beta)|_{\Omega_{h>t}}$ only, thus there holds

$$\begin{aligned}
\|(0,\beta_h)\|_V &\preceq \|\beta_h\|_1 + (h+t)^{-1}\|\beta_h\|_0 \\
&\preceq h^{-1}\|\beta_h\|_0 \\
&\preceq h^{-1}\|\nabla w - \beta\|_{0,\Omega_{h>t}} \\
&\preceq (h+t)^{-1}\|\nabla w - \beta\|_0 \preceq \|(w,\beta)\|_V.
\end{aligned}$$

By Theorem 2.9 we obtain stability of $B^\varepsilon(.,.)$ on X_h, and thus the best approximation property

$$\|(w,\beta) - (w_h,\beta_h)\|_V + \|p - p_h\|_Q \preceq \inf_{(v_h,\eta_h,q_h)\in X_h} \|(w,\beta) - (v_h,\eta_h)\|_V + \|p - q_h\|_Q$$

holds. In (2.90) we have estimated the $\|.\|_{Q,0}$ norm from above. For a finite element function $q_h \in Q_h$ we can estimate it from below. Therefore, let $\beta_h \in V_{h,b}$ in the space of element bubbles such that

$$\int_T \beta_h \,\mathrm{dx} = \int_T q_h \,\mathrm{dx} \qquad \forall\, T \in \mathcal{T}.$$

Then there holds

$$\|q_h\|_{Q,0} = \sup_{(w,\beta)\in V} \frac{c(\nabla w - \beta, q_h)}{\|\beta\|_1 + (h+t)^{-1}\|\nabla w - \beta\|_0} \geq \frac{c(\beta_h, q_h)_c}{h^{-1}\|\beta_h\|_0} \succeq h\frac{\|q_h\|_c^2}{\|q_h\|_0}. \qquad (2.91)$$

If one combines (2.90) and (2.91) one gets sharp bounds between $\|.\|_Q$ and $\|.\|_0$. One has to distinguish the cases $\mu < 1$ and $\mu = 1$, as well as $h < t$ or $h \geq t$. We collect the results in the table below. The estimates take over to locally refined meshes.

	$\mu < 1$		$\mu = 1$	
	$h \geq t$	$h < t$	$h \geq t$	$h < t$
$\|q_h\|_c/\|q_h\|_0$	1	1	1	h/t
$\|q_h\|_{Q,0}/\|q_h\|_0$	h	$\leq t$	h	$\leq h$
$\|q_h\|_Q/\|q_h\|_0$	h	t	h	h

Next, we will investigate approximation estimates and inverse estimates between the spaces V^-, V, and V^+, and between Q and Q^+. Let

$$I_h^V = (I_h^{V,w}, I_h^{V,\beta})$$

be Scott-Zhang type projection operators with the order of approximation

$$\begin{aligned} \|w - I_h^{V,w} w\|_k &\preceq h^{l-k} \|w\|_l \qquad 0 \leq k \preceq 1, \quad k \leq l \leq 3, \\ \|\beta - I_h^{V,\beta} \beta\|_k &\leq h^{l-k} \|\beta\|_l \qquad 0 \leq k \leq 1, \quad k \leq l \leq 2. \end{aligned}$$

The operators $I_h^{V,w}$ and $I_h^{V,\beta}$ have to preserve quadratic and linear functions, respectively. The operator I_h^Q performs element wise averaging and fulfills

$$\|q - I_h^Q q\|_k \preceq h^{l-k} \|q\|_l \qquad 0 \leq k \leq l \leq 1.$$

We need a different characterization of the norm $\|.\|_{V^{-,w}}$.

Lemma 2.17. *Let $I_t^w : V^{-,w} \to V^w$ be a local regularization operator at length scale t. It shall fulfill*

$$\begin{aligned} t^{-1}\|w - I_t^w w\|_0 + \|I_t^w w\|_1 &\preceq \|w\|_1 \\ \|I_t^w w\|_0 + t\,\|I_t^w w\|_1 &\preceq \|w\|_0. \end{aligned}$$

Then the following norms are equivalent:

$$\|w\|_{V^{-,w}} \simeq \|I_t^w w\|_1 + t^{-1}\,\|w - I_t^w w\|_0. \tag{2.92}$$

Proof. One estimate follows directly from the definition of the norm $\|.\|_{V^{-,w}}$, namely

$$\|w\|_{V^{-,w}} \simeq \inf_{w = w_0 + w_r} \left\{\|w_0\|_1 + t^{-1}\|w_r\|_0\right\} \preceq \|I_t^w w\|_1 + t^{-1}\|w - I_t^w w\|_0.$$

To estimate the other, take a splitting $w = w_0 + w_r$ such that

$$\|w_0\|_1 + t^{-1}\,\|w_r\|_0 \preceq \|w\|_{V^{-1,w}}.$$

Then we use the triangle inequality and the assumptions made for I_t^w to estimate

$$\begin{aligned} \|I_t^w w\|_1 + t^{-1}\,\|w - I_t^w w\|_0 &\leq \|I_t^w w_0\|_1 + t^{-1}\,\|w_0 - I_t^w w_0\|_0 \\ &\quad + \|I_t^w w_r\|_1 + t^{-1}\,\|w_r - I_t^w w_r\|_0 \\ &\preceq \|w_0\|_1 + t^{-1}\|w_r\|_0 \\ &\preceq \|w\|_{V^{-,w}}. \end{aligned}$$

□

For finite element functions, the splitting in the norm can be reduced to the finite element space:

Lemma 2.18. *On the finite element space V_h^w, the following norms are equivalent:*

$$\|w_h\|^2_{V^{-,w}} \simeq \inf_{\substack{w_h=w_{h,1}+w_{h,2}\\ w_{h,1},w_{h,2}\in V_h^w}} \left\{\|w_{h,1}\|_1^2 + t^{-2}\|w_{h,2}\|_0^2\right\}. \tag{2.93}$$

Proof. The right hind side trivially dominates the left hand side. Now, let I_h^w be a Scott-Zhang interpolation operator to V_h^w. It shall be continuous in $\|.\|_0$ as well as $\|.\|_1$-norm. Let I_t^w be a operator feasible for Lemma 2.17. Then the splitting of $w_h \in V_h$

$$w_{h,1} = I_h^w I_t^w w_h \qquad \text{and} \qquad w_{h,2} = I_h^w (I - I_t^w) w_h$$

is chosen to verify the other estimate:

$$\begin{aligned}
\inf_{\substack{w_h=w_{h,1}+w_{h,2}\\ w_{h,1},w_{h,2}\in V_h^w}} \left\{\|w_{h,1}\|_1^2 + t^{-2}\|w_{h,2}\|_0^2\right\} &\preceq \|I_h^w I_t^w w_h\|_1^2 + t^{-2}\,\|I_h^w(I - I_t^w)w_h\|_0^2 \\
&\preceq \|I_t^w w_h\|_1^2 + t^{-2}\,\|(I - I_t^w)w_h\|_0^2 \\
&\preceq \|w_h\|_{V^{-,w}}.
\end{aligned}$$

The last inequality followed by Lemma 2.17. □

Theorem 2.19. *The interpolation operators I_h^V and I_h^Q have full order of approximation, namely*

$$\begin{aligned}
h^{-1}\|(w,\beta) - I_h^V(w,\beta)\|_{V^-} + \|(w,\beta) - I_h^V(w,\beta)\|_V &\preceq h\,\|(w,\beta)\|_{V^+} && (2.94)\\
h^{-1}\|(w,\beta) - I_h^V(w,\beta)\|_{V^-} + \|I_h^V(w,\beta)\|_V &\preceq \|(w,\beta)\|_V && (2.95)\\
\|I_h^V(w,\beta)\|_{V^-} &\preceq \|(w,\beta)\|_{V^-}, && (2.96)
\end{aligned}$$

and

$$\|p - I_h^Q p\|_Q \preceq h\,\|p\|_{Q^+}. \tag{2.97}$$

Proof. We start with the V^- norm of (2.94). The estimate of the component β is the property of the interpolation operator $I_h^{V,\beta}$. Let $w = w_0 + w_r$ such that

$$\|w_0\|_3 + t^{-1}\,\|w_r\|_2 \le \|w\|_{V^{+,w}}$$

holds. Then

$$\begin{aligned}
\|w - I_h^{V,w} w\|_{V^{-,w}} &\simeq \inf_{w - I_h^{V,w} w = \tilde{w}_0 + \tilde{w}_r} \left\{\|\tilde{w}_0\|_1 + t^{-1}\|\tilde{w}_r\|_1\right\} \\
&\preceq \|w_0 - I_h^{V,w} w_0\|_1 + t^{-1}\|w_r - I_h^{V,w} w_r\|_0 \\
&\preceq h^2\,\|w_0\|_3 + h^2 t^{-1}\,\|w_r\|_2 \\
&\preceq h^2\,\|w\|_{V^{+,w}}.
\end{aligned}$$

The estimate of the V-norm of (2.94) can be simply split as

$$\begin{aligned}\|(w,\beta)-I_h^V(w,\beta)\|_V &\preceq \|\beta-I_h^{V,\beta}\beta\|_1+(h+t)^{-1}\|\nabla(w-I_h^{V,w}w)-(\beta-I_h^{V,\beta}\beta)\|_0\\ &\preceq h\,\|\beta\|_2+(h+t)^{-1}\|(w_0+w_r)-I_h^{V,w}(w_0+w_r)\|_1\\ &\preceq h\,\|\beta\|_2+(h+t)^{-1}h^2\,\|w_0\|_3+(h+t)^{-1}h\,\|w_r\|_2\\ &\preceq h\,\|\beta\|_2+h\,(\|w_0\|_3+t^{-1}\|w_r\|_2)\\ &\preceq h\,\|(w,\beta)\|_{V^+}.\end{aligned}$$

The estimate (2.96) is similar.

Next, we estimate continuity in the norm $\|.\|_V$. We have to apply the Bramble-Hilbert lemma. We define the semi-norm

$$|(w,\beta)|_V := (\|\nabla\beta\|_0^2+\|\nabla w-\beta\|_0)^{1/2}.$$

It has the kernel

$$V_{00}=\{(w,\beta)=(a+b^Tx,b): a\in\mathbb{R}, b\in\mathbb{R}^2\}.$$

On the reference element T^R, the following norms are equivalent:

$$\|(w,\beta)\|_1^2\simeq\|(w,\beta)\|_0^2+|(w,\beta)|_V.$$

The interpolation operator I_h^V is continuous in $\|.\|_1$, and it preserves the kernel V_{00}. By Theorem 2.4, we conclude that the interpolation operator is continuous in the semi-norm, i. e.

$$\|\nabla(\beta-I_h^{V,\beta}\beta)\|_0+\|\nabla(w-I_h^{V,w}w)-(\beta-I_h^{V,\beta}\beta)\|_0\preceq\|\nabla\beta\|_0+\|\nabla w-\beta\|_0.$$

Let $x(\xi)$ be the mapping from the reference element to the element. The functions (w,β) are transformed differently, namely

$$\begin{aligned}w(x(\xi)) &= w^R(\xi)\\ \beta(x(\xi)) &= (\nabla x(\xi))\,\beta(x(\xi)).\end{aligned}$$

Using transformation rules, and $\|\nabla x\|\simeq h$, we get

$$\begin{aligned}&\|\nabla(\beta-I_h^{V,\beta}\beta)\|_T+h^{-1}\|\nabla(w-I_h^{V,w}w)-(\beta-I_h^{V,\beta}\beta)\|_0\\ &\quad\preceq h^{-1}\left(\|\nabla(\beta^R-I_h^{V,\beta}\beta^R)\|_0+\|\nabla(w^R-I_h^{V,w}w^R)-(\beta^R-I_h^{V,\beta}\beta^R)\|_0\right)\\ &\quad\preceq h^{-1}\left(\|\nabla\beta^R\|_0+\|\nabla w^R-\beta^R\|_0\right)\\ &\quad\preceq \|\nabla\beta\|_0+h^{-1}\|\nabla w-\beta\|_0.\end{aligned}$$

We obtained that the interpolation operator is continuous with respect to the semi-norm

$$\|\nabla\beta\|_0^2+h^{-2}\,\|\nabla w-\beta\|_0^2.$$

Because it is also continuous with respect to the semi norm

$$\|\nabla\beta\|_0,$$

the interpolation operator is continuous with respect to the family of semi-norms

$$\|\nabla\beta\|_0^2 + \alpha\,\|\nabla w - \beta\|_0^2$$

with

$$\alpha \preceq h^{-2}.$$

The choice $\alpha = (h^2+t^2)^{-1}$ of Chapelle and Stenberg is in the allowed range. For $h > t$, it is the maximal choice such that the interpolation operator is continuous.

Next, we estimate the better approximation in the norm $\|.\|_{V^-}$. We use Lemma 2.17 to estimate the norm $\|.\|_{V^{-,w}}$. Then, we distinguish between $\Omega_{h>t}$ and $\Omega_{h\preceq t}$. Using properties of the regularization operator we construct an splitting for the norm $\|.\|_{V^{-,w}}$ as

$$\begin{aligned}\|w - I_h^{V,w} w\|_{V^{-,w}} &\preceq \|I_t^w(w - I_h^{V,w} w)\|_1 + t^{-1}\|(I - I_t^w)(w - I_h^{V,w} w)\|_0 \\ &\preceq \|w - I_h^{V,w} w\|_{1,\Omega_{h>t}} + t^{-1}\|(w - I_h^{V,w} w)\|_{0,\Omega_{h\leq t}}.\end{aligned}$$

We use the approximation properties of the operator $I_h^{V,w}$, and the relations between h and t to continue

$$\begin{aligned}\|w - I_h^{V,w} w\|_{V^{-,w}} &\preceq \|w - I_h^{V,w} w\|_{1,\Omega_{h>t}} + ht^{-1}\|(w - I_h^{V,w} w)\|_{1,\Omega_{h\leq t}} \\ &\preceq \frac{h}{h+t}\|(w - I_h^{V,w} w)\|_1.\end{aligned}$$

We insert the rotations $\beta - I_h^{V,\beta}\beta$ and obtain the result

$$\begin{aligned}\|w - I_h^{V,w} w\|_{V^{-,w}} &\preceq h\frac{1}{h+t}\left(\|\nabla(w - I_h^{V,w} w) - (\beta - I_h^{V,\beta}\beta)\|_0 + \|\beta - I_h^{V,\beta}\beta\|_0\right) \\ &\preceq h\left(\frac{1}{h+t}\|\nabla(w - I_h^{V,w} w) - (\beta - I_h^{V,\beta}\beta)\|_0 + \|\beta - I_h^{V,\beta}\beta\|_1\right) \\ &\simeq h\,\|(w,\beta) - I_h^V(w,\beta)\|_V \\ &\preceq h\,\|(w,\beta)\|_V + h\,\|I_h^V(w,\beta)\|_V \\ &\preceq h\,\|(w,\beta)\|_V.\end{aligned}$$

The estimate the dual variable, we use the estimation of Q by a scaled L_2 norm, and separate the domain into $\Omega_{h>t}$ and $\Omega_{h\leq t}$

$$\begin{aligned}\|q - I_h^Q q\|_Q &\preceq (h+t)\|q - I_h^Q q\|_0 \\ &\preceq (h+t)\,\|q - I_h^Q q\|_{0,\Omega_{h<t}} + (h+t)\,\|q - I_h^Q q\|_{0,\Omega_{h\geq t}} \\ &\preceq th\,\|q\|_{1,\Omega_{h<t}} + h\,\|q\|_{0,\Omega_{h\geq t}} \\ &\preceq h\,(t\,\|q\|_1 + \|q\|_0) \\ &\simeq h\,\|q\|_{Q^+}.\end{aligned}$$

We have proved the theorem. □

Multigrid theory is based on the interplay between approximation estimates and inverse estimates. We have to establish the inverse estimate

Theorem 2.20. *There holds the inverse estimate*

$$\|(w_h, \beta_h)\|_V \preceq h^{-1} \|(w_h, \beta_h)\|_{V^-} \qquad \forall\, (w_h, \beta_h) \in V_h. \tag{2.98}$$

Proof. We estimate $\|(w_h, \beta_h)\|_V$ by splitting into components

$$\begin{aligned} \|(w_h, \beta_h)\|_V &\preceq \|\beta_h\|_0 + (h+t)^{-1}\|w_h\|_1 + (h+t)^{-1}\|\beta_h\|_0 \\ &\preceq h^{-1}\|\beta_h\|_0 + (h+t)^{-1}\|w_h\|_1. \end{aligned}$$

The term $\|\beta_h\|_0$ is contained in the norm $\|.\|_{V^-}$. We estimate $\|w_h\|_1$ by using a splitting $w_h = w_0 + w_r$ such that

$$\|w_0\|_1 + t^{-1}\|w_r\|_0 \preceq \|w_h\|_{V^{-,w}}.$$

Let $I_h^{V,w}$ be a projection operator of Scott Zhang type. Then we estimate

$$\begin{aligned} (h+t)^{-1}\|w_h\|_1 &= (h+t)^{-1}\|I_h^{V,w}(w_0 + w_r)\|_1 \\ &\preceq (h+t)^{-1} \left(\|I_h^{V,w} w_0\|_1 + \|I_h^{V,w} w_r\|_1\right) \\ &\preceq h^{-1}\|w_0\|_1 + (h+t)^{-1}h^{-1}\|I_h^{V,w} w_r\|_0. \end{aligned}$$

Since $(h+t)^{-1}h^{-1} \leq h^{-1}t^{-1}$, we finish with

$$\begin{aligned} (h+t)^{-1}\|w_h\|_1 &\preceq h^{-1}\|w_0\|_1 + h^{-1}t^{-1}\|w_r\|_0 \\ &\preceq h^{-1}\left(\|w_0\|_1 + t^{-1}\|w_r\|_0\right) \\ &\preceq h^{-1}\|w_h\|_{V^{-,w}}. \end{aligned}$$

□

Chapter 3

Iterative Methods

In the last chapter we came up with the finite dimensional variational problem: Find $u_h \in V_h$ such that

$$A_h(u_h, v_h) = f_h(v_h) \qquad \forall\, v_h \in V_h, \tag{3.1}$$

where V_h is a finite element space of dimension N, $A_h(.,.)$ is a symmetric and elliptic bilinear form on V_h, and $f_h(.)$ is a linear form on V_h. By the choice of a basis $(\varphi_i)_{i=1}^N \in [V_h]^N$ for the finite element space, we can represent the finite element function $u_h \in V_h$ by the vector $\underline{u} = (u_i)_{i=1}^N \in \mathbb{R}^N$ via

$$u_h = \sum_{i=1}^{N} u_i \varphi_i. \tag{3.2}$$

Usually, the nodal basis is chosen. We define the symmetric and positive definite (spd) system matrix $\underline{A} \in \mathbb{R}^{N\times N}$ as

$$\underline{A} := (A_h(\varphi_i, \varphi_j))_{i,j=1}^N, \tag{3.3}$$

and the vector on the right hand side as

$$\underline{f} := (f_h(\varphi_i))_{i=1}^N. \tag{3.4}$$

The variational problem (3.1) is equivalent to the linear system of equations

$$\underline{A}\,\underline{u} = \underline{f}. \tag{3.5}$$

Since the dimension N of the system may be very large, solvers of optimal arithmetic complexity and memory complexity are required. Optimal memory complexity is achieved by most iterative methods like conjugate gradients (cg) iteration ([HS52], for a wide overview see [Hac91]). The iteration number and the time complexity of the iterative solver depend on the applied preconditioner. A spd matrix $\underline{C}$ is called preconditioner, if the operation

$$\underline{w} := \underline{C}^{-1} \times \underline{d} \tag{3.6}$$

is computable. The performance of the preconditioner $\underline{C}$ for the solution of (3.5) depends on the arithmetic operations needed to compute (3.6), and on the constants of the spectral inequalities

$$c_1\,\underline{C} \le \underline{A} \le c_2\,\underline{C}. \tag{3.7}$$

The relation $A_1 \leq A_2$ defined for two positive definite matrices is equivalent to $A_2 - A_1$ is positive semi-definite. It is well known, that an upper bound for necessary cg iterations for a fixed error reduction behaves like $O(\sqrt{c_2/c_1})$. The optimal time complexity for the approximative solution of (3.5) with fixed accuracy is $O(N)$. This is possible with constants c_1 and c_2 independent of N. Multigrid methods are optimal methods (see [Hac82], [Hac85], [Xu92], [Bra93]). We call a preconditioner robust with respect to the parameter ε, if there exists constants c_1 and c_2 independent of ε such that (3.7) holds.

In the next sections we will collect techniques to construct and analyze preconditioners for problems without parameters. In Section 4 we apply these methods to obtain robust preconditioners for parameter dependent problems.

3.1 Additive Schwarz techniques

The additive Schwarz (AS) framework provides us a simple and elegant technique for the construction and analysis of a class of preconditioners. A good introduction to the material is [SBG96].

For the analysis it is advantageously to stay in the fe space V_h instead of the Euclidean space $\mathbb{R}^N$. For sure, the obtained methods are the same, but the notation simplifies. By means of the inner product $(.,.)_h$ of the Hilbert space V_h we define the linear operator $A_h : V_h \rightarrow V_h$ by

$$(A_h u_h, v_h)_h = A_h(u_h, v_h) \qquad \forall\, u_h, v_h \in V_h.$$

Similarly, the vector $f_h \in V_h$ is defined by

$$(f_h, v_h)_h = f_h(v_h) \qquad \forall\, v_h \in V_h.$$

By means of A_h and f_h we can rewrite the variational problem (3.1) or the system of linear equations (3.5) equivalently as operator equation: Find $u_h \in V_h$ such that

$$A_h u_h = f_h. \tag{3.8}$$

A preconditioner $C_h : V_h \rightarrow V_h$ is a $(.,.)_h$ - self-adjoint and positive definite operator. The spectral estimates

$$c_1\, C_h \leq A_h \leq c_2\, C_h \tag{3.9}$$

correspond to (3.7). The relation $A_1 \leq A_2$ between two self-adjoint operators is defined by

$$(A_1 v_h, v_h)_h \leq (A_2 v_h, v_h)_h \quad \forall\, v_h \in V_h.$$

The idea of additive Schwarz preconditioning is to split one large problem into a set of smaller problems. Let

$$\{(V_i, (.,.)_i) : 1 \leq i \leq M\}$$

be a set of Hilbert spaces. Each of them is embedded by a lifting operator into the space V_h

$$R_i : V_i \rightarrow V_h. \tag{3.10}$$

The whole space V_h shall be decomposable into lifted local spaces, i.e.

$$V_h = \sum_{i=1}^{M} R_i V_i. \tag{3.11}$$

The sum is not necessarily a direct sum. On each one of the spaces V_i we require a symmetric, continuous and elliptic bilinear form

$$C_i(.,.) : V_i \times V_i \to \mathbb{R}.$$

As we will see below, it should be a local approximation of $A_h(.,.)$. We define the AS preconditioner $C_h : V_h \to V_h$ by the application of its inverse:

$$w_h = C_h^{-1} d_h,$$

which is computed by

$$w_h = \sum_{i=1}^{N} R_i w_i,$$

with $w_i \in V_i$ the unique solution of

$$C_i(w_i, v_i) = (d_h, R_i v_i)_h \qquad \forall\, v_i \in V_i.$$

It is easily seen that the operator $C_h^{-1} A_h : V_h \to V_h$ and the finite element function $C_h^{-1} f_h$ do not change, if the norms of V_h or V_i are replaced by equivalent ones.

By means of the inner product $(.,.)_i$ on V_i we define the linear operator $C_i : V_i \to V_i$:

$$(C_i u_i, v_i)_i = C_i(u_i, v_i) \qquad \forall\, u_i, v_i \in V_i.$$

The adjoint operator R_i^T of R_i is defined by

$$(R_i^T u_h, v_i)_i = (u_h, R_i v_i)_h \qquad \forall\, u_h \in V_h,\ \forall\, v_i \in V_i.$$

By these definitions we can rewrite the preconditioning operation C_h^{-1} in operator form as

$$C_h^{-1} = \sum_{i=1}^{M} R_i C_i^{-1} R_i^T. \tag{3.12}$$

Although we use the abstract notation in the Hilbert space, it is very close related to the implementation. We define the subspace solution operators

$$T_i := R_i C_i^{-1} R_i^T A_h. \tag{3.13}$$

A special choice of the subspace bilinear forms is the Galerkin setting

$$C_i(u_i, v_i) := A_i(u_i, v_i) := A(R_i u_i, R_i v_i).$$

The according definition of $A_i : V_i \to V_i$ evaluates to

$$A_i = R_i^T A_h R_i.$$

For the Galerkin choice the subspace solution operator

$$T_i = R_i A_i^{-1} R_i^T A_h. \tag{3.14}$$

is the $A_h(.,.)$-orthogonal projection to $R_i V_i$.

The following theorem provides the central technique for the analysis of AS preconditioners. It was formulated in different forms in many papers (see e.g.[MN85], [Lio88], [DW90], [Zha91], [BPWX91], [Xu92], [Nep92], [GO95]).

Theorem 3.1 (Additive Schwarz Lemma). *Let us define the splitting norm*

$$|||u_h|||^2 := \inf_{\substack{u_h = \sum R_i u_i \\ u_i \in V_i}} \sum_{i=1}^{M} C_i(u_i, u_i). \tag{3.15}$$

on V_h. It is equal to the norm $\|u_h\|_{C_h} := (C_h u_h, u_h)_h^{1/2}$ generated by the AS preconditioner, i.e. there holds

$$|||u_h||| = \|u_h\|_{C_h} \qquad \forall\, u_h \in V_h. \tag{3.16}$$

Proof. The constrained minimization problem (3.15) can be written as saddle point problem: Find $(u_1, \ldots, u_M) \in V_1 \times \ldots \times V_M$ and $\lambda_h \in V_h$ such that

$$\begin{array}{lcl} C_i(u_i, v_i) \quad + \quad (\lambda_h, R_i v_i)_h &=& 0 \qquad\qquad \forall\, v_i \in V_i,\ 1 \le i \le M, \\ (\sum R_i u_i, \mu_h)_h &=& (u_h, \mu_h)_h \qquad \forall\, \mu_h \in V_h. \end{array}$$

Eliminating the u_i by $u_i = -C_i^{-1} R_i^T \lambda_h$, we can rewrite the second row as

$$(\sum R_i C_i^{-1} R_i^T \lambda_h, \mu_h)_h = -(u_h, \mu_h)_h \qquad \forall\, \mu_h \in V_h.$$

By (3.12) we can replace the sum by C_h^{-1} and obtain

$$C_h^{-1} \lambda_h = -u_h.$$

The minimum of the constrained minimization problem (3.15) is taken at the solution (u_i) of the saddle point system and evaluates to

$$\begin{aligned} |||u_h|||^2 &= \sum (C_i u_i, u_i)_i = \sum (C_i C_i^{-1} R_i^T \lambda_h, C_i^{-1} R_i^T \lambda_h)_i = \sum (R_i C_i^{-1} R_i^T \lambda_h, \lambda_h)_h \\ &= (C_h^{-1} \lambda_h, \lambda_h)_h = (C_h^{-1} C_h u_h, C_h u_h)_h = (C_h u_h, u_h)_h = \|u_h\|_{C_h}^2, \end{aligned}$$

and the proof is complete. □

This theorem provides us with a constructive possibility to calculate or bound the constants of the spectral inequalities (3.9).

If only a finite number of spaces overlap, the right inequality of (3.9) is trivial We define the symmetric matrix $(g_{i,j})_{i,j=1}^M$ of overlapping subspaces by

$$g_{i,j} := \begin{cases} 1 & \text{if } \exists\, u_i \in V_i,\ v_j \in V_j :\ A_h(R_i u_i, R_j v_j) \neq 0 \\ 0 & \text{else} \end{cases}$$

Then we define the overlap of the space splitting as

$$\text{overlap}(\{V_i\}) := \max_i \sum_{j=1}^M g_{i,j}$$

Lemma 3.2 (Finite Overlap). *Let $c_c > 0$ be a constant such that*

$$\|R_i u_i\|_{A_h}^2 \leq c_c \, \|u_i\|_{C_i}^2 \qquad \forall\, u_i \in V_i,\ 1 \leq i \leq M \tag{3.17}$$

holds. Then the estimate

$$A_h \leq c_2 \, C_h$$

is fulfilled with the constant

$$c_2 := c_c \, overlap(\{V_i\}).$$

Proof. Let $u_h = \sum R_i u_i,\ u_i \in V_i$ be an arbitrary splitting. There holds

$$\|u_h\|_{A_h}^2 = A_h(u_h, u_h) = \sum_{i,j} A_h(R_i u_i, R_j u_j).$$

Using the symmetric matrix of overlapping subspaces and Young's inequality we get

$$\begin{aligned} \|u_h\|_{A_h}^2 &= \sum_{i,j} g_{i,j}\, A_h(R_i u_i, R_j u_j) \\ &\leq \sum_{i,j} g_{i,j}\, \frac{1}{2}(\|R_i u_i\|_{A_h}^2 + \|R_j u_j\|_{A_h}^2) \\ &= \sum_{i,j} g_{i,j} \|R_i u_i\|_{A_h}^2. \end{aligned}$$

Now we use the definition of overlap($\{V_i\}$), and (3.17) to get

$$\begin{aligned} \|u_h\|_{A_h}^2 &\leq \text{overlap}(\{V_i\}) \sum_i \|R_i u_i\|_{A_h}^2 \\ &\leq c_c \, \text{overlap}(\{V_i\}) \sum_i \|u_i\|_{C_i}^2. \end{aligned}$$

Because this estimate holds for any splitting, it holds also for the minimizer of the splitting norm

$$\begin{aligned}\|u_h\|^2_{A_h} &\leq c_c \operatorname{overlap}(\{V_i\}) \inf_{u_h=\sum R_i u_i} \sum_i \|u_i\|^2_{C_i} \\ &= c_c \operatorname{overlap}(\{V_i\}) \, |\!|\!| u_h |\!|\!|^2 \\ &= c_c \operatorname{overlap}(\{V_i\}) \, \|u_h\|^2_{C_h}.\end{aligned}$$

□

By the paper of [BPWX91], the multiplicative version of Schwarz methods is understood as well. Let us assume that the subspace problems are scaled such that

$$C_i \geq A_i. \tag{3.18}$$

Let

$$S_{mu} := (I - T_M)(I - T_{M-1}) \dots (I - T_2)(I - T_1) \tag{3.19}$$

be one step of the multiplicative Schwarz iteration. The adjoint with respect to the inner product $A_h(.,.)$ is

$$S^*_{mu} = (I - T_1)(I - T_2) \dots (I - T_{M-1})(I - T_M). \tag{3.20}$$

The symmetric version of multiplicative Schwarz is

$$S_{smu} := S^*_{mu} S_{mu}. \tag{3.21}$$

It defines the preconditioner C_{smu} by

$$S_{smu} = I - C^{-1}_{smu} A_h.$$

Usually, the multiplicative method leads to a faster convergent method then the additive version. But in general, one can prove only that the multiplicative version is not much worse:

Theorem 3.3 (Multiplicative Schwarz). *Let C_{add} and C_{smu} be the preconditioners obtained by the additive Schwarz method and the symmetric multiplicative Schwarz method, respectively. Assume the scaling 3.18 holds. Then the spectral estimates*

$$A_h \leq C_{smu} \leq \mathit{overlap}(\{V_i\})^2 \, C_{add}. \tag{3.22}$$

are fulfilled.

The proof is similar to [BPWX91] and [BP92], Chapter 5. Actually, estimate (3.18) can be relaxed to allow over-relaxation with parameter $\omega < 2$.

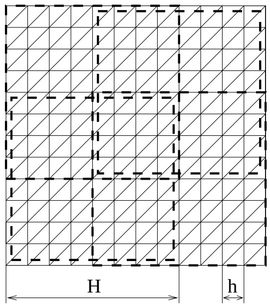


Figure 3.1: Local domain decomposition

3.1.1 The local domain decomposition preconditioner

First, we consider a domain decomposition (dd) method with local spaces, only. Let $V = H^m(\Omega)$, and let $A(.,.)$ be a symmetric, elliptic and continuous bilinear form on V. Let $V_h \subset V$ be a finite element subspace of mesh size h, and let $A_h(u_h, v_h) = A(u, v)$.

We will formulate and analyze the local dd preconditioner. We decompose the domain into an overlapping set of subdomains Ω_i of diameter $O(H)$ and overlap $O(H)$, see Figure 3.1. We assume that the number of overlapping subdomains is finite. Let $\omega_i \subset \Omega_i$ such that

$$\operatorname{dist}(\omega_i, \partial\Omega_i \setminus \partial\Omega) \succeq h, \tag{3.23}$$

and the construction of a partitioning of unity $\{\psi_i\}$, $\psi_i \in C^\infty(\Omega)$ with the following properties is possible:

$$\operatorname{supp} \psi_i \subset \omega_i,$$

and

$$\|\nabla^k \Psi\|_{L_\infty} \preceq H^{-k} \qquad 0 \leq k \leq m.$$

We define the local spaces

$$V_i = \{v_h \in V_h : v_h = 0 \text{ in } \Omega \setminus \Omega_i\}.$$

The lifting operators $R_i : V_i \to V_h$ are the trivial embedding operators. The inner products $(.,.)_i$ are inherited from V_h. We consider the space splitting

$$V_h = \sum R_i V_i.$$

The subspace forms are defined by the Galerkin approach

$$C_i(u_i, v_i) = A_i(u_i, v_i) = A_h(R_i u_i, R_i v_i).$$

Thus we get

$$C_i = A_i = R_i^T A_h R_i.$$

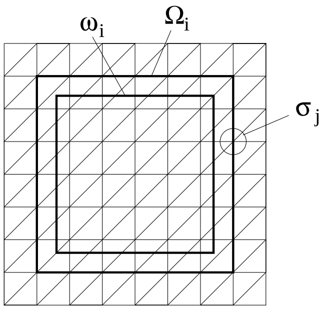


Figure 3.2: Relation of Ω_i, ω_i and σ_j

We define the overlapping domain decomposition preconditioner D_h without coarsegrid correction

$$D_h^{-1} := \sum R_i A_i^{-1} R_i^T. \tag{3.24}$$

We call this preconditioner D_h in analogy to **d**iagonal. It will be needed as a part of more complicated preconditioners, later.

Theorem 3.4. *The overlapping domain decomposition preconditioner D_h without coarse grid correction fulfills the spectral inequalities*

$$H^{2m} D_h \preceq A_h \preceq D_h. \tag{3.25}$$

Proof. The upper bound follows from the finite overlap of subspaces and Lemma 3.2. The lower bound is verified by the construction of an explicit decomposition $u_h = \sum u_i$, $u_i \in V_i$ fulfilling

$$\sum \|u_i\|_{A_h}^2 \preceq H^{-2m} \|u_h\|_{A_h}^2.$$

Then we can conclude by the Additive Schwarz lemma

$$H^{2m} \|u_h\|_{D_h}^2 = H^{2m} \|\!|u_h|\!\|^2 \leq H^{2m} \sum \|u_i\|_{A_h}^2 \preceq \|u_h\|_{A_h}^2.$$

Let $I_h : V \to V_h$ be a Scott-Zhang projection operator of the form

$$I_h v = \sum_{j=1}^{N} l_j(\Pi_j v) \varphi_j.$$

Each functional $l_j(\Pi_j .)$ is defined on $L_2(\sigma_j)$, where σ_j is a set of diameter $O(h)$.

Because of (3.23) we can chose σ_j such that there holds

$$I_h : H_0^m(\omega_i) \to V_i,$$

see Figure 3.2. We define

$$u_i = I_h(\psi_i u_h).$$

By using the linearity of I_h, the partitioning of unity and the projection property of I_h we verify

$$\sum_{i=1}^{M} u_i = \sum_{i=1}^{M} I_h(\psi_i u_h) = I_h u_h = u_h.$$

There holds $u_i \in V_i$ because $(\psi_i u_h) \in H_0^m(\omega_i)$. Thus, we constructed a splitting of u_h. Using the continuity of $A_h(.,.)$, the continuity of the Scott-Zhang projector I_h, and properties of the partition of unity (Lemma 2.6), we get

$$\begin{aligned} \|u_i\|_{A_h}^2 &\preceq \|u_i\|_{m,\Omega}^2 = \|I_h(\psi_i u_h)\|_{m,\Omega}^2 \\ &\preceq \|\psi_i u_h\|_{m,\Omega}^2 \\ &\preceq H^{-2m}\|u_h\|_{0,\omega_i}^2 + \|\nabla^m u_h\|_{0,\omega_i}^2. \end{aligned}$$

Using the finite overlap of subdomains and the ellipticity of $A_h(.,.)$, we get

$$\begin{aligned} \sum_{i=1}^{M} \|u_i\|_{A_h}^2 &\preceq \sum_{i=1}^{M} [H^{-2m}\|u_h\|_{0,\omega_i}^2 + \|\nabla^m u_h\|_{0,\omega_i}^2] \\ &\preceq H^{-2m}\|u_h\|_{0,\Omega}^2 + \|\nabla^m u_h\|_{0,\Omega}^2 \\ &\preceq H^{-2m}\|u_h\|_{m,\Omega}^2 \preceq H^{-2m}\|u_h\|_{A_h}^2, \end{aligned} \tag{3.26}$$

and the proof is complete. □

3.1.2 Domain decomposition preconditioner with coarse grid system

In this section we will consider the method of Dryja and Widlund [DW89]. A pure local space decomposition cannot give an preconditioner of optimal spectral bounds. One has to add a coarse grid system in order to avoid the factor H^{2m}. Let V_H be a finite element subspace of V of mesh-size H. It is called the coarse grid space. It is not necessarily a subspace of V_h.

Let I_H be an interpolation operator with the optimal approximation estimates

$$|v - I_H v|_{m-k} \preceq H^k \|v\|_m \qquad \forall\, v \in V,\ 0 \le k \le m. \tag{3.27}$$

Let $R_H : V_H \to V_h$ be the so called prolongation operator. If there holds $V_H \subset V_h$ we can use natural embedding. In general, it shall fulfill the approximation estimates

$$|v_H - R_H v_H|_{m-k} \preceq H^k \|v_H\|_m \qquad \forall\, v_H \in V_H, \quad 0 \le k \le m. \tag{3.28}$$

We use the space splitting

$$V_h = R_H V_H + \sum_{i=1}^{M} R_i V_i. \tag{3.29}$$

The spaces V_i are local spaces with natural embedding, and thus we can chose the Galerkin approach $A_i = R_i^T A_h R_i$. On the space V_H we define

$$C_H(u_H, v_H) = A_H(u_H, v_H) = A(u_H, v_H).$$

If $V_H \not\subset V_h$, the method differs from the Galerkin approach, because in general

$$A_H(u_H, v_H) \neq A_h(R_h u_H, R_h v_H).$$

We define the overlapping dd preconditioner with coarse grid system as

$$C_h^{-1} = R_H A_H^{-1} R_H^T + \sum R_i A_i^{-1} R_i^T. \tag{3.30}$$

By using the local dd preconditioner D_h of (3.24) it can be expressed by

$$C_h^{-1} = R_H A_H^{-1} R_H^T + D_h^{-1}.$$

Lemma 3.5 (Optimal two level preconditioner). *Assume the following is true:*

i. The overlap of local spaces is bounded by N_O.

ii. The prolongation is continuous, i.e.

$$\|R_H u_H\|_{A_h} \leq c_R \|u_H\|_{A_H} \qquad \forall\, u_H \in V_H. \tag{3.31}$$

iii. There exists an continuous interpolation operator $I_H : V_h \to V_H$, i.e.

$$\|I_H u_h\|_{A_H} \leq c_I \|u_h\|_{A_h} \qquad \forall\, u_h \in V_h. \tag{3.32}$$

iv. The local splitting of the difference $u_f := u_h - R_H I_H u_h$ is stable, i.e.

$$\inf_{\substack{u_f = \sum u_i \\ u_i \in V_i}} \sum \|u_i\|_{A_h}^2 \leq c_L \|u_h\|_{A_h}^2. \tag{3.33}$$

Then the two level preconditioner (3.30) fulfills the optimal spectral bounds

$$c_1\, C_h \leq A_h \leq c_2\, C_h \tag{3.34}$$

with

$$\begin{aligned} c_1 &:= (c_I^2 + c_L)^{-1}, \\ c_2 &:= (1 + N_O)\, \max\{c_R^2, 1\}. \end{aligned}$$

Assumptions 2-4 are also necessary to obtain an optimal preconditioner.

Proof. We apply Lemma 3.2 to estimate $A_h \leq c_2\, C_h$. The overlap of spaces is bounded by $(1 + N_O)$. Due to (3.31), and the Galerkin choice for A_i, assumption (3.17) is fulfilled with

$$c_c := \max\{c_R^2, 1\},$$

and the upper bound is proven.

On the other hand, assumptions (3) and (4) are formulated such that

$$\begin{aligned}\|u_h\|_{C_h}^2 = |||u_h|||^2 &= \inf_{u_h = R_H u_H + \sum u_i} \left\{ \|u_H\|_{A_H}^2 + \sum \|u_i\|_{A_h}^2 \right\} \\ &\leq \|I_H u_h\|_{A_H}^2 + \inf_{u_f = \sum u_i} \sum \|u_i\|_{A_h}^2 \\ &\leq (c_I^2 + c_L)\|u_h\|_{A_h}^2.\end{aligned}$$

Now, assume that C_h is a preconditioner with optimal bounds, i.e

$$\|u_h\| \simeq |||u_h|||.$$

We will derive conditions (2) - (4). We start with (2). Let $u_H \in V_H$, and set $u_h = R_H u_H$. Because C_h is an optimal preconditioner, there holds

$$\|u_h\|_{A_h}^2 \preceq |||u_h|||^2 = \inf_{R_H u_H = R_H v_H + \sum v_i} \left\{ \|v_H\|_{A_H}^2 + \sum_i \|v_i\|^2 \right\} \leq \|u_H\|_{A_H}^2.$$

We have chosen the splitting $v_H = u_H$ and $v_i = 0$. The continuous interpolation operator (3.32) and the stable splitting (3.33) is chosen as the minimizer of the splitting norm. □

Theorem 3.6. *The two level preconditioner C_h of (3.30) has optimal spectral bounds, i.e.*

$$C_h \preceq A_h \preceq C_h. \tag{3.35}$$

Proof. We apply the previous lemma. We have assumed finite overlap of subdomains. Assumptions (2) and (3) follow from (3.28) and (3.27), respectively, by the choice $k = 0$ and the triangle inequality.

We are left to verify (4). By using (3.28) and (3.27) we estimate

$$\begin{aligned}\|u_f\|_0 &= \|u_h - R_H I_H u_h\|_0 \\ &\leq \|u_h - I_H u_h\|_0 + \|I_H u_h - R_H I_H u_h\|_0 \\ &\preceq H^m\,\|u_h\|_m + H^m\,\|I_H u_h\|_m \\ &\preceq H^m\,\|u_h\|_m \end{aligned} \tag{3.36}$$

and

$$\|u_f\|_m \preceq \|u_h\|_m. \tag{3.37}$$

Proceeding similar to Theorem 3.4, we chose

$$u_i = I_H(\psi_i u_f).$$

According to the intermediate result (3.26) there holds

$$\sum_i \|u_i\|^2_{A_h} \preceq H^{-2m}\|u_f\|^2_0 + \|u_f\|^2_m.$$

But now, we have the improved L_2 estimate (3.36) and we can finish the proof with

$$\sum_i \|u_i\|^2_{A_h} \preceq \|u_h\|^2_m \preceq \|u_h\|^2_{A_h}.$$

□

The coarse grid mesh sizes H can be chosen such that the total work is minimized. We cannot obtain optimal complexity for a 2 level method. The dimension of the coarse grid system grows unless $H \simeq 1$, and the dimensions of the local systems grow unless $H \simeq h$.

We will analyze two level methods with $H \simeq h$ as preparation for the more involved multigrid analysis providing optimal preconditioners.

3.2 Multigrid Methods

Optimal preconditioners can be constructed by the use of variational problems on a sequence of finite-element spaces

$$V_1, V_2, \ldots, V_L = V_h$$

with according inner products $(.,.)_l$. We do not assume that the spaces are nested, so we need grid transfer operators

$$R_l : V_{l-1} \to V_l, \qquad 2 \leq l \leq L.$$

On each level, $1 \leq l \leq L$, we have the eventually modified bilinear form

$$A_l(.,.) : V_l \times V_l \to \mathbb{R}.$$

The operators $A_l : V_l \to V_l$ are defined by

$$(A_l u_l, v_l)_l = A_l(u_l, v_l) \qquad \forall\, u_l, v_l \in V_l.$$

The adjoint operators $R_l^T : V_l \to V_{l-1}$ of R_l are defined by

$$(R_l^T u_l, v_{l-1})_{l-1} = (u_l, R_l v_l) \qquad \forall\, u_l \in V_l,\ \forall\, v_{l-1} \in V_{l-1}.$$

On each level l we require a preconditioner $D_l : V_l \to V_l$, which is scaled such that

$$A_l \leq D_l \tag{3.38}$$

holds. It defines the smoothing iteration

$$S_l = I - D_l^{-1} A_l. \tag{3.39}$$

The preconditioner D_l might stem from a local additive Schwarz method on level l.

Procedure $MG(l, u_l, f_l)$
if $l = 1$

$$MG(l, u_l, f_l) = A_l^{-1} f_l$$

else

$u_l^{1,0} = u_l$
(* pre-smoothing *)
do $j = 1, \ldots, m_l$
$\quad u_l^{1,j} = u_l^{1,j-1} + D_l^{-1}(f_l - A_l u_l^{1,j-1})$
(* coarse grid corrections *)
$d_{l-1} = R_l^T (f_l - A_l u_l^{1,m_l})$
$u_{l-1}^{2,0} = 0$
do $j = 1, \ldots, q$
$\quad u_{l-1}^{2,j} = MG(l-1, u_{l-1}^{2,j-1}, d_{l-1})$
$u_l^{3,0} = u_l^{1,m_l} + R_l u_{l-1}^{2,q}$
(* post-smoothing *)
do $j = 1, \ldots, m_l$
$\quad u_l^{3,j} = u_l^{3,j-1} + D_l^{-1}(f_l - A_l u_l^{3,j-1})$
$MG(l, u_l, f_l) = u_l^{3,m_l}$

Algorithm 1: Multigrid method

3.2.1 Multigrid algorithm

One step of the multigrid iteration

$$\hat{u}_l := MG(l, u_l, f_l) \tag{3.40}$$

is defined in Algorithm 1. The parameter q defines the type of cycle. Usually, the V-cycle method with $q = 1$ or the W-cycle method with $q = 2$ is performed. The number of smoothing steps m_l on the level l may be fixed, or may depend on the level. For the variable V-cycle the number of smoothing steps increase geometrically on lower levels. The multigrid algorithm leads to the multigrid operator $M_l : V_l \to V_l$

$$M_l u_l := MG(l, u_l, 0) \tag{3.41}$$

and to the multigrid preconditioner $C_l : V_l \to V_l$ defined by

$$C_l^{-1} f_l := MG(l, 0, f_l), \tag{3.42}$$

(see [JLM+89]). The multigrid operator fulfills the recursion formula

$$\begin{aligned} M_1 &= 0, \\ M_l &= (S_l)^{m_l} \left(I - R_l (I - (M_{l-1})^q) A_{l-1}^{-1} R_l^T A_l \right) (S_l)^{m_l}. \end{aligned} \tag{3.43}$$

3.2.2 Multigrid analysis

The classical multigrid analysis by Hackbusch [Hac82] is based on the approximation property and the smoothing property. If these properties are available, one can prove optimal convergence of the W-cycle iteration (with sufficiently many smoothing steps), and of the variable V-cycle. Several more sophisticated techniques ([BH83], [MMB87], [BP93]) provide optimal convergence of the V-cycle, but, up to the knowledge of the author, only for the nested case. So, we will focus on the classical theory.

The interplay of the coarse grid correction step and of the smoother is measured by means of two different norms, the energy norm $\|.\|_{A_l}$ and the local norm

$$\|u_l\|_{l,\tilde{0}}. \tag{3.44}$$

For full regular, second order problems it is

$$\|u_l\|_{l,\tilde{0}} = h_l^{-1}\|u_l\|_0,$$

and therefore we will call it L_2 - like norm. For problems with partial regularity, one has to chose Sobolev norms of fractional order. For these two norms, one has to check the approximation property

$$\|u_l - R_l A_{l-1}^{-1} R_l^T A_l u_l\|_{l,\tilde{0}} \leq C_a \|u_l\|_{A_l} \tag{3.45}$$

and the smoothing property

$$\|(S_l)^m u\|_{A_l} \preceq \eta(m)\ \|u\|_{l,\tilde{0}}, \tag{3.46}$$

where $\eta(m)$ is a function with

$$\eta(m) \to 0 \quad \text{as} \quad m \to \infty.$$

One part of the proof of the smoothing property is the purely algebraic estimate

$$\|(S_l)^m u_l\|_{A_l} = \|(I - \tau D_l^{-1} A_l)^m u_l\|_{A_l} \preceq m^{-1}\|u_l\|_{D_l},$$

which is well established in multigrid theory. Since the preconditioner D_l is scaled such that $\|(S_l)^m\|_{A_l} \leq 1$, one can conclude by operator interpolation

$$\|(S_l)^m u_l\|_{A_l} \preceq m^{-\alpha}\ \|u_l\|_{[A_l,D_l]_\alpha}$$

for any $\alpha \in [0,1]$. The norm $\|.\|_{[A_l,D_l]_\alpha}$ is the interpolation norm between the energy norm ($\alpha = 0$) and the $\|.\|_{D_l}$-norm ($\alpha = 1$). For the AS smoother applied for second order problems one gets

$$\|u_l\|_{D_l} \simeq h_l^{-1}\|u_l\|_0.$$

The interpolation norm for this case is

$$\|u_l\|_{[A_l,D_l]_\alpha} = h_l^{-\alpha}\|u_l\|_{H^{1-\alpha}}.$$

If one can establish the *link* between the algebraic smoothing property and the approximation property, namely

$$\|u\|_{[A_l,D_l]_\alpha} \leq C_s \|u\|_{l,\tilde{0}}, \tag{3.47}$$

the smoothing property (3.46) is proved. We collect multigrid convergence estimates in the following theorem:

Theorem 3.7 (Abstract multigrid convergence). *Let the multigrid procedure be as defined in Algorithm 1. Assume there holds the scaling (3.38), the approximation property (3.45) and the estimate (3.47) used for the smoothing property. Then the following multigrid methods lead to optimal solvers:*

- *The W-cycle multigrid scheme with sufficiently many smoothing steps leads to a convergent method. The contraction number is bounded by*

$$\|M_L\| \leq C m^{-\alpha/2} \tag{3.48}$$

  with a constant C only depending on C_a and C_s, but independent of the level L.

- *The variable V-cycle with $m_L \geq 1$ and $\beta_0 m_l \leq m_{l-1} \leq \beta_1 m_l$ smoothing steps ($1 < \beta_0 < \beta_1$) leads to a preconditioner C_L. The condition number is bounded by*

$$\kappa(C_L^{-1} A_L) \leq 1 + C m^{-\alpha/2} \tag{3.49}$$

  with a constant C only depending on C_a, C_s, and β_0, β_1, but independent of the level L.

Proof. The proof of the W-cycle method follows from the approximation property (3.45) and the smoothing property (3.46) by induction. We refer to [Hac85].

The proof of the variable V-cycle method is provided in [BPX91], see also [Bra93]. First, we derive the type of approximation property used there, then we translate the notation.

Because $D_l^{-1} A_l$ is a symmetric and positive definite operator in the inner product $D_l(.,.)$, one can define the Hilbert scale

$$|||u_l|||_{l,s}^2 := D_l((D_l^{-1} A_l)^s u_l, u_l).$$

For a shorter notation, we neglect the index l indicating the level. Especially, there holds

$$\begin{aligned} |||u_l|||_0^2 &= \|u_l\|_{D_l}^2, \\ |||u_l|||_1^2 &= \|u_l\|_{A_l}^2, \\ |||u_l|||_2^2 &= A(D_l^{-1} A_l u_l, u_l). \end{aligned}$$

For $\alpha \in (0,1)$, the interpolation norm used above can be expressed by

$$\|u_l\|_{[A_l,D_l]_\alpha} = |||u_l|||_{1-\alpha}. \tag{3.50}$$

There holds the estimate of logarithmic convexity

$$A_l(u_l, v_l) \leq |||u_l|||_{1-s} |||v_l|||_{1+s} \tag{3.51}$$

for $s \in \mathbb{R}$, and the interpolation estimate

$$|||u_l|||_s \leq |||u_l|||_{s_1}^{\alpha} \, |||u_l|||_{s_2}^{1-\alpha} \tag{3.52}$$

for $s_1, s_2 \in \mathbb{R}$, $\alpha \in (0,1)$ and $s = \alpha s_1 + (1-\alpha)s_2$.

We use (3.51) to estimate

$$|A_l((I - R_l A_{l-1}^{-1} R_l^T A_l)u_l, u_l)| \leq |||(I - R_l A_{l-1}^{-1} R_l^T A_l)u_l|||_{1-\alpha} \, |||u_l|||_{1+\alpha}.$$

By (3.50), the approximation property (3.45) and estimate (3.47) we get

$$|||(I - R_l A_{l-1}^{-1} R_l^T A_l)u_l|||_{1-\alpha} \leq C_a C_s \|u_l\|_{A_l}.$$

We go on and use interpolation

$$\begin{aligned} |A_l((I - R_l A_{l-1}^{-1} R_l^T A_l)u_l, u_l)| &\leq C_a C_s \|u_l\|_{A_l} \, |||u_l|||_{1+\alpha} \\ &\leq C_a C_s |||u_l|||_1 \, |||u_l|||_2^{\alpha} \, |||u_l|||_1^{1-\alpha} \\ &\leq C_a C_s |||u_l|||_1^{2-\alpha} \, |||u_l|||_2^{\alpha}. \end{aligned} \tag{3.53}$$

This estimate corresponds to [Bra93], Assumption 10, page 62. We translate the present notation to the formulation used in [Bra93], which we will indicate by subscripts or superscripts Br. We are free to set the inner product on level l to

$$(u, v)_{Br,l} = (D_l u, v)_l.$$

Then the operator $A_l^{Br} : V_l \to V_l$ evaluates by

$$(A_l u_l, v_l)_l = A_l(u_l, v_l) = (A_l^{Br} u_l, v_l)_{Br,l} = (D_l A_l^{Br} u_l, v_l)_l$$

to

$$A_l^{Br} = D_l^{-1} A_l.$$

The smoother R_l^{Br} evaluates by

$$S_l = I - D_l^{-1} A_l = I - R_l^{Br} A_l^{Br}$$

to

$$R_l^{Br} = I.$$

We set $\lambda_l^{Br} = 1$. We want to apply [Bra93], Theorem 4.6. A.4 and A.4* are fulfilled for the symmetric smoother. A.10 is (3.53) with $\alpha_{Br} = \alpha/2$. A.12 is assumed in the theorem. Estimate (3.49) follows by dividing η_1^{Br} by η_0^{Br}. □

We have formulated the simple approximation property (3.45) instead of the stronger estimate

$$\|u_l - R_l A_{l-1}^{-1} R_l^T A_l u_l\|_{l,\tilde{0}} \leq C_a \sup_{v_l \in V_l} \frac{A_l(u_l, v_l)}{\|u_l\|_{l,\tilde{0}}}, \tag{3.54}$$

which is usually used. This is paid by reducing the dependency on the number of smoothing steps from $Cm^{-\alpha}$ to $Cm^{-\alpha/2}$. The reason for choosing (3.45) instead of (3.54) is that only the former one could be verified for parameter dependent problems.

If the smoothing property and the approximation property are available, the door is open for other multilevel algorithms, like full multigrid methods [Hac85], Cascadic multigrid methods [BD96], [Sha96], [BD99], and extension operators [HLMN94], [Haa97].

Chapter 4

Robust Preconditioning for Parameter Dependent Problems

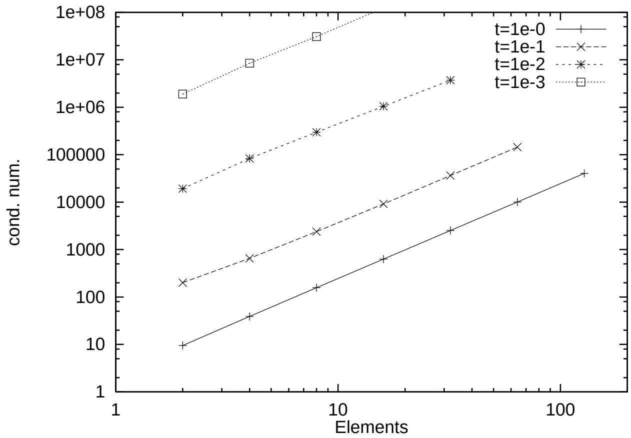
In this chapter we combine the results of the former two to construct preconditioners for parameter dependent problems. We will start with one level and two level preconditioners, and then go on with the multigrid case.

4.1 Local preconditioning

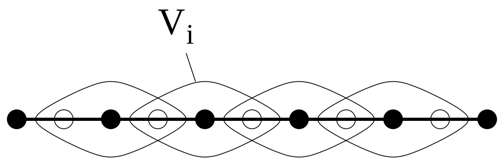
First, we show by means of an example that the condition number of standard preconditioners may deteriorate for parameter dependent problems. Let A_h be the operator obtained by a stable discretization for the Timoshenko beam model, i.e.

$$(A_h(w_h, \beta_h), (v_h, \delta_h))_h = \int_I \beta_h' \delta_h' + t^{-2}(\overline{w_h' - \beta_h}^h)(\overline{v_h' - \delta_h}) \, \mathrm{dx}.$$

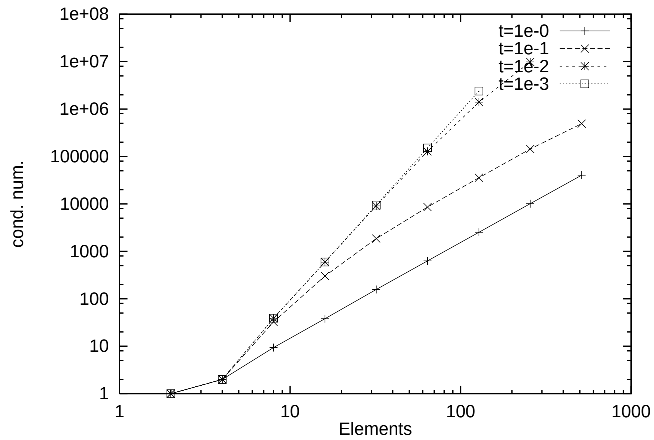
Let D_{Jac} be the Jacobi preconditioner for A_h. We measured the spectral condition number $\kappa(D_{Jac}^{-1} A_h)$ for different values of the mesh size h and the parameter t. The results are:



Now, we try the block Jacobi preconditioner D_{BJ} with the blocks indicated below. Three by three nodes are connected with an overlap of one node.



The corresponding results for $\kappa(D_{BJ}^{-1} A_h)$ are:



We see, the preconditioner is robust with respect to the small parameter, if the triangulation is fixed. For the interesting range $t < h$ it shows the same dependence on the mesh size as a local preconditioner for 4^{th} order problems.

The reason why the block Jacobi preconditioner is robust, while the pointwise Jacobi deteriorates deals with the kernel $V_{h,0}$ of the operator Λ_h. These are the functions

$$V_{h,0} = \{(w_h, \beta_h) : \overline{w_h' - \beta_h}^h = 0\}.$$

On the kernel, the energy norm $\|.\|_{A_h}$ behaves like

$$\|(w_h, \beta_h)\|_{A_h}^2 \simeq \|\beta_h\|_1^2 \qquad \forall\, (w_h, \beta_h) \in V_{h,0}.$$

The norm induced by the diagonal behaves always like

$$\|(w_h, \beta_h)\|_{D_{Jac}}^2 \simeq (h^{-2} + t^{-2})\|\beta_h\|_0^2 + h^{-2}t^{-2}\|w\|_0^2.$$

By choosing smooth functions $(w_h, \beta_h) \in V_{h,0}$, one sees that the estimate

$$h^2 t^2 \, D_{Jac} \preceq A_h$$

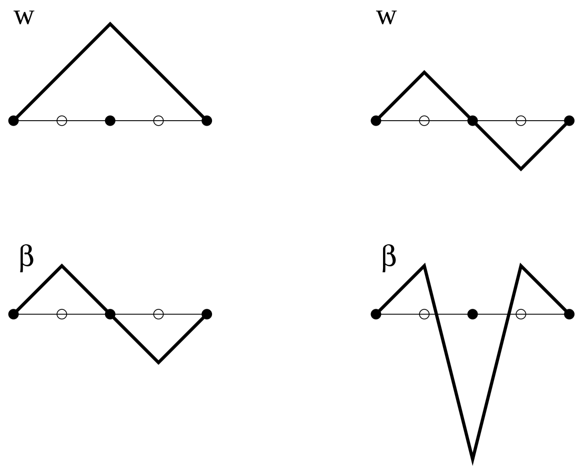


Figure 4.1: Basis functions for kernel $V_{h,0}$

is asymptotically sharp. Because also $A_h \preceq D_h$ is sharp, the condition number behaves like

$$\kappa(D_{Jac}^{-1} A_h) \simeq h^{-2} t^{-2}.$$

One observes, that errors in $V_{h,0}$ are smoothed out very slowly be the Jacobi preconditioner. The kernel is spanned by a set of basis functions of two different types, see Figure 4.1. The block Jacobi preconditioner is designed to capture these basis functions.

The idea of the block smoother is related to the methods of [AFW97b] and [Hip99] for problems in $H(\text{div})$ and $H(\text{rot})$

The following Theorem gives an abstract framework for one level preconditioners for parameter dependent problems.

Theorem 4.1. *Let $A_h(.,.) : V_h \times V_h \to \mathbb{R}$ be the bilinear form*

$$A_h(u_h, v_h) = a_h(u_h, v_h) + \varepsilon^{-1}\, c_h(\Lambda_h u_h, \Lambda_h v_h) \tag{4.1}$$

with the assumptions made in Section 2.3. Let $\{V_i\}$ be a local space splitting with overlap N_O such that

$$V = \sum V_i.$$

We assume that functions $u_h \in V_h$ and kernel functions $u_0 \in V_{h,0}$ can be split locally with estimates depending on the mesh size, i. e.

$$\inf_{\substack{u_h = \sum u_i \\ u_i \in V_i}} \sum_{i=1}^{M} \|u_i\|_{V_h}^2 \leq c_1(h)\, \|u\|_{V_h}^2 \tag{4.2}$$

$$\inf_{\substack{u_0 = \sum u_{0,i} \\ u_i \in V_i \cap V_{h,0}}} \sum_{i=1}^{M} \|u_{0,i}\|_{a_h}^2 \leq c_2(h)\, \|u_0\|_{V_h}^2. \tag{4.3}$$

Assume, there holds the inverse estimate

$$\|q_h\|_{c_h} \leq c_3(h)\|q_h\|_{Q_h,0} \qquad \forall\, q_h \in \Lambda_h V_h, \tag{4.4}$$

where $\|.\|_{Q_h,0}$ *is defined in (2.64). Then the additive Schwarz preconditioner* D_h *built on the space splitting* $\{V_i\}$ *fulfills*

$$(c_1(h) + c_2(h)c_3(h)^2)^{-1}\, D_h \preceq A_h \leq N_O\, D_h \tag{4.5}$$

on V_h*. The bounds are independent of the parameter* ε*.*

Proof. The upper estimate follows immediately from Lemma 3.2. We choose an $u_h \in V_h$. First, we split

$$u_h = u_0 + u_1$$

with $u_0 \in V_{h,0}$ by projection into $V_{h,0}$. Let $(u_0, p_0) \in X_h$ be the solution of the mixed variational problem with the bilinear form of the limit

$$B_h^0((u_0, p_0), (v_h, q_0)) = a(u_h, v_h) \qquad \forall\, (v_h, q_h) \in X_h.$$

Theorem 2.8 (Brezzi) bounds u_0 by

$$\|u_0\|_{V_h} \preceq \|u_h\|_{V_h}. \tag{4.6}$$

Then $u_1 = u_h - u_0$ fulfills

$$\begin{aligned} B_h^0((u_1, -p_0), (v_h, q_h)) &= B_h^0((u_h, 0), (v_h, q_h)) - B_h^0((u_0, p_0), (v_h, q_h)) \\ &= a_h(u_h, v_h) + c_h(\Lambda_h u_h, q_h) - a_h(u_h, v_h) \\ &= c_h(\Lambda_h u_h, q_h). \end{aligned}$$

We use Cauchy-Schwarz, the definition (4.1) of $A_h(.,.)$ and the inverse inequality (4.4) to proceed with

$$\begin{aligned} B_h^0((u_1, p_0), (v_h, q_h)) &\leq \|\Lambda_h u_h\|_{c_h}\, \|q_h\|_{c_h} \\ &\leq \varepsilon^{1/2}\|u_h\|_{A_h}\, c_3(h)\|q_h\|_{Q_h,0}. \end{aligned}$$

By Theorem 2.8 we get

$$\|u_1\|_{V_h} \preceq \varepsilon^{1/2} c_3(h)\|u_h\|_{A_h}. \tag{4.7}$$

We use assumptions (4.3) and (4.2) to split $u_0 = \sum u_{0,i}$ and $u_1 = \sum u_{1,i}$. Then we use the additive Schwarz lemma (Theorem 3.1) to estimate

$$\begin{aligned} \|u_h\|_{D_h}^2 &= \inf_{\substack{u_h = \sum_i u_i \\ u_i \in V_i}} \|u_i\|_{A_h}^2 \\ &\leq \sum \|u_{0,i} + u_{1,i}\|_{A_h}^2 \\ &\leq 2\sum \left\{ \|u_{0,i}\|_{A_h}^2 + \|u_{1,i}\|_{A_h}^2 \right\} \\ &\preceq \sum \left\{ \|u_{0,i}\|_{a_h}^2 + \varepsilon^{-1}\|u_{1,i}\|_{V_h}^2 \right\} \end{aligned}$$

We finish with the inverse estimates (4.3) and (4.2) and the stability estimates (4.6) and (4.7):

$$\begin{aligned}
\|u_h\|_{D_h}^2 &\preceq c_1(h)\|u_0\|_{V_h}^2 + \varepsilon^{-1}c_2(h)\|u_1\|_{V_h}^2 \\
&\preceq c_1(h)\|u_h\|_{A_h}^2 + c_2(h)c_3(h)^2\|u_h\|_{A_h}^2 \\
&\preceq \left\{c_1(h) + c_2(h)c_3(h)^2\right\}\|u_h\|_{A_h}^2.
\end{aligned}$$

□

4.1.1 Timoshenko beam and Reissner Mindlin plate

We will prove now the ε-robust condition number of the block Jacobi preconditioner for the beam model of Timoshenko. We will apply Theorem 4.1. Because $\|.\|_{V_h} = \|.\|_1$, assumptions (4.2) is fulfilled with

$$c_1(h) \preceq h^{-2},$$

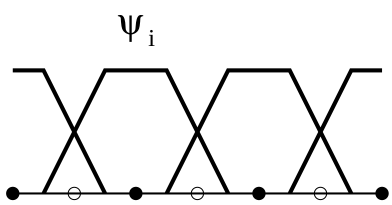
and assumption (4.4) is fulfilled with

$$c_3(h) \preceq 1,$$

since $\|.\|_{c_h} = \|.\|_{Q_h,0} = \|.\|_0$. We will verify the splitting inequality for kernel functions:

$$\inf_{\substack{(w_h,\beta_h)=\sum(w_i,\beta_i)\\(w_i,\beta_i)\in V_{h,0}\cap V_i}} \sum \|(w_i,\beta_i)\|_{a_h}^2 \preceq h^{-4}\|(w_h,\beta_h)\|_{V,h}^2. \tag{4.8}$$

We use a partitioning of unity $\{\psi_i\}$ as drawn in the figure below:



The support of each function ψ is inside four elements. Recall the construction of the Fortin operator (2.81). There we used the nodal interpolation operator $I_h^{F,1}$, and the projector $I_h^{F,2}$ adjusting bubbles. The co-projector maps onto $V_{h,0}$:

$$I - I_h^{F,2} : V_h \to V_{h,0}.$$

Thus, the combination

$$(I - I_h^{F,2})I_h^{F,1} : V \to V_{h,0}$$

is a projection, too. We chose the kernel splitting as

$$(w_i,\beta_i) := (I - I_h^{F,2})I_h^{F,1}(\psi_i \cdot (w_h,\beta_h)).$$

We have obtained in Section 2.4.1 that $I_h^{F,2}$ is continuous with respect to the norm $h^{-1}\|w_h\|_0 + \|\beta_h\|_0$. Thus we get

$$\begin{aligned}\|(w_i,\beta_i)\|_{a_h} &= |\beta_i|_1 \preceq h^{-1}\|\beta\|_0 \\ &\preceq h^{-2}\|\psi_i w_h\|_0 + h^{-1}\|\psi_i\beta_h\|_0 \\ &\preceq h^{-2}\|(w_h,\beta_h)\|_{0,\omega_i}\end{aligned}$$

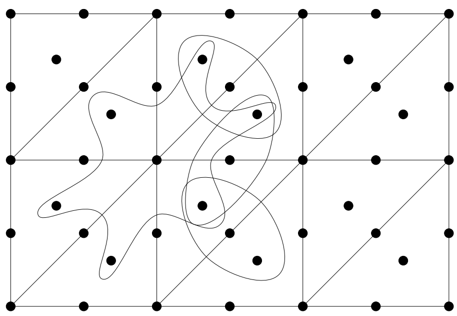
Summing up gives the inverse estimate

$$\sum_i \|(w_i,\beta_i)\|_{a_h}^2 \preceq h^{-4}\|(w_h,\beta_h)\|_0^2 \preceq h^{-4}\|(w_h,\beta_h)\|_{V_h}^2.$$

The local smoother for the Reissner Mindlin plate is analogous. The inverse inequalities (4.2), (4.3) and (4.4) hold with

$$\begin{aligned}c_1(h) &\simeq h^{-2}, \\ c_2(h) &\simeq h^{-4}, \\ c_3(h) &\simeq h^{-1}.\end{aligned}$$

Thus, the same upper bound h^{-4} results as for the beam. The subspaces $V_{l,i}$ have to contain $(I - I_h^{F,2})\varphi_i$, where φ_i is a nodal base function. This is possible with subspaces spanned by the nodal base functions clustered as follows:



4.1.2 Nearly incompressible materials

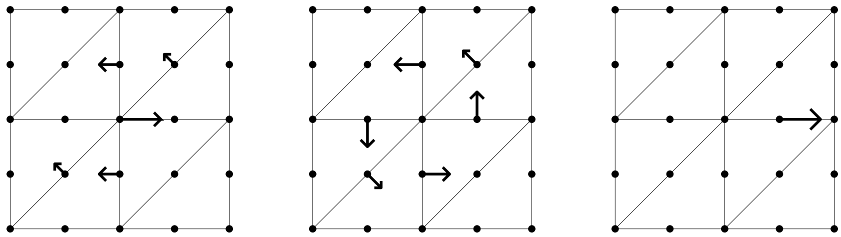
We will check the conditions of Theorem 4.1 for the bilinear form of nearly incompressible materials

$$A_h(u_h,v_h) = (e(u_h),e(v_h)) + \varepsilon^{-1}(\overline{\operatorname{div} u_h}^h, \overline{\operatorname{div} v_h}^h)_0.$$

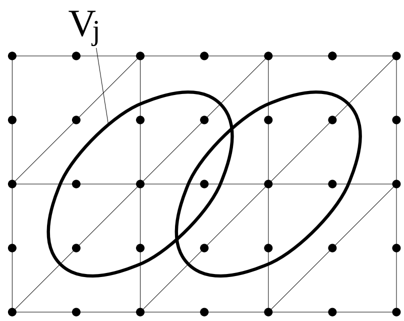
The estimates (4.2) and (4.4) follow directly form the spaces $V_h \subset H^1$ and $Q_h \subset L_2$, namely

$$\begin{aligned}c_1(h) &\simeq h^{-2}, \\ c_3(h) &\simeq 1.\end{aligned}$$

We will prove the splitting estimate for the kernel functions. We are guided by a basis for the kernel, which consists of the following type of functions:



These basis functions cannot be split into kernel functions of smaller support. Thus, we will use subspaces spanned by the nodal base functions connected as follows:



The plan for the analysis is to lift the divergence free function to the potential space. For this, we have to assume that Ω is simply connected, and there is only one simply connected part with natural boundary conditions. Then the potential function is split into local functions. The splitting of the original function is obtained by interpolating the curl of the local potential functions by the Fortin operator.

We fix $u_h \in V_{h,0}$. We construct the lifting operator $E : V_h \to V$ as follows. Define the spaces

$$\begin{aligned} V_{loc} &:= \{v \in V : v = 0 \text{ on } \partial T \; \forall T \in \mathcal{T}_h\} \\ Q_{loc} &:= \{q \in Q : \overline{q}^h = 0\}. \end{aligned}$$

They split into local spaces on the elements. The pair $V_{loc} \times Q_{loc}$ is stable for Stokes. Now let

$$u := Eu_h := u_h + w$$

with $(w, r) \in V_{loc} \times Q_{loc}$ be the solution of the variational problem

$$B^0((w, r), (v, q)) = -B^0((u_h, 0), (v, q)) \qquad \forall (v, q) \in V_{loc} \times Q_{loc}.$$

We observe that

$$(\operatorname{div}(w + u_h), q)_0 = B^0((w + u_h, r), (0, q)) = 0 \qquad \forall q \in Q_{loc}.$$

Applying Gauss' theorem to w on each element, and using the assumption $u_h \in V_{h,0}$, we also get

$$\overline{\operatorname{div}(w + u_h)}^h = 0.$$

We have constructed an $u \in V$ such that

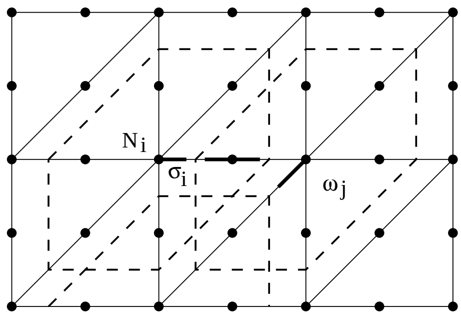
$$\|u\|_V \preceq \|u_h\|_V,$$

$$\operatorname{div} u = 0,$$

and there holds

$$u = u_h \qquad \text{on } \partial T \; \forall \, T \in \mathcal{T}_h.$$

Let $I_h^F = I_h^{F,1} + I^{F,2}(I - I_h^{F,1})$ be a Fortin operator based on a Scott-Zhang projection operator. It shall be constructed such that it uses only function values at $\partial T, T \in \mathcal{T}_h$, see the Figure below:



The Fortin operator is an left inverse to the extension operator, i.e.

$$I_h^F E u_h = u_h \qquad \forall \, u_h \in V_h.$$

Due to the assumptions on the domain, there exists an $\varphi \in H^2_{0,D}(\Omega)$ such that

$$\operatorname{rot} \varphi = u.$$

Let $\{\psi_i\}$ be a partitioning of unity with $\operatorname{supp}\{\psi_i\} \subset \omega_i$. We can define the splitting as

$$u_i = I_h^F \operatorname{rot}(\psi_i \varphi).$$

It is a splitting, because

$$\sum u_i = \sum I_h^F \operatorname{rot}(\psi_i \varphi) = I_h^F \operatorname{rot}(\sum \psi_i \varphi) = I_h^F \operatorname{rot} \varphi = I_h^F u = I_h^F E u_h = u_h.$$

Using the continuity of the Fortin operator, the properties of the partitioning of unity, and the construction of φ we can bound the kernel splitting by

$$\begin{aligned} \sum \|u_i\|_{a_h}^2 &= \sum \|I_h^F \operatorname{rot}(\psi_i \varphi)\|_1^2 \\ &\preceq \sum \|\operatorname{rot}(\psi_i \varphi)\|_1^2 \\ &\preceq \sum \|(\psi_i \varphi)\|_2^2 \preceq h^{-4} \|\varphi\|_2^2 \\ &\preceq h^{-4} \|u\|_1^2 \preceq h^{-4} \|u_h\|_{V_h}^2. \end{aligned}$$

Thus, we obtained the inverse estimate

$$c_2(h) \preceq h^{-4}.$$

We mention that the original problem is of second order, but the lifting to the potential space lets it behave like a fourth order problem. We have made assumption onto the domain. They are really essential, because curls around a hole cannot be split into local curls. When we use a coarse grid system, these global curls are caught by the coarse grid.

4.2 Two Level Preconditioning

In the last section we have constructed ε-robust local preconditioners. Now, we will see how to combine the local preconditioner with a coarse grid correction to obtain a preconditioner with robust and optimal condition number.

The bilinear form

$$A_h(u_h, v_h) = a_h(u_h, v_h) + \frac{1}{\varepsilon} c_h(\Lambda_h u_h, \Lambda_h v_h) \tag{4.9}$$

stems from the reduction of a mixed finite element problem with the bilinear form $B_h(.,.) : X_h \times X_h \to \mathbb{R}$. We need an additional coarse grid bilinear form

$$A_H(u_H, v_H) = a_H(u_H, v_H) + \frac{1}{\varepsilon} c_H(\Lambda_H u_H, \Lambda_H v_H) \tag{4.10}$$

coming from a mixed finite element method with the bilinear form $B_H(.,.) : X_H \times X_H \to \mathbb{R}$ on the coarser space $X_H = V_H \times Q_H$. The forms $B_h(.,.)$ and $B_H(.,.)$ may differ. They might be obtained by stabilized methods with different weights. We do not assume that the primal spaces are nested, i.e.

$$V_H \subset V_h \qquad \text{or} \qquad V_H \not\subset V_h.$$

Thus we need a prolongation operator

$$R_H^{V,0} : V_H \to V_h.$$

If the spaces are nested, we may take the natural embedding. We call $R_H^{V,0}$ trivial prolongation operator, because we will actually use a more complicated one. The trivial prolongation operator shall be bounded by

$$\|R_H^{V,0} u_H\|_{V,h} \preceq \|u_H\|_{V,H} \qquad \forall\, u_H \in V_H. \tag{4.11}$$

The spaces for the dual variable shall be nested

$$Q_H \subset Q_h.$$

This is not a strong assumption, because the space Q is L_2 or weaker for our applications. We assume that the according norms are equivalent, i.e.

$$\|q_H\|_{Q_H} \simeq \|q_H\|_{Q_h} \qquad \forall\, q_H \in Q_H \tag{4.12}$$

and

$$\|q_H\|_{c_H} \simeq \|q_H\|_{c_h} \qquad \forall\, q_H \in Q_H. \tag{4.13}$$

We have fixed the coarse grid operator A_H. Next, have to define the actual prolongation operator

$$R_H^V : V_H \to V_h.$$

In Lemma 3.5 we have shown, that the continuity in energy is a necessary condition for an optimal two level preconditioner. Thus

$$\|R_H^V u_H\|_{A_h} \preceq \|u_H\|_{A_H}$$

must be fulfilled uniformly with respect of the small parameter ε. To get a feeling for this condition, we take a coarse grid kernel function

$$u_H \in V_{H,0} = \text{kern } \Lambda_H$$

The coarse grid energy

$$\|u_H\|_{A_H}^2 = \|u_H\|_{a_H}^2$$

does not depend on ε. The fine grid energy of the prolongated function

$$\|R_H^V u_H\|_{A_h}^2 = \|R_H^V u_H\|_{a_h}^2 + \frac{1}{\varepsilon}\|\Lambda_h R_H^V u_H\|_{c_h}^2$$

has to be bounded by a constant independent of ε. This is only possible, if $\|\Lambda_h R_H^V u_H\|_{c_h}^2 \leq \varepsilon$. Essentially, this means that the prolongation has to map the coarse grid kernel to the fine grid kernel:

$$R_H^V : V_{H,0} \to V_{h,0}. \tag{4.14}$$

Indeed, the prolongation may produce a small perturbation out of $V_{h,0}$.

We consider the Timoshenko beam as example, see Figure 4.2. Let (w_H, β_H) be a function in $V_{H,0}$. This means that

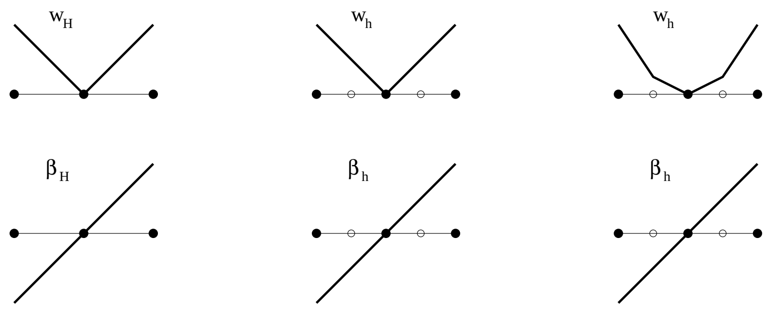
$$\overline{w_H' - \beta_H}^H = 0.$$

The prolongation has to construct functions (w_h, β_h) such that

$$\overline{w_h' - \beta_h}^h = 0, \tag{4.15}$$

i.e. the averaging is done on the fine grid. The trivial prolongation $R_H^{V,0}$ can be chosen as embedding. It does not fulfill that condition. But, by a local modification of the functions, condition (4.15) can be achieved. We can adjust the nodal values of the deflection w_h on the new nodes.

It is always the question which degrees of freedom have to be adjusted. The following theorem gives an abstract guide for the construction of robust prolongation operators:



$\mathcal{T}_H, (w_H, \beta_H)$ $\quad\quad$ $\mathcal{T}_h, R_H^{V,0}(w_H, \beta_H)$ $\quad\quad$ $\mathcal{T}_h, R_H^V(w_H, \beta_H)$

Figure 4.2: Prolongation for Timoshenko beam

Theorem 4.2 (Robust prolongation operator). *Assume that the dual space Q_h is split $c_h(.,.)$-orthogonal into*

$$Q_h = \widetilde{Q}_H \oplus Q_T \tag{4.16}$$

with $\widetilde{Q}_H \subset Q_H$. The corresponding projection operators are $P^{c_h}_{\widetilde{Q}_H}$ and $P^{c_h}_{Q_T}$. The trivial prolongation operator $R_H^{V,0}$ is assumed to satisfy

$$c_h(\Lambda_h R_H^{V,0} u_H, \widetilde{q}_H) = c_H(\Lambda_H u_H, \widetilde{q}_H) \qquad \forall\, \widetilde{q}_h \in \widetilde{Q}_H. \tag{4.17}$$

Assume there exists a space $V_T \subset V_h$ such that the bilinear form $B_h(.,.)$ is stable on $X_T := V_T \times Q_T$, i.e.

$$\sup_{(u_T,p_T)\in X_T} \frac{B_h((u_T,p_T),(v_T,q_T))}{\|(u_T,p_T)\|_{X,h}} \succeq \|(v_T,q_T)\|_{X_h} \qquad \forall\,(v_T,q_T) \in X_T. \tag{4.18}$$

Additionally, we assume the orthogonality condition

$$c_h(\Lambda_h v_T, \widetilde{q}_H) = 0 \qquad \forall\, v_T \in V_T,\ \forall\, \widetilde{q}_h \in \widetilde{Q}_H. \tag{4.19}$$

Let $w_T \in V_T$ be the solution of the variational problem

$$A_h(w_T, v_T) = A_h(R_H^{V,0} u_H, v_T) \qquad \forall\, v_T \in V_T. \tag{4.20}$$

Then the prolongation R_H^V defined by

$$R_H^V u_H := R_H^{V,0} u_H - w_T \tag{4.21}$$

is continuous in energy norm:

$$\|R_H^V u_H\|_{A_h} \preceq \|u_H\|_{A_H} \qquad \forall\, u_H \in V_H. \tag{4.22}$$

Remark 4.3. *The splitting of Q_h has the following meaning. The constraints according to $\widetilde{Q}_H$ are inherited from the coarse grid. This is only possible, if the trivial prolongation preserves the constraints. The constraints belonging to Q_T are fulfilled by construction. For an efficient implementation it is important that the problem can be solved locally.*

Proof. First, we observe that the orthogonality relations (4.16) and (4.19) imply that

$$\Lambda_h V_T \subset Q_T.$$

Using the relations

$$\begin{aligned} B_h((u_h, \varepsilon^{-1}\Lambda_h u_h), (v_h, 0)) &= A_h(u_h, v_h), \\ B_h((u_h, \varepsilon^{-1}\Lambda_h u_h), (0, q_h)) &= 0 \end{aligned}$$

for all $u_h, v_h \in V_h$ and $q_h \in Q_h$ we can rewrite (4.20) in mixed variables:

$$B_h((w_T, \varepsilon^{-1}\Lambda_h w_T), (v_T, q_T)) = B_h((R_H^{V,0} u_H, \varepsilon^{-1}\Lambda_h R_H^{V,0} u_H), (v_T, q_T)) \qquad \forall\, (v_T, q_T) \in X_T.$$

By the orthogonality relations (4.16) and (4.19) we can insert the projection $P_{Q_T}^{c_h}$:

$$B_h((w_T, \varepsilon^{-1} P_{Q_T}^{c_h}\Lambda_h w_T), (v_T, q_T)) = B_h((R_H^{V,0} u_H, \varepsilon^{-1} P_{Q_T}^{c_h}\Lambda_h R_H^{V,0} u_H), (v_T, q_T)).$$

We bring the dual variable to the left hand side, and use continuity of $B_h(.,.)$ to obtain

$$\begin{aligned} B_h((w_T, \varepsilon^{-1} P_{Q_T}^{c_h}\Lambda_h(w_T - R_H^{V,0} u_H)), (v_T, q_T)) &= B_h((R_H^{V,0} u_H, 0), (v_T, q_T)) \\ &\preceq \|R_H^{V,0} u_H\|_{V_h} \, \|(v_T, q_T)\|_{X_h}. \end{aligned}$$

By stability of $B_h(.,.)$ on X_T we get

$$\|w_T\|_{V_h} + \|\varepsilon^{-1} P_{Q_T}^{c_h}\Lambda_h(w_T - R_H^{V,0} u_H)\|_{Q_h} \preceq \|R_H^{V,0} u_H\|_{V_h} \preceq \|u_H\|_{V_H}. \tag{4.23}$$

This bounds the first term of estimate (4.22):

$$\|R_H^{V} u_H\|_{a_h} \preceq \|R_H^{V} u_H\|_{V_h} \preceq \|R_H^{V,0} u_H - w_h\|_{V_h} \preceq \|u_H\|_{V_H}. \tag{4.24}$$

We are left to bound

$$\varepsilon^{-1}\,\|\Lambda_h u_h\|_{c_h}^2 = \varepsilon^{-1}\,\|P_{Q_T}^{c_h}\Lambda_h u_h\|_{c_h}^2 + \varepsilon^{-1}\,\|P_{\widetilde{Q}_H}^{c_h}\Lambda_h u_h\|_{c_h}^2.$$

By the definition of the norm $\|q_h\|_{Q_h}^2 := \|q_h\|_{Q_h,0}^2 + \varepsilon\,\|q_h\|_{c_h}^2$ and (4.23) we bound the first term

$$\begin{aligned} \varepsilon^{-1}\,\|P_{Q_T}^{c_h}\Lambda_h u_h\|_{c_h}^2 &\leq \varepsilon^{-2}\,\|P_{Q_T}^{c_h}\Lambda_h u_h\|_{Q_h}^2 \\ &= \varepsilon^{-2}\,\|P_{Q_T}^{c_h}\Lambda_h(R_H^{V,0} u_H - w_T)\|_{Q_h}^2 \\ &\preceq \|u_H\|_{V_H}. \end{aligned} \tag{4.25}$$

We use $P^{c_h}_{\widetilde{Q}_h} \Lambda_h w_T = 0$, property (4.17) of the prolongation and norm equivalence $\|.\|_{c_h} \simeq \|.\|_{c_H}$ to bound the second term

$$\begin{aligned}
\|P^{c_h}_{\widetilde{Q}_H} \Lambda_h u_h\|_{c_h} &= \|P^{c_h}_{\widetilde{Q}_H} \Lambda_h R_H^{V,0} u_H\|_{c_h} \\
&= \sup_{\widetilde{q}_H \in \widetilde{Q}_H} \frac{c_h(\Lambda_h R_H^{V,0} u_H, \widetilde{q}_H)}{\|\widetilde{q}_H\|_{c_h}} \\
&= \sup_{\widetilde{q}_H \in \widetilde{Q}_H} \frac{c_H(\Lambda_H u_H, \widetilde{q}_H)}{\|\widetilde{q}_H\|_{c_h}} \\
&\leq \sup_{\widetilde{q}_H \in \widetilde{Q}_H} \frac{\|\Lambda_H u_H\|_{c_H} \|\widetilde{q}_H\|_{c_H}}{\|\widetilde{q}_H\|_{c_h}} \\
&\leq \|\Lambda_H u_H\|_{c_H}.
\end{aligned}$$

This gives

$$\varepsilon^{-1} \|P^{c_h}_{\widetilde{Q}_H} \Lambda_h u_h\|^2_{c_h} \preceq \varepsilon^{-1} \|\Lambda_H u_H\|^2_{c_H} \preceq \|u_H\|^2_{A_H}. \tag{4.26}$$

Combining (4.24), (4.25), and (4.26) proves the theorem. □

According to Lemma 3.5 we have to prove the existence of an interpolation operator

$$I_H^V : V_h \to V_H$$

which is continuous in energy, i.e.

$$\|I_H^V u_h\|_{A_H} \preceq \|u_h\|_{A_h} \qquad \forall\, u_h \in V_h. \tag{4.27}$$

As in Theorem 2.13, the Fortin operator I_H^F has the desired properties. Let

$$Q_H = Q_{H,0} \oplus Q_{H,1}$$

be a stable decomposition with respect to $\|.\|_{c_H}$. The subspace $Q_{H,0}$ fulfills

$$\|q_H\|^2_{Q_H} \preceq \varepsilon \, \|q_H\|^2_{c_H} \qquad \forall\, q_H \in Q_H.$$

The Fortin operator shall be continuous in the norm

$$\|I_H^F u_h\|_{V,H} \preceq \|u_h\|_{V,h}$$

and preserve constraints belonging to $Q_{H,1}$, i.e.

$$c_h(\Lambda_h u_h, q_1) = c_H(\Lambda_H I_H^F u_h, q_1) \qquad \forall\, q_1 \in Q_{H,1}.$$

Then Theorem 2.13 proves its continuity (4.27). The last assumption of Lemma 3.5, namely the stable decomposability of

$$u_f := u_h - R_H^V I_H^F u_h,$$

will be formulated for the specific examples.

4.2.1 Nearly incompressible materials

Let us consider the case of nearly incompressible materials. We have the fine grid bilinear-form

$$A_h(u_h, v_h) = (e(u_h), e(v_h))_0 + \varepsilon^{-1} (\overline{\text{div} u_h}^h, \overline{\text{div} v_h}^h)_0$$

at the fine grid space V_h, and the coarse grid form

$$A_H(u_H, v_H) = (e(u_H), e(v_H))_0 + \varepsilon^{-1} (\overline{\text{div} u_H}^H, \overline{\text{div} v_H}^H)_0$$

at a coarse grid space V_H of P_2 elements. The coarse triangulation $\mathcal{T}_H$ may be arbitrarily coarse. The triangulations must be nested.

We construct the prolongation operator R_H^V following Theorem 4.2. Therefore, we split

$$Q_h = Q_H \times Q_T.$$

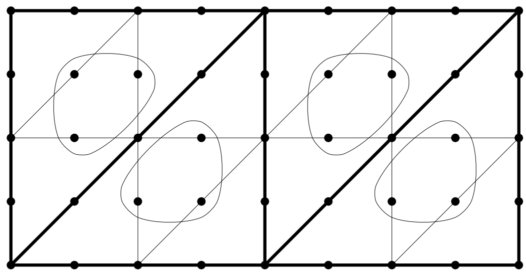
The space Q_H consists of the piecewise constants on the coarse grid, factored by the global constant. The space Q_T is

$$Q_T = \{q_h \in Q_h : \overline{q_h}^H = 0\}.$$

The space V_T is chosen as

$$V_T = \{v_h \in V_h : v_h|_{\partial T_H} = 0 \, \forall T_H \in \mathcal{T}_H\}.$$

If the fine mesh is obtained by bisection, the prolongation requires the solution of the local problems indicated below:



This is a stable pair for the Stokes problem. The orthogonality (4.19) follows by the theorem of Gauss, i.e

$$(\text{div}\, v_T, q_H)_0 = \sum_{T_H} \int_{T_H} \text{div}\, v_T q_H \,\text{dx} = \sum_{T_H} q_H|_T \int_{\partial T_H} (n^T v_T) \,\text{ds} = 0.$$

Thus the prolongation operator consists of solving local Stokes problems in the coarse grid elements T_H. The Fortin operator is such that the function values at coarse grid edges E_H are preserved. By construction, the local complement

$$u_f = u_h - R_H^V I_H^F u_h$$

is bounded by

$$\|u_f\|_1 \preceq \|u_h\|_1.$$

and fulfills

$$\int_{E_H} u_f \, \mathrm{ds} = 0$$

for each coarse grid edge E_H. By the Bramble Hilbert lemma we get

$$\|u_f\|_0 \preceq H \, \|u_f\|_1 \preceq H \, \|u_h\|_1.$$

We will construct a local splitting of mesh size h_{loc} with $h \preceq h_{loc} \preceq H$. There are the two extreme cases $h_{loc} \simeq h$, and $h_{loc} \simeq H$. The first one gives cheap local problems, while the other one leads to an preconditioner with optimal bounds. For $H \simeq h$ both properties are fulfilled.

We split u_f into two parts

$$u_f = u_0 + u_1,$$

with $u_0 V_{h,0}$. This is done using the space X_T once more. Let $(u_1, p_1) \in X_T$ be the solution of

$$B^0((u_1, p_1), (v_T, q_T)) = c(\Lambda_h u_f, q_T) \qquad \forall \, (v_T, q_T) \in X_T. \tag{4.28}$$

Then there holds

$$c(\Lambda_h u_0, q_T) = c(\Lambda_h (u_f - u_1), q_T) = 0$$

per construction. Additionally, applying Gauss' theorem on each element of $\mathcal{T}_H$ and using $\int_{E_H} u_f \, \mathrm{ds} = \int_{E_H} u_1 \, \mathrm{ds} = 0$ we get

$$c(\Lambda_h u_0, q_H) = c(\Lambda_h u_f, q_H) - c(\Lambda_h u_1, q_H) = 0.$$

We got

$$\overline{\operatorname{div} u_0}^h = 0.$$

We proceed as in Section 4.1.2. We lift to $u \in V$ such that

$$\operatorname{div} u = 0 \qquad \text{and} \qquad u = u_0 \text{ on } E_h,$$

and construct $\varphi \in H_0^2$ such that

$$\operatorname{rot} \varphi = u.$$

But now, we have the additional property

$$\int_{E_H} \frac{\partial \varphi}{\partial t} \mathrm{ds} = \int_{E_H} n^T \operatorname{rot} \varphi \, \mathrm{ds} = \int_{E_H} n^T u \, \mathrm{ds} = \int_{E_H} n^T u_f \, \mathrm{ds} = 0.$$

This means that φ has a constant value in all corner nodes of $\mathcal{T}_H$. By the choice of boundary conditions it is 0. An other consequence is that the potential φ exists also for general domains. Applying the Bramble Hilbert lemma to φ, we get the optimal L_2 estimate

$$\|\varphi\|_0 \preceq H^2 \|\varphi\|_2.$$

By multiplication with the p.o.u. at mesh size h_{loc} we get

$$\|\psi_i \varphi\|_2 \preceq h_{loc}^{-2} H^2 \|\varphi\|_2,$$

and as in Section 4.1.2

$$\begin{aligned} \|u_0\|_{D_h}^2 &\preceq \sum_i \|I_h^F \operatorname{rot}(\psi_i \varphi)\|_1 \preceq \sum_i \|(\psi_i \varphi)\|_2^2 \\ &\preceq h_{loc}^{-2} H^2 \|\varphi\|_2^2 \preceq h_{loc}^{-2} H^2 \|u_0\|_1^2. \end{aligned}$$

The other component u_1 vanishes at ∂T_H, thus

$$\|u_1\|_0 \preceq H \|u_1\|_1$$

holds. By the definition (4.28) of u_1 and Theorem 2.8 (Brezzi) there holds

$$\|u_1\|_1 \preceq \|\overline{\operatorname{div} u_f}^h\|_0.$$

Combining we get

$$\|u_1\|_0^2 \preceq H^2 \varepsilon \, \|u_f\|_{A_h}^2.$$

Thus we have optimal L_2 estimate, with the additional factor ε. We can use the partitioning of unity method starting with the original form A_h, namely

$$\begin{aligned} \|u_1\|_{D_h}^2 &= \inf_{u_1 = \sum u_i} \sum \|u_i\|_{A_h}^2 \\ &\preceq \varepsilon^{-1} \sum \|\psi_i u_1\|_1 \preceq \varepsilon^{-1} \{ \|u_1\|_1^2 + h_{loc}^{-2} \|u_1\|_0^2 \} \preceq h_{loc}^{-2} H^2 \|u_f\|_{A_h} \end{aligned}$$

Summing up, we get

$$\|u_f\|_{D_h}^2 \preceq \|u_1\|_{D_h}^2 + \|u_2\|_{D_h}^2 \preceq h_{loc}^{-2} H^2 \, \|u_f\|_{A_h}^2 \preceq h_{loc}^{-2} H^2 \|u_h\|_{A_h}^2.$$

If we have $h_{loc} \simeq H$, we have constructed and analyzed a preconditioner with robust and optimal spectral bounds.

4.2.2 Reissner Mindlin plate

We analyze the two level method for the Reissner Mindlin plate model. The fine grid bilinear form $A_h(.,.)$ is

$$\begin{aligned} A_h((w_h, \beta_h), (v_h, \delta_h)) &= a(\beta_h, \delta_h) + \frac{1}{h^2 + t^2} (\nabla w_h - \beta_h, \nabla v_h - \delta_h)_0 \\ &+ \left\{ \frac{1}{t^2} - \frac{1}{h^2 + t^2} \right\} (\overline{\nabla w_h - \beta_h}^h, \overline{\nabla v_h - \delta_h}^h)_0. \end{aligned}$$

The coarse grid form is defined by stabilization and averaging at the coarse grid, i.e.

$$\begin{aligned} A_H((w_H,\beta_H),(v_H,\delta_H)) &= a^b(\beta_H,\delta_H) + \frac{1}{H^2+t^2}(\nabla w-\beta,\nabla v-\delta)_0 \\ &+ \left\{\frac{1}{t^2} - \frac{1}{H^2+t^2}\right\} (\overline{\nabla w-\beta}^H, \overline{\nabla v-\delta}^H)_0. \end{aligned}$$

We recall the norms of the primal and dual spaces, on the fine and on the coarse level:

$$\begin{aligned} \|(w_h,\beta_h)\|^2_{V_h} &= \|\beta_h\|_1^2 + (h+t)^{-2}\|\nabla w_h - \beta_h\|_0^2 \\ \|(w_H,\beta_H)\|^2_{V_H} &= \|\beta_H\|_1^2 + (H+t)^{-2}\|\nabla w_H - \beta_H\|_0^2 \end{aligned}$$

and

$$\begin{aligned} \|q_h\|^2_{Q_h} &= \{h^2+t^2\}\,\|q_h\|^2_{L_2} \\ \|q_H\|^2_{Q_H} &= \{H^2+t^2\}\,\|q_H\|^2_{L_2}. \end{aligned}$$

To obtain a stable prolongation, we need that the norms $\|.\|_{Q_h}$ and $\|.\|_{Q_H}$ are equivalent. Thus, we have to restrict us to the case

$$H \simeq h. \tag{4.29}$$

The construction of the prolongation operator follows Theorem 4.2. We split the dual space trivially into

$$Q_h = Q_T.$$

The space

$$V_T = \{(0,\beta_h) \in V_h : \beta_h|_T \in B_3\}$$

is enough to obtain stability (see Section 2.4.3). By the trivial choice $\widetilde{Q}_H = \{0\}$, assumptions (4.17) and (4.19) are clearly satisfied.

The prolongation is implemented very cheap by just performing one step of Point-Jacobi iteration on the subspace of β-bubbles. The bubble-space splits $A_h(.,.)$ orthogonal into two-dimensional subspaces inside each element on the fine grid. Thus, the Jacobi step becomes and exact solver.

The interpolation is the Fortin operator I_H^F. We can chose a nodal interpolation operator $I_H^{F,1}$ for the corner nodes, because we had to assume $H \simeq h$. The second step, $I_H^{F,2}$ adjusts bubble functions for elements with $H > t$. The fine grid complement

$$(w_f,\beta_f) = (w_h,\beta_h) - R_H^V I_H^F(w_h,\beta_h)$$

has to be split locally. It vanishes at corner nodes, thus the Bramble Hilbert lemma gives

$$\|\beta_f\|_0 \preceq h\,\|\beta_f\|_1,$$

and

$$\begin{aligned}\|w_f\|_0 &\preceq h\,\|\nabla w_f\|_0 \le h\,\{\|\nabla w_f - \beta_f\|_0 + \|\beta_f\|_0\}\\ &\preceq h\,\|\nabla w_f - \beta_f\|_0 + h^2\,\|\beta_f\|_1\\ &\preceq h(h+t)\|(w_f,\beta_f)\|_{V_h}.\end{aligned}$$

We define the local norm

$$\|(w_h,\beta_h)\|^2_{V_h,loc} := \frac{1}{h^2}\|\beta_h\|_0^2 + \frac{1}{h^2(h+t)^2}\|w_h\|_0^2. \tag{4.30}$$

Combining both estimates above, we obtain that

$$\|(w_f,\beta_f)\|^2_{V_h,loc} \preceq \|(w_f,\beta_f)\|^2_{A_h}. \tag{4.31}$$

By continuity in energy of the interpolation I_H^F and of the prolongation R_H^V, there holds

$$\|(w_f,\beta_f)\|_{A_h} \preceq \|(w_h,\beta_h)\|_{A_h}. \tag{4.32}$$

All information we will need is contained now in

$$\|(w_f,\beta_f)\|^2_{V_h,loc} + t^{-2}\|\overline{\nabla w_f - \beta_f}^h\|_0^2 \preceq \|(w_h,\beta_h)\|^2_{A_h}. \tag{4.33}$$

This is the approximation property for the two-level method.

We will estimate

$$\|(w_f,\beta_f)\|^2_{D_h} = \inf_{(w_f,\beta_f)=\sum(w_i,\beta_i)} \|(w_i,\beta_i)\|^2_{A_h} \preceq \|(w_f,\beta_f)\|^2_{V_h,loc} + t^{-2}\|\overline{\nabla w_f - \beta_f}^h\|_0^2$$

Let $I_h^{F,2}$ be the operator for adjusting β_h-bubbles for elements with $h > t$, i.e.

$$c_h(\Lambda_h I_h^{F,2}(w_h,\beta_h), q_h) = c_h(\Lambda_h(w_h,\beta_h), q_h) \qquad \forall\, q_h \in Q_{h,h>t}.$$

It is continuous with respect to the local norm (4.30). The co-projector $I - I_h^{F,2}$ maps into the (reduced) kernel, thus there holds

$$\|(w_h,\beta_h) - I_h^{F,2}(w_h,\beta_h)\|^2_{A_h} \preceq \|(w_h,\beta_h) - I_h^{F,2}(w_h,\beta_h)\|^2_{V_h} \preceq \|(w_h,\beta_h)\|^2_{V_h,loc}. \tag{4.34}$$

We split $(w_f,\beta_f) = (w_0,\beta_0) + (w_1,\beta_1)$ by

$$(w_0,\beta_0) = (w_f,\beta_f) - I_h^{F,2}(w_f,\beta_f) \qquad \text{and} \qquad (w_1,\beta_1) = I_h^{F,2}(w_f,\beta_f).$$

into a kernel function and into the rest. The splitting is stable in $V_{h,loc}$-norm. The following identity using a partitioning of unity $\{\psi_i\}$ adjusted to the Scott-Zhang projector $I_h^{F,1}$

$$(w_0,\beta_0) = (I - I_h^{F,2})I_h^{F,1}(\sum \psi_i)(w_0,\beta_0)$$

leads to the splitting of (w_0, β_0):

$$(w_{0,i}, \beta_{0,i}) = (I - I_h^{F,2}) I_h^{F,1} \psi_i(w_0, \beta_0).$$

Combining (4.34) and estimates related to the partitioning of unity we get

$$\begin{aligned} \|(w_0, \beta_0)\|_D^2 &\preceq \sum \|(w_{0,i}, \beta_{0,i})\|_{A_h}^2 \\ &= \sum \|(I - I_h^{F,2}) I_h^{F,1} \psi_i(w_0, \beta_0)\|_{A_h}^2 \\ &\preceq \sum \|I_h^{F,1} \psi_i(w_0, \beta_0)\|_{V_h,loc}^2 \\ &\preceq \sum \|(w_0, \beta_0)\|_{V_h,loc,\omega_i}^2 \\ &\preceq \|(w_0, \beta_0)\|_{V_h,loc}^2. \end{aligned} \tag{4.35}$$

The other component, (w_1, β_1), consists only of β-bubbles on elements with $h > t$, thus we have

$$\|(w_1, \beta_1)\|_D^2 \simeq t^{-2} \|\beta_1\|_0^2 \leq t^{-2} \|\overline{\nabla w_f - \beta_f}^h\|_0^2.$$

Combining the last two estimates with (4.33), we obtain the result

$$\|(w_f, \beta_f)\|_{D_h}^2 \preceq \|(w_0, \beta_0)\|_{D_h}^2 + \|(w_1, \beta_1)\|_{D_h}^2 \leq \|(w_h, \beta_h)\|_{A_h}.$$

Thus, the two level preconditioner is robust with respect to h and t.

4.3 Multigrid Methods

In this section we formulate multigrid methods for parameter dependent problems and give an outline of the analysis. The proofs of the approximation property and of the smoothing properties will be postponed to the next sections.

The multigrid formulation is based on a sequence of possibly non-nested non-conforming finite element spaces

$$V_1, V_2, \ldots, V_L$$

and on a chain of nested spaces

$$Q_1 \subset Q_2 \subset \ldots Q_L \subset Q$$

with norms $\|.\|_{V_l}$ and $\|.\|_{Q_l}$. The norms may and will depend on the level as well as on the parameter. We set $X_l = V_l \times Q_l$ with the product norm $\|(u,p)\|_{X_l} := (\|u\|_{V_l}^2 + \|p\|_{Q_l}^2)^{1/2}$. The index l denotes the level and takes the values 1 to L, unless explicitely defined differently. On each level l we need the symmetric and positive definite bilinear form in primal variables

$$A_l(u_l, v_l) = a_l(u_l, v_l) + \varepsilon^{-1} c_l(\Lambda_l u_l, \Lambda_l v_l). \tag{4.36}$$

It is related to the symmetric bilinear form in mixed variables

$$B_l((u_l, p_l), (v_l, q_l)) = a_l(u_l, v_l) + c_l(\Lambda_l u_l, q_l) + c_l(\Lambda_l v_l, p_l) - \varepsilon\, c_l(p_l, q_l) \tag{4.37}$$

defined on $X_l \times X_l$. The bilinear form $B_l(.,.)$ shall be consistent with the form $B(.,.) : X \times X \to \mathbb{R}$ in the sense of

$$B_l((u, \varepsilon^{-1}\Lambda u), (v, q)) = B((u, \varepsilon^{-1}\Lambda u), (v, q)) \qquad \forall\, u \in V\ \forall\, (v, q) \in X.$$

We assume that $a_l(.,.)$ is symmetric and non-negative, i.e.

$$a_l(u_l, u_l) \geq 0 \qquad \forall\, u_l \in V_l. \tag{4.38}$$

The form $c_l(.,.)$ is assumed to be symmetric and positive, i.e.

$$c_l(p_l, p_l) > 0 \qquad \forall\, 0 \neq p_l \in Q_l. \tag{4.39}$$

The bilinear form $B_l(.,.)$ is assumed to be continuous

$$B_l((u_l, p_l), (v_l, q_l)) \preceq \|(u_l, p_l)\|_{X_l} \|(v_l, q_l)\|_{X_l} \tag{4.40}$$

and stable

$$\sup_{(v_l, q_l) \in X_l} \frac{B_l((u_l, p_l), (v_l, q_l))}{\|(v_l, q_l)\|_{X_l}} \succeq \|(u_l, p_l)\|_{X_l} \qquad \forall\, (u_l, p_l) \in X_l. \tag{4.41}$$

Let $V^+ \subset V$ and $Q^+ \subset Q$ be subspaces with stronger norms $\|.\|_{V^+}$ and $\|.\|_{Q^+}$, respectively. Let $V^- \supset V$ be a larger space with norm $\|.\|_{V^-}$. Its dual is $(V^-)^*$ with norm $\|.\|_{(V^-)^*}$. We assume that $V_l \subset V^-$. The continuous variational problem: Find $(u, p) \in X$ such that

$$B((u, p), (v, q)) = f(v) \qquad \forall\, (v, q) \in X \tag{4.42}$$

is assumed to fulfill the regularity shift

$$\|u\|_{V^+} + \|p\|_{Q^+} \preceq \|f\|_{(V^-)^*}. \tag{4.43}$$

We require interpolation operators

$$I_l^X = (I_l^V, I_l^Q) : (V^+ \times Q^+) \to X_l \tag{4.44}$$

fulfilling the approximation estimates

$$\hbar_l^{-1} \|v - I_l^V v\|_{V^-} + \|(v, q) - I_l^X (v, q)\|_{X_l} \preceq \hbar_l \left(\|v\|_{V^+} + \|q\|_{Q^+}\right) \tag{4.45}$$

for given $(v, q) \in V^+ \times Q^+$. The coefficient $\hbar_l$ is related to the mesh size on level l. Additionally, on each space V_l the inverse estimate

$$\|u_l\|_{V_l} \preceq \hbar_l^{-1} \|u_l\|_{V^-} \tag{4.46}$$

shall be fulfilled with the same coefficient $\hbar_l$. Let $(u_l, p_l) \in X_l$ be the solution of the mixed finite element problem

$$B_l((u_l, p_l), (v_l, q_l)) = f(v_l) \qquad \forall\, (v_l, q_l) \in X_l. \tag{4.47}$$

Then, by Theorem 2.15, the a priori error estimate to the solution (u,p) of (4.42)

$$\hbar_l^{-1}\|u-u_l\|_{V^-}+\|u-u_l\|_{V_l}+\|p-p_l\|_{Q_l}\preceq \hbar_l\,\|f\|_{(V^-)^*} \tag{4.48}$$

is valid.

As in Section 4.2 we define the trivial prolongation operators

$$R_l^{V,0}:V_{l-1}\to V_l, \tag{4.49}$$

with $2\le l\le L$. They are supposed to fulfill the continuity and approximation estimates

$$\begin{aligned} \|R_l^{V,0}u_{l-1}\|_{V_l}+\hbar_l^{-1}\,\|u_{l-1}-R_l^{V,0}u_{l-1}\|_{V^-} &\preceq \|u_{l-1}\|_{V_{l-1}},\\ \|I_l^V u-R_l^{V,0}I_{l-1}^V u\|_{V^-} &\preceq \hbar_l^2\,\|u\|_{V^+} \end{aligned} \tag{4.50}$$

for all $u_{l-1}\in V_{l-1}$ and $u\in V^+$. Since $Q_{l-1}\subset Q_l$, we do not need explicite prolongation operators for the dual variable. We combine the $R_l^{V,0}$ with the identity to the trivial mixed prolongation operator

$$R_l^{X,0}=(R_l^{V,0},id):X_{l-1}\to X_l.$$

We assume norm equivalence across two levels

$$\|p_{l-1}\|_{Q_{l-1}}\simeq\|p_{l-1}\|_{Q_l}\qquad\forall\,p_{l-1}\in Q_{l-1}. \tag{4.51}$$

The actual grid transfer operator $R_l^{V,0}$ requires the space

$$X_{l,T}=V_{l,T}\times Q_{l,T}\subset X_l \tag{4.52}$$

with the properties defined below. We have to solve a variational problem with the bilinear form $A_l(.,.)$ on the space $V_{l,T}$. To be efficient it is necessary that $V_{l,T}$ splits into subspaces of small dimension which are orthogonal with respect to $A_l(.,.)$. The space Q_l is decomposed as

$$Q_l=Q_{l,T}\oplus\widetilde{Q}_{l-1} \tag{4.53}$$

with $\widetilde{Q}_{l-1}\subset Q_{l-1}$. The decomposition is assumed to be orthogonal with respect to the inner product $c_l(.,.)$. The space $\widetilde{Q}_{l-1}$ corresponds to coarse grid constraints which will be inherited on the fine grid. The complement $Q_{l,T}$ characterizes the constraints which can be fulfilled by local projections. We define the c_l ortho-projector $I_{l-1}^{\widetilde{Q}}:Q_l\to\widetilde{Q}_{l-1}$. It is assumed to be continuous with norm $\|.\|_{Q_l\to Q_{l-1}}$. The bilinear-from $B_l(.,.)$ is assumed to be stable on $X_{l,T}$, namely for all $(u,p)\in X_{l,T}$ there holds

$$\sup_{(v,q)\in X_{l,T}}\frac{B_l((u,p),(v,q))}{\|(v,q)\|_{X_l}}\succeq\|(u,p)\|_{X_l}. \tag{4.54}$$

We assume the inclusion

$$\Lambda_l V_{l,T}\subset Q_{l,T}. \tag{4.55}$$

The inclusion ensures that corrections in $V_{l,T}$ will not disturb constraints corresponding to $\widetilde{Q}_{l-1}$. The trivial prolongation is supposed to satisfy

$$B_{l-1}((u_{l-1},p_{l-1}),(0,\tilde{q}_{l-1})) = B_l((R_l^{V,0}u_{l-1},p_{l-1}),(0,\tilde{q}_{l-1})) \tag{4.56}$$

for all $(u_{l-1},p_{l-1}) \in X_{l-1}$ and $\tilde{q}_{l-1} \in \widetilde{Q}_{l-1}$. This condition guarantees that constraints corresponding to $\widetilde{Q}_{l-1}$ are preserved by the grid transfer. On the space $V_{l,T}$ not only the inverse estimate (4.46), but also the norm equivalence

$$\|u_{l,T}\|_{V^-} \simeq \hbar_l \, \|u_{l,T}\|_{V_l} \qquad \forall \, u_{l,T} \in V_{l,T} \tag{4.57}$$

shall be fulfilled. For $2 \leq l \leq L$, we define the projection

$$P_{l,T}^A : V_l \to V_{l,T} : \qquad A_l(P_{l,T}^A u_l, v_{l,T}) = A_l(u_l, v_{l,T}) \qquad \forall \, u_l \in V_l, \; \forall v_{l,T} \in V_{l,T}. \tag{4.58}$$

Then the actual grid transfer operator is defined as

$$R_l^V := (I - P_{l,T}^A) R_l^{V,0} \tag{4.59}$$

The construction of the smoother needs the decomposition

$$V_l = \sum_{i=1}^{M_l} V_{l,i}, \qquad Q_l = \sum_{i=1}^{M_l} Q_{l,i}. \tag{4.60}$$

into subspaces of small dimension. The decomposition may be overlapping. The inclusions

$$\Lambda_l V_{l,i} \subset Q_{l,i} \tag{4.61}$$

must be fulfilled for all $1 \leq i \leq M_l$. Let $\widetilde{a}_l(.,.) : V_l \times V_l \to \mathbb{R}$ be symmetric bilinear forms such that

$$\widetilde{a}_l(u_l, u_l) \geq a_l(u_l, u_l) \qquad \forall \, u_l \in V_l. \tag{4.62}$$

We define the parameter dependent primal form

$$\widetilde{A}_l(u_l, v_l) := \widetilde{a}_l(u_l, v_l) + \varepsilon^{-1} \, c_l(\Lambda_l u_l, \Lambda_l v_l). \tag{4.63}$$

The smoother is an additive Schwarz method with subspace problems derived from the bilinear form $\widetilde{A}_l(.,.)$. That is

$$S_l := I - \tau \sum_{i=1}^{M_l} T_{l,i}^A, \tag{4.64}$$

where the operator $T_{l,i}^A$ is defined by

$$T_{l,i}^A : V_l \to V_{l,i} : \qquad \widetilde{A}_l(T_{l,i}^A u_l, v_{l,i}) = A_l(u_l, v_{l,i}) \qquad \forall \, u_l \in V_l, \; \forall \, v_{l,i} \in V_{l,i}. \tag{4.65}$$

The smoother is an iteration with symmetric preconditioner D_l, i.e.

$$S_l = I - \tau D_l^{-1} A_l.$$

By the Additive Schwarz lemma, the norm induced by the preconditioner can be represented by

$$\|u_l\|_{D_l}^2 = \inf_{\substack{u_l=\sum u_{l,i} \\ u_{l,i}\in V_{l,i}}} \sum_{i=1}^{M_l} \|u_{l,i}\|_{\widetilde{A}_l}^2. \tag{4.66}$$

We will prove the approximation property for the norm

$$\|u_l\|_{V_l,\tilde{0}}^2 := \hbar_l^{-2}\|u\|_{V^-}^2 + \varepsilon^{-1}\|\Lambda_l u_l\|_{c_l}^2 + \varepsilon^{-2}\|I_{l-1}^{\widetilde{Q}}\Lambda_l u_l\|_{Q_l}^2. \tag{4.67}$$

The norm combines three different terms. The first one is the improved convergence in a weaker norm stemming from the usual duality argument. The second term is essential to reflect continuity in energy norm. The last term bounds the coarse grid part of the dual variable $\varepsilon^{-1}\Lambda_l u_l$.

In Section 4.5 and Section 4.6 we will establish the estimate

$$\|u_l\|_{[A_l,D_l]_\alpha} \preceq \|u_l\|_{V_l,\tilde{0}} \qquad \forall\, u_l \in V_l \tag{4.68}$$

for two different types of smoothers for the considered problems. These components lead to optimal and robust multigrid solvers:

Theorem 4.4 (Robust multigrid method). *Define a multigrid method with the components introduced above. Then a W-cycle scheme with sufficiently many smoothing steps lead to an iterative process with contraction number independent of the level l and the parameter ε. A variable V-cycle lead to a preconditioner for A_l with spectral bounds independent of the level l and the parameter ε.*

Proof. According to Theorem 3.7 we have to check 3 conditions: The approximation property

$$\|u_l - R_l^V A_{l-1}^{-1}[R_l^V]^T A_l u_l\|_{V_l,\tilde{0}} \preceq \|u_l\|_{A_l} \tag{4.69}$$

will be verified in Section 4.4. The local preconditioner $\tau^{-1}D_l$ of the smoother is scaled such that

$$A_l \leq \tau^{-1} D_l.$$

Estimate (4.68) will be verified in Section 4.5 and Section 4.6 for two different type of smoothers. All estimates are independent of the level l and the parameter ε. Together, these conditions prove the theorem. □

4.3.1 Multigrid methods in mixed variables

Before we go into the details of the multigrid analysis, we derive an equivalent multigrid procedure in mixed variables. Let the operator $B_l : X_l \to X_l$ be defined by an inner product on X_l. According to (4.63), we define the inexact mixed form $\widetilde{B}_l(.,.) : X_l \times X_l \to \mathbb{R}$ by

$$\widetilde{B}_l((u_l,p_l),(v_l,q_l)) := a_l(u_l,v_l) + c_l(\Lambda_l u_l, q_l) + c_l(\Lambda_l v_l, p_l) - \varepsilon\, c_l(p_l,q_l). \tag{4.70}$$

The projection $P_{l,T}^B : X_l \to X_{l,T}$ and the operators $T_{l,i}^B : X_l \to X_{l,i}$ are defined by

$$B_l(P_{l,T}^B(u_l,p_l),(v_{l,T},q_{l,T})) = B_l((u_l,p_l),(v_{l,T},q_{l,T})) \quad \forall\,(v_{l,T},q_{l,T}) \in X_{l,T} \tag{4.71}$$

and

$$\widetilde{B}_l(T_{l,i}^B(u_l,p_l),(v_{l,i},q_{l,i})) = B_l((u_l,p_l),(v_{l,i},q_{l,i})) \quad \forall\,(v_{l,i},q_{l,i}) \in X_{l,i} \tag{4.72}$$

for all $(u_l,p_l) \in X_l$, respectively. Then the smoother and the prolongation operator are

$$\begin{aligned} S_l^B &:= I - \tau \sum T_{l,i}^B, \\ R_l^X &:= (I - P_{l,T}^B) R_l^{X,0}. \end{aligned}$$

The smoother requires the solution of local saddle point problems. That is like a Vanka - smoother [Van86]. But usual versions of the family of Vanka smoothers do not fulfill our assumption (4.61). The multigrid method for the mixed problem is defined corresponding to Algorithm 1.

We define the subspaces $X_{l,0} \subset X_l$

$$X_{l,0} = \{(u_l,p_l) \in X_l : \Lambda_l u_l = \varepsilon p_l\}. \tag{4.73}$$

For all $(u_l,p_l) \in X_{l,0}$ there holds

$$B_l((u_l,p_l),(0,q_l)) = 0 \qquad \forall\, q_l \in Q_l \tag{4.74}$$

as well as

$$A_l(u_l,v_l) = B_l((u_l,p_l),(v_l,0)) \qquad \forall\, v_l \in V_l, \tag{4.75}$$

and corresponding relations for $\widetilde{A}_l(.,.)$ and $\widetilde{B}_l(.,.)$. The multigrid components in mixed variables preserve the spaces $X_{l,0}$. On the subspace, they reduce to the components in primal variables. This is collected in the following lemma:

Lemma 4.5. *The multigrid components fulfill the following properties:*

1. *The smoother S_l^B preserves $X_{l,0}$. On the subspace it is equivalent to the smoother S_l^A in primal variables. This means for $(u_l,p_l) \in X_{l,0}$ and*

   $$(\hat{u}_l,\hat{p}_l) = S_l^B(u_l,p_l)$$

   there holds $(\hat{u}_l,\hat{p}_l) \in X_{l,0}$ and

   $$\hat{u}_l = S_l^A u_l.$$

2. *The prolongation R_l^X maps $X_{l-1,0}$ into $X_{l,0}$. On the subspace it is equivalent to the prolongation R_l^V in primal variables. This means for $(u_{l-1},p_{l-1}) \in X_{l-1,0}$ and*

   $$(\hat{u}_l,\hat{p}_l) = R_l^X(u_{l-1},p_{l-1})$$

   there holds $(\hat{u}_l,\hat{p}_l) \in X_{l,0}$ and

   $$\hat{u}_l = R_l^V u_{l-1}.$$

3. *The coarse grid solution operator maps $X_{l,0}$ into $X_{l-1,0}$. On the subspace it is equivalent to the coarse grid solution operator in primal variables. This means for $(u_l, p_l) \in X_{l,0}$ and*
$$(\hat{u}_{l-1}, \hat{p}_{l-1}) = B_{l-1}^{-1}[R_l^X]^T B_l\,(u_l, p_l)$$
there holds $(\hat{u}_{l-1}, \hat{p}_{l-1}) \in X_{l-1,0}$ and
$$\hat{u}_{l-1} = A_{l-1}^{-1}[R_l^A]^T A_l u_l.$$

Proof. 1. It is sufficient to prove the corresponding properties for the operators $T_{l,i}^B$. Let $(u_l, p_l) \in X_{l,0}$ and $(\hat{u}_l, \hat{p}_l) = T_{l,i}^B(u_l, p_l)$. From (4.61), $\Lambda_l V_{l,i} \subset Q_{l,i}$, there follows $\Lambda_l \hat{u} - \varepsilon \hat{p} \in Q_{l,i}$. Using the definition (4.72) of the operator $T_{l,i}^B$, we get

$$\|\Lambda_l \hat{u}_l - \varepsilon \hat{p}_l\|_{c_l} = \sup_{q_{l,i} \in Q_{l,i}} \frac{(\Lambda_l \hat{u}_l - \varepsilon \hat{p}_l, q_{l,i})_{c_l}}{\|q_{l,i}\|_{c_l}} = \sup_{q_{l,i} \in Q_{l,i}} \frac{(\Lambda_l u_l - \varepsilon p_l, q_{l,i})_{c_l}}{\|q_{l,i}\|_{c_l}} = 0,$$

i. e. $(\hat{u}_l, \hat{p}_l) \in X_{l,0}$. By $(\hat{u}_l, \hat{p}_l) \in X_{l,0} \cap X_{l,i}$ and (4.75) we obtain

$$\widetilde{A}_l(\hat{u}_l, v_{l,i}) = \widetilde{B}_l((\hat{u}_l, \hat{p}_l), (v_{l,i}, 0)) = B_l((u_l, p_l), (v_{l,i}, 0)) = A_l(u_l, v_{l,i}) \qquad \forall\, v_{l,i} \in V_{l,i},$$

what is the definition of $T_{l,i}^A$.

2. Let $(u_{l-1}, p_{l-1}) \in X_{l-1,0}$. We define

$$\begin{aligned} (\tilde{u}_l, \tilde{p}_l) &= R_l^{X,0}(u_{l-1}, p_{l-1}), \\ (u_{l,T}, p_{l,T}) &= P_{l,T}^B(\tilde{u}_l, \tilde{p}_l), \\ (\hat{u}_l, \hat{p}_l) &= (\tilde{u}_l, \tilde{p}_l) - (u_{l,T}, p_{l,T}). \end{aligned}$$

By the definition of the projection $P_{l,T}^B$ we have

$$(\Lambda_l u_{l,T} - \varepsilon p_{l,T}, q_{l,T})_{c_l} = (\Lambda_l \tilde{u}_l - \varepsilon \tilde{p}_l, q_{l,T})_{c_l} \qquad \forall\, q_{l,T} \in Q_{l,T},$$

and by $\Lambda_l u_{l,T} - \varepsilon p_{l,T} \in Q_{l,T} \perp_{c_l} \widetilde{Q}_{l-1}$ we have

$$(\Lambda_l u_{l,T} - \varepsilon p_{l,T}, \tilde{q}_{l-1})_{c_l} = 0 \qquad \forall\, \tilde{q}_{l-1} \in \widetilde{Q}_{l-1}.$$

Thus we get
$$(\Lambda_l \hat{u}_l - \varepsilon \hat{p}_l, q_l)_{c_l} = (\Lambda_l \tilde{u}_l - \varepsilon \tilde{p}_l, I_{l-1}^{\widetilde{Q}} q_l)_{c_l} \qquad \forall\, q_l \in Q_l.$$

We use assumption (4.56), $I_{l-1}^{\widetilde{Q}} q_l \in Q_{l-1}$, and (4.74) to show

$$(\Lambda_l \hat{u}_l - \varepsilon \hat{p}_l, q_l)_{c_l} = B_l((\tilde{u}_l, \tilde{p}_l), (0, I_{l-1}^{\widetilde{Q}} q_l)) = B_{l-1}((u_{l-1}, p_{l-1}), (0, I_{l-1}^{\widetilde{Q}} q_l)) = 0,$$

i.e. $(\hat{u}_l, \hat{p}_l) \in X_{l,0}$. We apply (4.75) to $(\hat{u}_l, \hat{p}_l)$, namely

$$A_l(\hat{u}_l, v_l) = B_l((\hat{u}_l, \hat{p}_l), (v_l, 0)) \qquad \forall\, v_l \in V_l.$$

For $v_{l,T} \in V_{l,T}$ this gives

$$A_l(\tilde{u}_l - u_{l,T}, v_{l,T}) = 0,$$

i.e. $u_{l,T} = P^A_{l,T}\tilde{u}_l$, and

$$\hat{u}_l = (I - P^A_{l,T})\tilde{u}_l = (I - P^A_{l,T})R_l^{V,0} u_{l-1},$$

and the equivalence is proved.

3. Let $(u_l, p_l) \in X_l$. The definition of $(\hat{u}_{l-1}, \hat{p}_{l-1})$ is equivalent to find $(\hat{u}_{l-1}, \hat{p}_{l-1}) \in X_{l-1}$ such that

$$B_{l-1}((\hat{u}_{l-1}, \hat{p}_{l-1}), (v_{l-1}, q_{l-1})) = B_l((u_l, p_l), R_l^{X,0}(v_{l-1}, q_{l-1})) \qquad \forall\, (v_{l-1}, q_{l-1}) \in X_{l-1}.$$

We set $(v_{l-1}, q_{l-1}) = (0, q_{l-1})$ and obtain $(\hat{u}_{l-1}, \hat{p}_{l-1}) \in X_{l-1,0}$. Thus we can apply (4.75) on both sides to get

$$A_{l-1}(\hat{u}_{l-1}, v_{l-1}) = A_l(u_l, R_l^{V,0} v_{l-1}) \qquad \forall\, v_{l-1} \in V_{l-1},$$

what is equivalent to $\hat{u}_{l-1} = A_{l-1}^{-1}[R_l^{V,0}]^T A_l u_l$. □

Theorem 4.6 (Equivalence of algorithms). *The multigrid algorithm in mixed variables preserves the space $X_{l,0}$. On this subspace it is equivalent to the multigrid algorithm in primal variables. This means for $(u_l, p_l) \in X_{l,0}$ and $(\hat{u}_l, \hat{p}_l) = M_l^B(u_l, p_l)$ there holds $(\hat{u}_l, \hat{p}_l) \in X_{l,0}$ and*

$$\hat{u}_l = M_l^A u_l.$$

Proof. The multigrid operator M_l^A fulfills the recursion

$$\begin{aligned} M_1^A &= 0, \\ M_l^A &= (S_l^A)^{m_l} \left(I - R_l^V (I - (M_{l-1}^A)^q) A_{l-1}^{-1} [R_l^V]^T A_l\right) (S_l^A)^{m_l}, \end{aligned}$$

and the mixed operator M_l^B fulfills a corresponding one. We apply Lemma 4.5 and the theorem is proved by induction. □

We define the norm

$$\|(u_l, p_l)\|_{B_l}^2 := a_l(u_l, u_l) + \varepsilon c_l(p_l, p_l). \tag{4.76}$$

On the space $X_{l,0}$ it is identic to the primal energy norm $\|.\|_{A_l}$, i.e.

$$\|u_l\|_{A_l} = \|(u_l, p_l)\|_{B_l} \qquad \forall\, (u_l, p_l) \in X_{l,0}. \tag{4.77}$$

For $\varepsilon = 0$ the norm $\|.\|_{B_l}$ degenerates to a semi-norm.

Theorem 4.7 (Optimal convergence rate of multigrid in mixed variables). *Let $\varepsilon > 0$. The initial error is assumed to be in $X_{l,0}$. Perform either a W-cycle with sufficiently many smoothing steps or a variable V-cycle with sufficient damping for the system in mixed variables. Then the iterates converge with rates uniformly in L and ε with respect to the norm $\|.\|_{B_L}$.*

Assume additionally $B_{l,\varepsilon}(.,.) \to B_{l,\varepsilon=0}(.,.)$, and the limit form $B_{l,\varepsilon=0}(.,.)$ is stable with respect to a limit norm $\|.\|_{X_{l,\varepsilon=0}}$ on the subspaces defined above. Then the algorithm in mixed variables can be performed for $\varepsilon = 0$ and the primal variable converges with respect to the norm $\|.\|_{a_l}$ uniformly in L and ε.

Proof. First, we assume $\varepsilon > 0$. The errors of the iterates stay in $X_{l,0}$. Thus, the algorithm is equivalent to the corresponding one in primal variables. The last theorem stated uniform convergence rates with respect to the primal energy norm. Due to (4.77), this is equivalent to convergence in the norm $\|.\|_{B_l}$.

If the additional assumption is fulfilled, we can apply continuity arguments and pass to the limit $\varepsilon = 0$. □

4.4 The Approximation Property

In this section we will prove the approximation property (4.69) for parameter dependent problems. The coarse grid correction step

$$u_5 := u_1 - R_l^V A_{l-1}^{-1} [R_l^V]^T A_l u_1 \tag{4.78}$$

in primal variables is equivalent to the coarse grid correction in mixed variables. We set $p_1 = \varepsilon^{-1} \Lambda_l u_1$, and Lemma 4.5 provides

$$(u_5, p_5) = (I - R_l^X B_{l-1}^{-1} [R_l^X]^T B_l)(u_1, p_1). \tag{4.79}$$

On X_l, we define the local norm in mixed variables as

$$\|(u_l, p_l)\|^2_{X_l,\tilde{0}} = \hbar_l^{-2} \|u_l\|^2_{V^-} + \varepsilon \|p_l\|^2_{c_l} + \|I_{l-1}^{\widetilde{Q}} p_l\|^2_{Q_l}.$$

On the subspace $X_{l,0}$ it reduces to $\|.\|_{V_l,\tilde{0}}$, i.e.

$$\|(u_l, p_l)\|_{X_l,\tilde{0}} = \|u_l\|_{V_l,\tilde{0}} \qquad \forall\, (u_l, p_l) \in X_{l,0}.$$

The property to be proven is

$$\|u_5\|_{V_l,\tilde{0}} \preceq \|u_1\|_{A_l}, \tag{4.80}$$

which reads in mixed variables as

$$\|(u_5, p_5)\|_{X_l,\tilde{0}} \preceq \|(u_1, p_1)\|_{B_l}.$$

We recall the definition of R_l^X and evaluate its adjoint

$$\begin{aligned} R_l^X &= (I - P_{l,T}^B) R_l^{X,0}, \\ [R_l^X]^T &= [R_l^{X,0}]^T B_l (I - P_{l,T}^B) B_l^{-1}. \end{aligned}$$

The operation (4.79) is split into the steps

$$\begin{aligned} (u_2, p_2) &= (I - P_{l,T}^B)(u_1, p_1), && (4.81) \\ (u_3, p_3) &= B_{l-1}^{-1} [R_l^{X,0}]^T B_l (u_2, p_2), && (4.82) \\ (u_4, p_4) &= (I - P_{l,T}^B) R_l^{X,0} (u_3, p_3), && (4.83) \\ (u_5, p_5) &= (u_1, p_1) - (u_4, p_4). \end{aligned}$$

The first step is called preprocessing step. The third step (4.83) is similar, it is called post-processing. Both modify the functions by local operations, such that the classical coarse grid correction (4.82) becomes efficient.

We apply the triangle inequality in the form

$$\begin{aligned} \|(u_5, p_5)\|_{X_l, \tilde{0}} &= \|(u_1, p_1) - (u_4, p_4)\|_{X_l, \tilde{0}} \\ &\leq \|(u_1, p_1) - (u_2, p_2)\|_{X_l, \tilde{0}} + \|(u_2, p_2) - R_l^{X,0}(u_3, p_3)\|_{X_l, \tilde{0}} + \\ &\quad \|R_l^{X,0}(u_3, p_3) - (u_4, p_4)\|_{X_l, \tilde{0}} \end{aligned}$$

Each one of the three lemmas below estimates one of the three terms. We start with the preprocessing step.

Lemma 4.8. *There holds the approximation estimate*

$$\|(u_2, p_2) - (u_1, p_1)\|_{X_l, \tilde{0}} \preceq \|(u_1, p_1)\|_{B_l}.$$

Additionally, the stability estimate

$$\|u_2\|_{V_l} + \|p_2 - I_{l-1}^{\tilde{Q}} p_2\|_{Q_l} \preceq \|(u_1, p_1)\|_{B_l}.$$

is fulfilled.

Proof. We use $(u_2, p_2) \in X_{l,0}$, and thus

$$\|(u_2, p_2)\|_{B_l} = \|u_2\|_{A_l} = \|(I - P_{l,T}^A) u_1\|_{A_l} \leq \|u_1\|_{A_l} = \|(u_1, p_1)\|_{B_l}.$$

This bounds the first term of the stability estimate. Next, we will establish the estimate

$$\|w\|_{V_l} \preceq \|w\|_{A_l}. \tag{4.84}$$

We use stability (4.41) and continuity (4.40) of $B_l(.,.)$ on X_l to obtain

$$\begin{aligned} \|w_l\|_{V_l} &\preceq \sup_{(v_l,q_l)\in X_l} \frac{B_l((w_l,0),(v_l,q_l))}{\|(v_l,q_l)\|_{X_l}} = \sup_{(v_l,q_l)\in X_l} \frac{a_l(w_l,v_l) - c_l(\Lambda_l w_l, q_l)}{\|(v_l,q_l)\|_{X_l}} \\ &\preceq \sup_{(v_l,q_l)\in X_l} \frac{\left[a_l(w_l,w_l) + \varepsilon^{-1} c_l(\Lambda_l w_l, \Lambda_l w_l)\right]^{1/2} \left[a_l(v_l,v_l) + \varepsilon\, c_l(q_l,q_l)\right]^{1/2}}{\|(v_l,q_l)\|_{X_l}} \\ &= \sup_{(v_l,q_l)\in X_l} \frac{A_l(w_l,w_l)^{1/2}\, B_l((v_l,q_l),(v_l,-q_l))^{1/2}}{\|(v_l,q_l)\|_{X_l}} \preceq \|w_l\|_{A_l}. \end{aligned}$$

For $(u_l,p_l) \in X_{l,0}$, inequality (4.84) can be rewritten as

$$\|u_l\|_{V_l} \preceq \|(u_l,p_l)\|_{B_l}.$$

To bound the approximation term we use $(u_2 - u_1, p_2 - p_1) \in X_{l,T} \cap X_{l,0}$, the norm equivalence (4.57), the orthogonal decomposition (4.53), and the estimate above to obtain

$$\begin{aligned} \|(u_2,p_2) - (u_1,p_1)\|^2_{X_l,\tilde{0}} &= \hbar_l^{-2} \|u_2 - u_1\|^2_{V^-} + \varepsilon\, \|p_2 - p_1\|^2_{c_l} + \|I^{\tilde{Q}}_{l-1}(p_2 - p_1)\|^2_{Q_l} \\ &\preceq \|u_2 - u_1\|^2_{V_l} + \varepsilon\, \|p_2 - p_1\|^2_{c_l} \preceq \|(u_2 - u_1, p_2 - p_1)\|^2_{B_l} \preceq \|(u_1,p_1)\|^2_{B_l}. \end{aligned}$$

From $p_2 - I^{\tilde{Q}}_{l-1} p_2 \in Q_{l,T}$, stability (4.54) of $X_{l,T}$, inclusion (4.55), orthogonality (4.53), and the definition of $P^B_{l,T}$ we obtain

$$\begin{aligned} \|p_2 - I^Q_{l-1} p_2\|_{Q_l} &\preceq \sup_{(v,q)\in X_{l,T}} \frac{B_l((0, p_2 - I^{\tilde{Q}}_{l-1} p_2),(v,q))}{\|(v,q)\|_{X_l}} \\ &= \sup_{(v,q)\in X_{l,T}} \frac{B_l((0,p_2),(v,q)) - c_l(I^{\tilde{Q}}_{l-1} p_2, \Lambda_l v - q)}{\|(v,q)\|_{X_l}} = \sup_{(v,q)\in X_{l,T}} \frac{B_l((0,p_2),(v,q))}{\|(v,q)\|_{X_l}} \\ &= \sup_{(v,q)\in X_{l,T}} \frac{B_l((-u_2,0),(v,q))}{\|(v,q)\|_{X_l}} \preceq \|u_2\|_{V_l} \preceq \|(u_2,p_2)\|^2_{B_l}. \end{aligned}$$

□

Lemma 4.9. *The classical coarse grid correction fulfills the approximation inequality*

$$\|R^{X,0}_l(u_3,p_3) - (u_2,p_2)\|_{X_l,\tilde{0}} \preceq \|(u_1,p_1)\|_{B_l}.$$

Additionally, the stability estimate

$$\|u_3\|_{V_{l-1}} + \|p_3 - I^{\tilde{Q}}_{l-1} p_3\|_{Q_l} \preceq \|(u_1,p_1)\|_{B_l}.$$

is fulfilled.

Proof. From the variational definition of $(u_3, p_3) \in X_{l-1}$

$$B_{l-1}((u_3, p_3), (v, q)) = B_l((u_2, p_2), R_l^{X,0}(v, q)) \qquad \forall\, (v, q) \in X_{l-1}, \tag{4.85}$$

and property (4.56) we get

$$B_{l-1}((u_3, p_3 - I_{l-1}^{\widetilde{Q}} p_2), (v, q)) = B_l((u_2, p_2 - I_{l-1}^{\widetilde{Q}} p_2), R_l^{X,0}(v, q)) \qquad \forall\, (v, q) \in X_{l-1}.$$

Stability (4.41) on X_{l-1}, continuity (4.40) of $B_l(.,.)$, continuity (4.50)+(4.51) of the prolongation operator $R_l^{X,0}$, and Lemma 4.8 give

$$\begin{aligned}
\|u_3\|_{V_{l-1}} + \|p_3 - I_{l-1}^{\widetilde{Q}} p_2\|_{Q_{l-1}} &\preceq \sup_{(v,q)\in X_{l-1}} \frac{B_{l-1}((u_3, p_3 - I_{l-1}^{\widetilde{Q}} p_2), (v, q))}{\|(v, q)\|_{X_{l-1}}} \\
&= \sup_{(v,q)\in X_{l-1}} \frac{B_l((u_2, p_2 - I_{l-1}^{\widetilde{Q}} p_2), R_l^{X,0}(v, q))}{\|(v, q)\|_{X_{l-1}}} \\
&\preceq \|u_2\|_{V_l} + \|p_2 - I_{l-1}^{\widetilde{Q}} p_2\|_{Q_l} \preceq \|(u_1, p_1)\|_{B_l}.
\end{aligned}$$

Norm equivalence (4.51) on Q_{l-1} and Q_l, and Lemma 4.8 give

$$\|p_3 - p_2\|_{Q_l} \preceq \|p_3 - I_{l-1}^{\widetilde{Q}} p_2\|_{Q_{l-1}} + \|p_2 - I_{l-1}^{\widetilde{Q}} p_2\|_{Q_l} \preceq \|(u_1, p_1)\|_{B_l}$$

and the estimate

$$\|p_3 - I_{l-1}^{\widetilde{Q}} p_3\|_{Q_l} \leq \|p_3 - p_2 + I_{l-1}^{\widetilde{Q}}(p_2 - p_3)\|_{Q_l} + \|p_2 - I_{l-1}^{\widetilde{Q}} p_2\|_{Q_l} \preceq \|(u_1, p_1)\|_{B_l}.$$

We use continuity (4.40) of $B_l(.,.)$ to bound the term $\varepsilon \|.\|_{c_l}^2$ of the approximation estimate, namely

$$\begin{aligned}
\varepsilon\, \|p_2 - p_3\|_{c_l}^2 &\preceq \varepsilon\, \|p_2 - I_{l-1}^{\widetilde{Q}} p_2\|_{c_l} + \varepsilon\, \|p_3 - I_{l-1}^{\widetilde{Q}} p_2\|_{c_l} \\
&\preceq \|p_2 - I_{l-1}^{\widetilde{Q}} p_2\|_{Q_l} + \|p_3 - I_{l-1}^{\widetilde{Q}} p_2\|_{Q_l} \\
&\preceq \|(u_1, p_1)\|_{B_l}
\end{aligned}$$

The only term left is the improved estimate in the weaker norm, which we will bound by

$$\|u_2 - R_l^{V,0} u_3\|_{V^-} \preceq \hbar_l\, \|(u_2, p_2 - I_{l-1}^{\widetilde{Q}} p_2)\|_{X_l}. \tag{4.86}$$

We define the dual problem

$$B((\varphi, \psi), (v, q)) = (u_2 - R_l^{V,0} u_3, v)_{V^-}.$$

The right hand side is a linear functional in $(V^-)^*$ with norm $\|u_2 - R_l^{V,0} u_3\|_{V^-}$. We define the finite element problems

$$\begin{aligned}
B_l((\varphi_l, \psi_l), (v_l, q_l)) &= (u_2 - R_l^{V,0} u_3, v_l)_{V^-}, &(4.87)\\
B_{l-1}((\varphi_{l-1}, \psi_{l-1}), (v_{l-1}, q_{l-1})) &= (u_2 - R_l^{V,0} u_3, v_{l-1})_{V^-}, &(4.88)\\
B_{l-1}((\widetilde{\varphi}_{l-1}, \widetilde{\psi}_{l-1}), (v_{l-1}, q_{l-1})) &= (u_2 - R_l^{V,0} u_3, R_l^{V,0} v_{l-1})_{V^-}. &(4.89)
\end{aligned}$$

with test functions $(v_l, q_l) \in X_l$ and $(v_{l-1}, q_{l-1}) \in X_{l-1}$.

First, we estimate the difference of the solutions of (4.88) and (4.89) by

$$\begin{aligned} &\|(\varphi_{l-1} - \widetilde{\varphi}_{l-1}, \psi_{l-1} - \widetilde{\psi}_{l-1})\|_{X_{l-1}} \qquad (4.90)\\ &\preceq \sup_{(v_{l-1},q_{l-1})\in X_{l-1}} \frac{B_{l-1}((\varphi_{l-1} - \widetilde{\varphi}_{l-1}, \psi_{l-1} - \widetilde{\psi}_{l-1}), (v_{l-1}, q_{l-1}))}{\|(v_{l-1}, q_{l-1})\|_{X_{l-1}}} \\ &= \sup_{(v_{l-1},q_{l-1})\in X_{l-1}} \frac{(u_2 - R_l^{V,0} u_3, v_{l-1} - R_l^{V,0} v_{l-1})_{V^-}}{\|(v_{l-1}, q_{l-1})\|_{X_{l-1}}} \\ &\le \|u_2 - R_l^{V,0} u_3\|_{V^-} \sup_{v_{l-1}\in V_{l-1}} \frac{\|v_{l-1} - R_l^{V,0} v_{l-1}\|_{V^-}}{\|v_{l-1}\|_{V_{l-1}}} \preceq h\,\|u_2 - R_l^{V,0} u_3\|_{V^-}. \end{aligned}$$

We have used the approximation (4.50). Using the variational problems (4.87) and (4.89), and the variational specification of the coarse grid correction (4.85) we get

$$\begin{aligned} \|u_2 - R_l^{V,0} u_3\|_{V^-}^2 &= (u_2 - R_l^{V,0} u_3, u_2)_{V^-} - (u_2 - R_l^{V,0} u_3, R_l^{V,0} u_3)_{V^-} \\ &= B_l((\varphi_l, \psi_l), (u_2, p_2)) - B_{l-1}((\widetilde{\varphi}_{l-1}, \widetilde{\psi}_{l-1}), (u_3, p_3)) \\ &= B_l((\varphi_l, \psi_l), (u_2, p_2)) - B_l(R_l^{X,0}(\widetilde{\varphi}_{l-1}, \widetilde{\psi}_{l-1}), (u_2, p_2)) \\ &= B_l((\varphi_l, \psi_l) - R_l^{X,0}(\widetilde{\varphi}_l, \widetilde{\psi}_l), (u_2, p_2)). \end{aligned}$$

Next, we use (4.56) and observe that

$$\begin{aligned} &B_l((\varphi_l, \psi_l) - R_l^{X,0}(\widetilde{\varphi}_{l-1}, \widetilde{\psi}_{l-1}), (0, I_{l-1}^{\widetilde{Q}} p_2)) \qquad (4.91)\\ &\quad = B_l((\varphi_l, \psi_l), (0, I_{l-1}^{\widetilde{Q}} p_2)) - B_{l-1}((\widetilde{\varphi}_{l-1}, \widetilde{\psi}_{l-1}), (0, I_{l-1}^{\widetilde{Q}} p_2)) = 0. \end{aligned}$$

We continue and obtain

$$\begin{aligned} \|u_2 - R_l^{V,0} u_3\|_{V^-}^2 &= B_l((\varphi_l, \psi_l) - R_l^{X,0}(\widetilde{\varphi}_l, \widetilde{\psi}_l), (u_2, p_2 - I_{l-1}^{\widetilde{Q}} p_2)) \qquad (4.92)\\ &\le \|(\varphi_l, \psi_l) - R_l^{X,0}(\widetilde{\varphi}_l, \widetilde{\psi}_l)\|_{X_l} \;\; \|(u_2, p_2 - I_{l-1}^{\widetilde{Q}} p_2)\|_{X_l}. \qquad (4.93) \end{aligned}$$

We use the continuity of the prolongation $R_l^{X,0}$, and (4.90) to get

$$\begin{aligned} &\|(\varphi_l, \psi_l) - R_l^{X,0}(\widetilde{\varphi}_l, \widetilde{\psi}_l)\|_{X_l} \\ &\quad\preceq \|(\varphi_l, \psi_l) - R_l^{X,0}(\varphi_{l-1}, \psi_{l-1})\|_{X_l} + \|R_l^{X,0}(\varphi_{l-1}, \psi_{l-1}) - R_l^{X,0}(\widetilde{\varphi}_{l-1}, \widetilde{\psi}_{l-1})\|_{X_l} \\ &\quad\preceq \|\varphi_l - R_l^{V,0}\varphi_{l-1}\|_{V_l} + \|\psi_l - \psi_{l-1}\|_{Q_l} + h\,\|u_2 - R_l^{V,0} u_3\|_{V^-}. \end{aligned}$$

If the spaces V_{l-1} and V_l are nested, and $R_l^{V,0}$ is the embedding operator, we can apply the a priori estimate (4.48) and we are done. We did not use the regularity estimate (4.43) explicitely for this case. If $R_l^{V,0}$ is not the embedding, we need the intermediate steps

$$\|\varphi_l - R_l^{V,0}\varphi_{l-1}\|_{V_l} \le \|\varphi_l - I_l^V \varphi\|_{V_l} + \|I_l^V \varphi - R_l^{V,0} I_{l-1}^V \varphi\|_{V_l} + \|R_l^{V,0} I_{l-1}^V \varphi - R_l^{V,0}\varphi_{l-1}\|_{V_l}.$$

Now we apply the approximation estimate (4.45) of the interpolation operator, the a priori estimate (4.48), the approximation estimate of the prolongation operator (4.50), and the regularity estimate (4.43) to obtain

$$\|\varphi_l - R_l^{V,0}\varphi_{l-1}\|_{V_l} + \|\psi_l - \psi_{l-1}\|_{Q_l} \preceq h\,\|u_2 - R_l^{V,0}u_3\|_{V^-}.$$

Using these estimates in (4.92) and canceling one factor $\|u_2 - R_l^{V,0}u_3\|_{V^-}$ gives the result.

□

Lemma 4.10. *The post-processing step (4.83) fulfills the approximation estimate*

$$\|(u_4,p_4) - R_l^{X,0}(u_3,p_3)\|_{X_l,\tilde{0}} \preceq \|(u_1,p_1)\|_{B_l}. \tag{4.94}$$

Proof. Norm equivalence (4.57) on $V_{l,T}$, orthogonality (4.53), stability (4.54) of $X_{l,T}$ and the definition of the projection (4.71) give

$$\begin{aligned}
\|(u_4,p_4) - R_l^{X,0}(u_3,p_3)\|_{X_l,\tilde{0}} &\simeq \|u_4 - R_l^{V,0}u_3\|_{V_l} + \varepsilon^{1/2}\|p_4 - p_3\|_{c_l} \\
&\preceq \|(u_4,p_4) - R_l^{X,0}(u_3,p_3)\|_{X_l} \preceq \sup_{(v,q)\in X_{l,T}} \frac{B_l((R_l^{V,0}u_3 - u_4, p_3 - p_4),(v,q))}{\|(v,q)\|_{X_l}} \\
&= \sup_{(v,q)\in X_{l,T}} \frac{B_l((R_l^{V,0}u_3, p_3),(v,q))}{\|(v,q)\|_{X_l}} = \sup_{(v,q)\in X_{l,T}} \frac{B_l((R_l^{V,0}u_3, p_3 - I_{l-1}^{\widetilde{Q}}p_3),(v,q))}{\|(v,q)\|_{X_l}} \\
&\preceq \|(R_l^{V,0}u_3, p_3 - I_{l-1}^{\widetilde{Q}}p_3)\|_{X_l} \preceq \|(u_3, p_3 - I_{l-1}^{\widetilde{Q}}p_3)\|_{X_{l-1}} \preceq \|(u_1,p_1)\|_{B_l},
\end{aligned}$$

and the proof if complete.

□

4.5 The Smoother of Braess and Sarazin

Recently, Braess and Sarazin [BS97] suggested a new smoothing iteration for saddle point problems. It is relatively simple to implement, and it is the only known method providing an optimal dependency $O(m^{-1})$ on the number of smoothing steps. It was observed in [BS97] that the iteration depends on the primal variable, only. Although the iteration is called smoother, it is responsible for the grid transfer as well. In [Wie99] the scheme is applied for parameter dependent problems in primal variables arising from nearly incompressibility. The task of this section is the integration of the smoother by Braess and Sarazin into the theory based on the norm related to primal variables.

The construction of the smoother is based on a bilinear form $\tilde{a}_l(.,.) : V_l \times V_l$ such that the problem: Find $u_l \in V_l$ such that

$$\tilde{a}_l(u_l, v_l) = f(v_l) \qquad \forall\, v_l \in V_l \tag{4.95}$$

can be easily solved. E.g., the matrix arising from $\tilde{a}_l$ may be diagonal. It is assumed that $\tilde{a}_l$ is scaled such that

$$\tilde{a}_l(u_l, u_l) \geq a_l(u_l, u_l) \qquad \forall\, u_l \in V_l. \tag{4.96}$$

The bilinear form

$$\widetilde{A}_l(u_l, v_l) := \tilde{a}_l(u_l, v_l) + \varepsilon^{-1} c_l(\Lambda_l u_l, \Lambda_l v_l) \tag{4.97}$$

is feasible for (4.63). A special choice of space decomposition is to choose just one space ($M = 1$). Then the preconditioner D_l used in the smoother becomes a global operation. With

$$D_l(u_l, v_l) = \widetilde{A}_l(u_l, v_l)$$

and the according definition of the operator D_l, the smoothing iteration $S_{BS} : V_l \to V_l$ is written as

$$S_{BS} = I - D_l^{-1} A_l. \tag{4.98}$$

Before stepping into the analysis, we shortly comment on the implementation. One step of the iteration

$$\hat{u}_l = S_{BS} u_l$$

is written as

$$\hat{u}_l = u_l - w_l,$$

where w_l is the solution of the problem

$$\widetilde{D}_l(w_l, v_l) = A_l(u_l, v_l) \qquad \forall\, v_l \in V_l. \tag{4.99}$$

One has to solve a global saddle point problem for w_l as well, but by transforming to the (discrete) dual problem, it (may) become much simpler. First, we formulate an according mixed problem. Therefore, set $p_l = \varepsilon^{-1} \Lambda_l u_l$, and $r_l = \varepsilon^{-1} \Lambda_l w_l$. Then $(w_l, r_l) \in X_l$ is the solution of the problem in mixed variables:

$$\begin{array}{lclclclcl} \tilde{a}_l(w_l, v_l) & + & c_l(\Lambda_l v_l, r_l) & = & a_l(u_l, v_l) & + & c_l(\Lambda_l v_l, p_l) & & \forall\, v_l \in V_l, \\ c_l(\Lambda_l w_l, q_l) & - & \varepsilon\, c_l(r_l, q_l) & = & 0 & & & & \forall\, q_l \in Q_l. \end{array} \tag{4.100}$$

With the definition of $\widetilde{B}_l((u_l, p_l), (v_l, q_l)) = \tilde{a}_l(u_l, v_L) + c_l(\Lambda_l u_l, q_l) + c_l(\Lambda_l v_l, p_l) - \varepsilon\, c_l(p_l, q_l)$, the variational problem is rewritten as

$$\widetilde{B}_l((w_l, r_l), (v_l, q_l)) = B_l((u_l, p_l), (v_l, 0)) \qquad \forall\, (v_l, q_l) \in X_l. \tag{4.101}$$

In the notation of linear algebra, it reads as

$$\begin{pmatrix} \underline{\tilde{a}} & \underline{\Lambda}^T \\ \underline{\Lambda} & -\varepsilon \underline{c} \end{pmatrix} \begin{pmatrix} \underline{w} \\ \underline{r} \end{pmatrix} = \begin{pmatrix} \underline{a}\, \underline{u} + \underline{\Lambda}^T \underline{p} \\ 0 \end{pmatrix}.$$

One can pass to the Schur complement problem:

$$\left(\underline{\Lambda}\, \underline{\tilde{a}}^{-1} \underline{\Lambda}^T + \varepsilon \underline{c}\right) \underline{r} = \Lambda \underline{\tilde{a}}^{-1} (\underline{a}\, \underline{u} + \underline{\Lambda}^T \underline{p}) \tag{4.102}$$

It is assumed, that this problem is simple, and can be solved by a few steps of a preconditioned iterative scheme. The corresponding norm can be expressed by

$$\underline{p}^T (\underline{\Lambda} \underline{\tilde{a}}^{-1} \underline{\Lambda}^T + \varepsilon \underline{c}) \underline{p} = \sup_{v_l \in V_l} \frac{c_l(\Lambda_l v_l, p_l)^2}{\|v_l\|_{\tilde{a}_l}^2} + \varepsilon\, \|p_l\|_{c_l}^2. \tag{4.103}$$

If the multigrid iteration is performed in mixed variables, an approximative solution of (4.102) is theoretically understood by the work of [Zul98b]. We will assume that (4.102) will be solved exactly. If the dual variable $\underline{r}$ is calculated, the primal variable $\underline{w}$ is obtained by

$$\underline{w} = \underline{\tilde{a}}^{-1} \left(\underline{a}\, \underline{u} + \underline{\Lambda}^T \underline{p} - \underline{\Lambda}^T \underline{r} \right).$$

We will return to the Schur complement problems later, when we consider the specific examples of nearly incompressibility and Reissner Mindlin plates.

Lemma 4.11. *Assume that there holds the full regularity estimate*

$$\|u_l\|_{\tilde{a}_l} \preceq \hbar^{-1} \|u_l\|_{V^-} \qquad \forall\, u_l \in V_l. \tag{4.104}$$

Then estimate (4.68) holds with $\alpha = 1$, i.e.

$$\|u_l\|_{D_l} \preceq \|u_l\|_{V_l,\tilde{0}}. \tag{4.105}$$

Proof. The estimate follows immediately by the definition of the norm (4.67)

$$\|u_l\|^2_{V_l,\tilde{0}} = \hbar_l^{-2} \|u_l\|^2_{V^-} + \varepsilon^{-1} \|\Lambda_l u_l\|^2_{c_l} + \varepsilon^{-2} \|I^{\widetilde{Q}}_{l-1} \Lambda_l u_l\|^2_{Q_l}.$$

The last term is not used. □

Theorem 4.4 proves robust multigrid convergence with rate $O(m^{-1/2})$, if the coarse grid correction from Section 4.4 is used. The iteration of Braess and Sarazin can be used to construct robust grid transfer operations, too. Using it, the rate can be improved to $O(m^{-1})$.

The smoother of Braess Sarazin maps into the kernel space. Thus, it is bounded as a mapping from a parameter free norm to a parameter dependent norm. This is formulated in the following lemma:

Lemma 4.12. *There holds the estimate*

$$\|S_{BS} u_l\|_{D_l} \leq \|u_l\|_{\tilde{a}_l} \qquad \forall\, u_l \in V_l. \tag{4.106}$$

Proof. We fix $u_l \in V_l$ and define $w_l \in V_l$ as solution of

$$D_l(w_l, v_l) = A_l(u_l, v_l) \qquad \forall\, v_l \in V_l.$$

We represent the norm by

$$\|S_{BS} u_l\|_{D_l} = \sup_{v_l \in V_l} \frac{D_l(S_{BS} u_l, v_l)}{\|v_l\|_{D_l}}. \tag{4.107}$$

The numerator is estimated by

$$\begin{aligned} D_l(S_{BS} u_l, v_l) &= D_l(u_l - w_l, v_l) = D_l(u_l, v_l) - A_l(u_l, v_l) \\ &= \tilde{a}_l(u_l, v_l) - a_l(u_l, v_l) \\ &\leq \|u_l\|_{\tilde{a}_l - a_l} \|v_l\|_{\tilde{a}_l - a_l} \\ &\leq \|u_l\|_{\tilde{a}_l}\, \|v_l\|_{D_l}. \end{aligned}$$

Here, we have applied Cauchy-Schwarz to the (semi)-definite form $\tilde{a}(.,.) - a(.,.)$. Inserting the estimate in (4.107) proves the lemma. □

It is mentioned, that we used only algebraic properties to obtain a stable grid transfer operator. Especially, we do not require stability of the discretization. Ch. Wieners pointed out this advantage, if not 100 percent stable finite element schemes like $Q_1 - P_0$ elements are used.

By a slight modification of the multigrid scheme, one can obtain the improved dependency $O(m^{-1})$ on smoothing steps. This improvement could not be obtained by local grid transfer operators. For technical reasons we assume now, that the spaces V_l are nested, and the trivial prolongation operator is just embedding, i.e.

$$R_l^{V,0} = I. \tag{4.108}$$

But nevertheless, the forms are non-nested. Then we obtain the improved approximation property.

Theorem 4.13. *Assume, there holds the regularity and approximation estimate*

$$\|u - u_l\|_{V^-} \le \hbar_l^2 \|f\|_{(V^-)^*} \tag{4.109}$$

derived from (4.48). It is assumed that there holds the full regularity estimate:

$$\|u_l\|_{\tilde{a}_l} \preceq \hbar^{-1} \|u_l\|_{V^-} \qquad \forall\, u_l \in V_l. \tag{4.110}$$

Let $R_l^{V,0} = I$. Assume, there exists an interpolation operator $I_l : V \to V_l$ which is a projection and which is continuous with respect to $\|.\|_{V^-}$. The alternative coarse grid correction step is defined by the procedure

$$u_5 = S_{BS}(I - R_l^{V,0} A_{l-1}^{-1} [R_l^{V,0}]^T A_l) S_{BS} u_1. \tag{4.111}$$

Then there holds the improved approximation property:

$$\|u_5\|_{D_l} \preceq \sup_{v_l \in V_l} \frac{A_l(u_1, v_l)}{\|v_l\|_{D_l}}. \tag{4.112}$$

Proof. We fix $u_1 \in V_l$. The coarse grid approximation is split into the steps

$$\begin{aligned} u_2 &:= S_{BS} u_1, \\ u_3 &:= R_l^{V,0} A_{l-1}^{-1} [R_l^{V,0}]^T A_l u_2, \\ u_4 &:= u_2 - u_3, \\ u_5 &= S_{BS} u_4. \end{aligned}$$

We define the linear functional $\tilde{f} \in (V^-)^*$ by

$$\tilde{f}(\tilde{v}) := A_l(u_2, I_l \tilde{v}) \qquad \forall\, \tilde{v} \in V^-.$$

Using the symmetry of the smoother we obtain for an arbitrary $\tilde{v} \in V^-$:

$$\begin{aligned} \tilde{f}(\tilde{v}) &= A_l(u_2, I_l\tilde{v}) = A_l(S_{BS}u_1, I_l\tilde{v}) \\ &= A_l(u_1, S_{BS}I_l\tilde{v}) \\ &\leq \|S_{BS}I_l\tilde{v}\|_{D_l} \sup_{v_l \in V_l} \frac{A_l(u_1, v_l)}{\|v_l\|_{D_l}} \end{aligned} \tag{4.113}$$

Next, we use Lemma 4.12, norm estimate (4.110), and continuity of I_l to obtain

$$\|S_{BS}I_l\tilde{v}\|_{D_l} \leq \|I_l\tilde{v}\|_{\tilde{a}_l} \preceq \hbar^{-1}\|I_l\tilde{v}\|_{V^-} \preceq \hbar^{-1}\|\tilde{v}\|_{V^-}. \tag{4.114}$$

Combining (4.113) with (4.114) bounds $\|\tilde{f}\|_{(V^-)^*}$ by the right hand side of (4.112), i.e.

$$\|\tilde{f}\|_{(V^-)^*} = \sup_{\tilde{v} \in V^-} \frac{\tilde{f}(\tilde{v})}{\|\tilde{v}\|_{V^-}} \preceq \hbar^{-1} \sup_{v_l} \frac{A_l(u_1, v_l)}{\|v_l\|_{D_l}}.$$

By definition of $\tilde{f}$, and by the variational definition of the coarse grid solution, there holds

$$\begin{aligned} A_l(u_2, v_l) &= \tilde{f}(v_l) \qquad \forall\, v_l \in V_l, \\ A_{l-1}(u_3, v_{l-1}) &= \tilde{f}(v_{l-1}) \qquad \forall\, v_{l-1} \in V_{l-1}. \end{aligned}$$

We define the artificial problem: Find $\tilde{u} \in V$ such that

$$A(\tilde{u}, v) = \tilde{f}(v) \quad \forall\, v \in V.$$

Then we can use the a priori estimate (4.109) to obtain

$$\begin{aligned} \|u_4\|_{V^-} &= \|u_2 - u_3\|_{V^-} \leq \|u_2 - u\|_{V^-} + \|u_3 - u\|_{V^-} && (4.115) \\ &\preceq \hbar_l^2 \|f\|_{(V^-)^*}. && (4.116) \end{aligned}$$

Finally, we apply Lemma 4.12 once more to obtain the result

$$\begin{aligned} \|u_5\|_{D_l} &= \|S_{BS}u_4\|_{D_l} \preceq \|u_4\|_{\tilde{a}_l} \\ &\preceq \hbar_l^{-1}\|u_4\|_{V^-} \\ &\preceq \hbar_l\|\tilde{f}\|_{(V^-)^*} \\ &\preceq \sup_{v_l} \frac{A_l(u_1, v_l)}{\|v_l\|_{D_l}}. \end{aligned}$$

The proof is complete. □

This approximation property gives optimal rate of convergence of order $O(m^{-1})$. The technique to proof the approximation property for non-nested forms was used by [BV90] for problems without parameters. In [Wie99] a similar technique was applied for a parameter dependent problem. Wieners used the intermediate estimate (4.115) as approximation property, and verified an according smoothing property.

Both versions lead to equivalent results for problems with full regularity. But, for problems with less regularity there may occur differences. The formulation presented here suggests to use interpolation norms based on the parameter dependent norms $\|.\|_{A_l}$ and $\|.\|_{D_l}$, while Wieners analysis suggests to use interpolation between the norms without parameters $\|.\|_{a_l}$ and $\|.\|_{\tilde{a}_l}$. This point is only mentioned, and has to be analyzed in further work.

4.5.1 Application to nearly incompressible materials

The smoother of Braess and Sarazin was applied in [Wie99] to the problem of nearly incompressible materials. We will shortly collect the involved bilinear forms. The form $\tilde{a}(.,.)$ is

$$\tilde{a}(u_h, v_h) = \sum_{i=1}^{N_l} a(\varphi_i, \varphi_j) u_i v_i \qquad \forall\, u_h = \sum u_i \varphi_i, v_h = \sum v_i \varphi_j,$$

i.e. the diagonal of the system matrix assembled with respect to the nodal basis $(\varphi_i)_{i=1}^{N_l}$. By an appropriate scaling factor $\simeq 1$ the condition (4.96) is achieved. By an inverse inequality and by Friedrichs' inequality there holds the equivalence

$$\|u_l\|_{\tilde{a}_l} \simeq h_l^{-1} \|u_l\|_0 \qquad \forall\, u_l \in V_l.$$

For problems with full elliptic regularity, the norm $\|.\|_{V^-}$ was chosen as $\|.\|_0$ and $\hbar = h$, such that

$$\|u_l\|_{\tilde{a}_l} \preceq \hbar_l^{-1} \|u_l\|_{V^-}$$

holds. The discrete Schur complement problem behaves like

$$\underline{\Lambda}\, \underline{\tilde{a}}^{-1}\, \underline{\Lambda}^T + \varepsilon \underline{c} \simeq -h_l^2 \Delta_h + \varepsilon I$$

such that a Poisson solver can be used.

4.5.2 Application to Reissner Mindlin plates

In order to get an efficient smoother, one has to define to preconditioners, namely one arising from $\tilde{a}_l(.,.)$, and one for the Schur complement. We suggest the following components. Let $\tilde{a}_l$ be the block diagonal form

$$\|(w_l, \beta_l)\|_{\tilde{a}_l}^2 = \tau^{-1} \left\{ \|w_l\|_{\tilde{a}_{l,w}}^2 + \|\beta_l\|_{\tilde{a}_{l,\beta}}^2 \right\} \tag{4.117}$$

with a proper scaling factor $\tau \simeq 1$. The preconditioners are chosen by

$$\|\beta_l\|_{\tilde{a}_{l,\beta}}^2 \simeq h^{-2} \|\beta_l\|_0^2 \tag{4.118}$$

and

$$\tilde{a}_{l,w}^{-1} := C_1^{-1} + C_2^{-1} \tag{4.119}$$

with

$$\|w_l\|_{C_1}^2 \simeq h^{-2}\, \|w_l\|_1^2 \qquad \text{and} \qquad \|w_l\|_{C_2}^2 \simeq h^{-2} t^2\, \|w_l\|_0^2.$$

This means, we choose a sum of a preconditioner for the Poisson equation and of a diagonal preconditioner. For the Schur complement, we choose a preconditioner $\hat{S}$ such that

$$\|p_l\|_{\hat{S}}^2 \simeq (h_l^2 + t^2) \|p\|_{c_l}^2. \tag{4.120}$$

Theorem 4.14. *The smoother of Braess and Sarazin with the components (4.119) and (4.120) leads to an efficient smoother.*

Proof. First, we check conditions (4.96) and (4.110), namely

$$\|(w_l,\beta_l)\|_{a_l}^2 \leq \|(w_l,\beta_l)\|_{\tilde a_l}^2 \preceq h^{-2}\|(w_l,\beta_l)\|_{V^-}^2.$$

By the additive Schwarz lemma, and Lemma 2.18 we have

$$\begin{aligned}
\|w_l\|_{\tilde a_{l,w}}^2 &= \inf_{w_l=w_{l,1}+w_{l,2}} \left\{\|w_{l,1}\|_{C_1}^2 + \|w_{l,2}\|_{C_2}^2\right\} \\
&\simeq \inf_{w_l=w_{l,1}+w_{l,2}} \left\{h^{-2}\|w_{l,1}\|_1^2 + h^{-2}t^2\|w_{l,2}\|_0^2\right\} \\
&\simeq \inf_{w_l=w_1+w_2} \left\{h^{-2}\|w_1\|_1^2 + h^{-2}t^2\|w_2\|_0^2\right\} \\
&= h^{-2}\|w_l\|_{V^-,w}^2.
\end{aligned}$$

Additionally, there holds

$$\|\beta_l\|_{\tilde a_{l,\beta}} \simeq h^{-1}\|\beta_l\|_0 \simeq h^{-1}\|\beta\|_{V^-,\beta},$$

and thus

$$\|(w_l,\beta_l)\|_{\tilde a_l} \simeq h^{-1}\|(w_l,\beta_l)\|_{V^-}.$$

By the inverse inequality $\|(w_l,\beta_l)\|_{V-l} \preceq h^{-1}\|(w_l,\beta_l)\|_{V^-}$ proved in Theorem 2.20, and a proper scaling factor $\tau \simeq 1$, we get

$$\|(w_l,\beta_l)\|_{a_l} \leq \|(w_l,\beta_l)\|_{\tilde a_l}.$$

Additionally, we have to check that the diagonal preconditioner for the Schur complement is an optimal one. We have to check that

$$(h^2+t^2)\|p_l\|_{c_l}^2 \simeq \sup_{(w_l,\beta_l)\in V_l} \frac{c_l(\nabla w_l - \beta_l, p_l)^2}{\|(w_l,\beta_l)\|_{\tilde a_l}^2} + t^2\|p_l\|_{c_l}^2. \tag{4.121}$$

We start with bounding the right hand side from above:

$$\begin{aligned}
\sup_{(w_l,\beta_l)} \frac{c_l(\nabla w_l - \beta_l, p_l)^2}{\|(w_l,\beta_l)\|_{\tilde a_l}^2} + t^2\|p_l\|_{c_l}^2 &\preceq \left\{\sup_{(w_l,\beta_l)} \frac{\|\nabla w_l - \beta_l\|_{c_l}^2}{\|(w_l,\beta_l)\|_{\tilde a_l}^2} + t^2\right\}\|p_l\|_{c_l}^2 \\
&\preceq \left\{\sup_{(w_l,\beta_l)} \frac{(h^2+t^2)\|(w_l,\beta_l)\|_{a_l}^2}{\|(w_l,\beta_l)\|_{\tilde a_l}^2} + t^2\right\}\|p_l\|_{c_l}^2 \\
&\preceq (h^2+t^2)\,\|p_l\|_{c_l}^2.
\end{aligned}$$

We bound the right hand side from below, by choosing $w_l = 0$ and η_l in the bubble space $V_{h,bub}^{\beta}$ such that

$$\int_T \eta_l \,\mathrm{dx} = \int_T p_l \,\mathrm{dx} \quad \text{for } T \text{ with } h_T > t, \quad \text{and } \eta_l|_T = 0 \quad \text{else.}$$

This gives

$$\begin{aligned}\sup_{(w_l,\beta_l)} \frac{c_l(\nabla w_l - \beta_l, p_l)^2}{\|(w_l,\beta_l)\|^2_{\tilde a_l}} + t^2\|p_l\|^2_{c_l} &\succeq \frac{c_l(\beta_l,p_l)^2}{\|\eta_l\|^2_{\tilde a_{l,\beta}}} + t^2\,\|p_l\|^2_{c_l} \\ &\simeq \frac{\|p_l\|^4_{0,h>t}}{h^{-2}\|p_l\|^2_{0,h>t}} + t^2\,\|p_l\|^2_{c_l} \\ &\simeq (h^2+t^2)\,\|p_l\|^2_{c_l},\end{aligned}$$

and the equivalence is proved. □

4.6 Local Smoothers

In this section we will verify the smoothing property for local smoothers. The estimate to be proven is (3.47), namely

$$\|u_l\|_{[A_l,D_l]_\alpha} \preceq \|u_l\|_{V_l,\tilde 0} \qquad \forall\, u_l \in V_l. \tag{4.122}$$

The interpolation norm of index $\alpha \in (0,1]$ is defined by interpolating between the energy norm $\|.\|_{A_l}$, and the splitting norm

$$\|u_l\|^2_{D_l} = \inf_{\substack{u_l=\sum u_i \\ u_i\in V_{l,i}}} \sum_{i=1}^{M_l} \|u_l\|^2_{A_l}. \tag{4.123}$$

Since we will use only small local problems, there is no need to replace $A_l(.,.)$ by a simplified form $\widetilde{A}_l(.,.)$. But there is no difficulty, when $a_l(.,.)$ is replaced by a spectrally equivalent form $\tilde a_{l,i}(.,.)$ on the subspace $V_{l,i}$. The norm on the right hand side was defined in (4.67):

$$\|u_l\|^2_{V_l,\tilde 0} = \hbar_l^{-2}\|u\|^2_{V^-} + \varepsilon^{-1}\|\Lambda_l u_l\|^2_{c_l} + \varepsilon^{-2}\|I^{\widetilde Q}_{l-1}\Lambda_l u_l\|^2_{Q_l}. \tag{4.124}$$

The general strategy to prove estimate (4.122) is to split a function u_l into three parts. The splitting has to be stable in $\|.\|_{V_l,\tilde 0}$, and each of the three terms is estimated by one of the terms of $\|.\|_{V_l,\tilde 0}$. For the case of nearly incompressible materials we use a global projection by the limit problem, while for the Reissner Mindlin plate we can use local interpolation operators.

4.6.1 A local smoother for nearly incompressible materials

We verify the smoothing property for the local preconditioner introduced in Section 4.1.2. We recall the bilinear form for the limit case $\varepsilon = 0$

$$B^0((u,p),(v,q)) = (e(u),e(v))_0 + (\operatorname{div} u, q)_0 + (\operatorname{div} v, p)_0. \tag{4.125}$$

The abstract norm $\|.\|_{V_l,\tilde 0}$ has the specific form

$$\|u\|^2_{V_l,\tilde 0} = h^{-2}\|u\|^2_0 + \varepsilon^{-1}\|\overline{\operatorname{div} u}^{\,l}\|^2_0 + \varepsilon^{-2}\|\overline{\operatorname{div} u}^{\,l-1}\|^2_0.$$

Theorem 4.15 (Smoothing Property). *Assume that the domain is simply connected, and pure Dirichlet conditions are posed on V. Assume there holds full elliptic regularity for the limit problem, i.e. the solution $(\tilde{u}, \tilde{p}) \in X$ of*

$$B^0((\tilde{u}, \tilde{p}), (v, q)) = (\tilde{f}, v) \qquad \forall\, (v, q) \in X$$

fulfills

$$\|\tilde{u}\|_2 + \|\tilde{p}\|_1 \leq \|\tilde{f}\|_0.$$

Then estimate (4.122) holds with $\alpha = 0.5$.

Proof. We split $u_l = u^1 + u^2 + u^3$ by solving for $(u^i, p^i) \in X_l$ such that

$$\begin{aligned} B^0((u^1, p^1), (v, q)) &= B^0((u_l, 0), (v, 0)), \\ B^0((u^2, p^2), (v, q)) &= B^0((u_l, 0), (0, q - I_{l-1}^Q q)), \\ B^0((u^3, p^3), (v, q)) &= B^0((u_l, 0), (0, I_{l-1}^Q q)) \qquad \forall (v, q) \in X_l. \end{aligned} \tag{4.126}$$

The splitting is constructed such that u^1 is discrete divergence free, u^2 has non-smooth divergence and u^3 has smooth divergence. Then we apply the triangle inequality, Lemma 4.17 - 4.19, and Lemma 4.16 below to obtain (4.122) by

$$\begin{aligned} \|u_l\|_{[D_l, A_l]_{1/2}} &\leq \|u^1\|_{[D_l, A_l]_{1/2}} + \|u^2\|_{[D_l, A_l]_{1/2}} + \|u^3\|_{[D_l, A_l]_{1/2}} \\ &\preceq \|u^1\|_{V_l, \tilde{0}} + \|u^2\|_{V_l, \tilde{0}} + \|u^3\|_{V_l, \tilde{0}} \\ &\preceq \|u_l\|_{V_l, \tilde{0}}. \end{aligned}$$

□

Lemma 4.16. *The decomposition (4.126) is stable in $\|.\|_{V_l, \tilde{0}}$ norm, namely*

$$\|u^1\|_{V_l, \tilde{0}} + \|u^2\|_{V_l, \tilde{0}} + \|u^3\|_{V_l, \tilde{0}} \preceq \|u_l\|_{V_l, \tilde{0}}. \tag{4.127}$$

Proof. By Theorem 2.8 (Brezzi) we get the bounds $\|u^1\|_1 + \|p^1\|_0 \preceq \|u_l\|_1$ and $\|u^2\|_1 + \|p^2\|_0 \preceq \|I_l^Q \operatorname{div} u_l\|_0$. First, we bound $\|u^1\|_{V_l, \tilde{0}}^2 = h^{-2} \|u^1\|_0^2$. The solution of the dual problem find $(\varphi, \psi) \in X$ such that

$$B^0((\varphi, \psi), (v, q)) = (u^1, v)_0 \qquad \forall\, (v, q) \in X,$$

is bounded by $\|\varphi\|_2 + \|\psi\|_1 \preceq \|u^1\|_0$. By Galerkin orthogonality, approximation, regularity, the inverse inequality $h\, \|u_l\|_1 \preceq \|u_l\|_0$, and integration by parts we obtain

$$\begin{aligned} \|u^1\|_0^2 &= B^0((\varphi, \psi), (u^1, p^1)) \\ &= B^0((\varphi, \psi) - I_l^X (\varphi, \psi), (u^1, p^1)) + B^0(I_l^X (\varphi, \psi) - (\varphi, \psi), (u_l, 0)) + B^0((\varphi, \psi), (u_l, 0)) \\ &\preceq h\, (\|\varphi\|_2 + \|\psi\|_1)\, (\|u^1\|_1 + \|p^1\|_0) + h\, (\|\varphi\|_2 + \|\psi\|_1)\, \|u_l\|_1 + (\|\varphi\|_2 + \|\psi\|_1) \|u_l\|_0 \\ &\preceq h\, \|u^1\|_0\, \|u_l\|_1 + \|u^1\|_0 \|u_l\|_0 \preceq \|u^1\|_0\, \|u_l\|_0. \end{aligned}$$

Next, we verify the estimate

$$h^{-2}\|u^2\|_0^2 \preceq \|I_l^Q \operatorname{div} u_l\|_0^2 \preceq \varepsilon\, \|u_l\|_{V_l,\tilde{0}}^2. \tag{4.128}$$

Therefore let

$$B^0((\varphi,\psi),(v,q)) = (u^2,v)_0 \qquad \forall\,(v,q)\in X,$$

then we get by $B((u^2,p^2),(v_l,q_{l-1})) = 0 \quad \forall\, v_l\in V_l,\ q_{l-1}\in Q_{l-1}$ the bound

$$\begin{aligned}\|u^2\|_0^2 &= B^0((\varphi,\psi)-(I_l^V\varphi, I_{l-1}^Q\psi),(u^2,p^2))\\ &\preceq\ h(\|\varphi\|_2+\|\psi\|_1)\,(\|u^2\|_1+\|p^2\|_0) \preceq \|u^2\|_0\, h\,\|I_l^Q \operatorname{div} u_l\|_0.\end{aligned}$$

Dividing by $\|u^2\|_0$ proves (4.128). The last term u^3 is bounded by the triangle inequality. □

The discrete divergence free part u^1 is estimated by lifting to the potential space and Sobolev-Space interpolation in the next lemma.

Lemma 4.17. *Let u^1 be defined in (4.126). Then the estimate*

$$\|u^1\|_{[D_l,A_l]_{1/2}} \preceq \|u^1\|_{V_l,\tilde{0}} \tag{4.129}$$

is valid.

Proof. We recall the lifting operator $E: V_l\to V$ and the Fortin operator $I_l^F: V\to V_l$ we have introduced in Section 4.1.1. The operator E solves Stokes problems in each triangle to map $V_{l,0}$ to V_0. The interpolation operator I_l^F uses only values from ∂T, such that it is a left inverse to E.

By stability of the Stokes problem, and Friedrichs' inequality there holds

$$\|Eu^1\|_1 + h^{-1}\|Eu^1\|_0 \preceq h^{-1}\|u^1\|_0.$$

By the assumptions onto the domain and boundary conditions, and $\operatorname{div} Eu_1 = 0$, there exists a potential $\varphi\in H_0^2(\Omega)$ such that

$$Eu^1 = \operatorname{rot}\varphi \qquad \|\varphi\|_2 + h^{-1}\|\varphi\|_1 \preceq h^{-1}\|u^1\|_0.$$

Different as in Section 4.2.1, we have no optimal bound for $h^{-2}\|\varphi\|_0$. Thus we have to apply operator interpolation.

On one hand, we have

$$\|I_l^F \operatorname{rot}\tilde{\varphi}\|_{A_l}^2 \simeq \|I_l^F \operatorname{rot}\tilde{\varphi}\|_1^2 \preceq \|\tilde{\varphi}\|_2^2 \qquad \forall\,\tilde{\varphi}\in H_0^2(\Omega). \tag{4.130}$$

We use the partition of unity from Section 4.1.2 to estimate for any $\tilde{\varphi}\in H_0^2$

$$\begin{aligned}\|I_l^F \operatorname{rot}\tilde{\varphi}\|_D^2 &\preceq \sum \|I_l^F \operatorname{rot}(\psi_i\tilde{\varphi})\|_{A_l}^2 \preceq \sum \|\operatorname{rot}(\psi_i\tilde{\varphi})\|_{A_l}^2 \preceq \sum \|\psi_i\tilde{\varphi}\|_2^2\\ &\preceq\ \sum \left(h^{-4}\|\tilde{\varphi}\|_{0,\Omega_i}^2 + \|\tilde{\varphi}\|_{2,\Omega_i}^2\right) \preceq \left(h^{-4}\|\tilde{\varphi}\|_0^2 + \|\tilde{\varphi}\|_2^2\right)\end{aligned}$$

Let $I_h^{H^2} : L_2 \to H_0^2$ be a local regularization operator at mesh size h such that

$$\begin{aligned} h^{-2}\|(I - I_h^{H^2})\varphi\|_0 + \|I_h^{H^2}\varphi\|_2 &\preceq \|\varphi\|_2, \\ \|I_h^{H^2}\varphi\|_0 + h^2\|I_h^{H^2}\varphi\|_2 &\preceq \|\varphi\|_0 \end{aligned}$$

holds. One can choose a Clément operator mapping to a finite element subspace of H_0^2. Then there holds

$$\|I_l^F \mathrm{rot} I_h^{H^2}\tilde{\varphi}\|_{A_l} \preceq \|\tilde{\varphi}\|_2 \qquad \forall\, \tilde{\varphi} \in H_0^2,$$

and

$$\|I_l^F \mathrm{rot}\, I_h^{H^2}\tilde{\varphi}\|_{D_l} \preceq h^{-2}\|I_h^{H^2}\tilde{\varphi}\|_0 + \|I_h^{H^2}\tilde{\varphi}\|_2 \preceq h^{-2}\|\tilde{\varphi}\|_0 \qquad \forall\, \tilde{\varphi} \in L_2.$$

We use operator interpolation, and $H_0^1 = [L_2, H_0^2]_{1/2}$ (see [Bra95]) to conclude

$$\|I_l^F \mathrm{rot}\, I_h^{H^2}\tilde{\varphi}\|_{[A_l,D_l]_{1/2}} \preceq h^{-1}\|\tilde{\varphi}\|_1 \qquad \forall\, \tilde{\varphi} \in H_0^1.$$

Since $A_l \preceq D_l$, there holds for the other component

$$\begin{aligned} \|I_l^F \mathrm{rot}(\tilde{\varphi} - I_h^{H^2}\tilde{\varphi})\|_{[A_l,D_l]_{1/2}} &\preceq \|I_l^F \mathrm{rot}(\tilde{\varphi} - I_h^{H^2}\tilde{\varphi})\|_{D_l} \\ &\preceq h^{-2}\|\tilde{\varphi} - I_h^{H^2}\tilde{\varphi}\|_0 + \|\tilde{\varphi} - I_h^{H^2}\tilde{\varphi}\|_2 \\ &\preceq \|\tilde{\varphi}\|_2. \end{aligned}$$

Thus we obtain

$$\begin{aligned} \|u^1\|_{[D_l,A_l]_{1/2}} &= \|I_l^F \,\mathrm{rot}\, \varphi\|_{[D_l,A_l]_{1/2}} \\ &\leq \|I_l^F \,\mathrm{rot}\, I_h^{H^2}\varphi\|_{[D_l,A_l]_{1/2}} + \|I_l^F \,\mathrm{rot}\,(\varphi - I_h^{H^2}\varphi)\|_{[D_l,A_l]_{1/2}} \\ &\preceq h^{-1}\|\varphi\|_1 + \|\varphi\|_2 \\ &\preceq h^{-1}\|u^1\|_0 \preceq \|u^1\|_{V_l,\tilde{0}}. \end{aligned}$$

The lemma is proven. □

The component u^2 is orthogonal to divergence free functions and has non-smooth divergence.

Lemma 4.18. *Let u^2 be defined in (4.126). Then the estimate*

$$\|u^2\|_{[D_l,A_l]_{1/2}} \preceq \|u^2\|_{V_l,\tilde{0}} \tag{4.131}$$

is valid.

Proof. We use $\|.\|_{A_l}^2 \preceq \|.\|_{D_l}^2$, $\|.\|_{A_l}^2 \preceq h^{-2}\varepsilon^{-1}\|.\|_0^2$ and the intermediate result (4.128) to obtain

$$\begin{aligned} \|u^2\|_{[D_l,A_l]_{1/2}}^2 &\preceq \|u^2\|_{D_l}^2 = \inf_{u^2=\sum u_i} \sum \|u_i\|_{A_l}^2 \\ &\preceq \inf_{u^2=\sum u_i} h^{-2}\varepsilon^{-1}\|u_i\|_0^2 \preceq h^{-2}\varepsilon^{-1}\|u^2\|_0^2 \preceq \|u_l\|_{V_l,\tilde{0}}^2. \end{aligned}$$

□

The part u^3 with smooth divergence will now be estimated by better approximation of the coarse grid interpolant of the dual variable.

Lemma 4.19. *Let u^3 be defined in (4.126). Then the estimate*

$$\|u^3\|_{[D_l,A_l]_{1/2}} \preceq \|u^3\|_{V_l,\tilde{0}} \tag{4.132}$$

is valid.

Proof. By definition of u^3 we have $I_l^Q \operatorname{div} u^3 = I_{l-1}^Q \operatorname{div} u$, and together with stability of X_{l-1} we get $\|u^3\|_1 \leq \|I_{l-1}^Q \operatorname{div} u_l\|_0$. This gives

$$\|u^3\|_{A_l}^2 \preceq (\|u^3\|_1^2 + \varepsilon^{-1}\|I_l^Q \operatorname{div} u^3\|_0^2) \preceq \varepsilon^{-1}\|I_{l-1}^Q \operatorname{div} u_l\|_0^2 \preceq \varepsilon\, \|u^3\|_{V_l,\tilde{0}}^2.$$

On the other hand, we have

$$\|u^3\|_{D_l}^2 \preceq \inf_{u^3=\sum u_i} h^{-2}\varepsilon^{-1}\|u_i\|_0^2 \preceq \varepsilon^{-1}\, h^{-2}\, \|u^3\|_0^2 \preceq \varepsilon^{-1}\, \|u^3\|_{V_l,\tilde{0}}^2.$$

By operator interpolation we finish the proof.

□

4.6.2 A local smoother for Reissner Mindlin plates

We verify the smoothing property of the local preconditioner analyzed in Section 4.2.2 for the two level method. Here, the abstract norm $\|.\|_{V_l,\tilde{0}}$ has the specific form

$$\begin{aligned} \|(w,\beta)\|_{V_l,\tilde{0}}^2 &= h^{-2}\|(w,\beta)\|_{V^-}^2 + t^{-2}\|\Lambda_l(w,\beta)\|_{c_l}^2 \\ &= h^{-2} \inf_{w=w_0+w_r}\{\|w_0\|_1^2 + t^{-2}\|w_r\|_0^2\} + h^{-2}\|\beta\|_0^2 + t^{-2}\|\overline{\nabla w - \beta}^l\|_0. \end{aligned}$$

Theorem 4.20. *Assume that pure Dirichlet boundary conditions are posed on V. Then the condition (4.122) holds with $\alpha = 0.5$, i.e. the estimate*

$$\|(w_l,\beta_l)\|_{[A_l,D_l]_{1/2}} \preceq \|(w_l,\beta_l)\|_{V_l,\tilde{0}} \tag{4.133}$$

is valid for all $(w_l,\beta_l) \in V_l$.

Proof. We will split the finite element function (w_l,β_l) into a Kirchhoff part, and the rest. For this we define two interpolation operators. Let $I_l^{H^2} : L_2 \to H_0^2$ be a local regularization operator at length-scale h. It shall fulfill the approximation inequalities

$$\|w - I_l^{H^2} w\|_k \preceq h^{m-k}\|w\|_m \qquad \forall\, w \in H_0^l \quad 0 \leq k < m \leq 2,$$

and the continuity and inverse inequalities

$$\|I_l^{H^2} w\|_k \preceq h^{m-k}\|w\|_m \qquad \forall\, w \in H_0^l \quad 0 \leq m \leq k \leq 2.$$

Further, let $I_t^w : L_2 \to V^w$ a regularization operator at length-scale t. It shall fulfill the approximation

$$\|w - I_t^w w\|_0 \preceq t\,\|w\|_1 \qquad \forall\, w \in H_0^1,$$

and the continuity and inverse inequalities

$$\|I_t^w w\|_k \preceq t^{m-k}\|w\|_m \qquad \forall\, w \in H_0^l \quad 0 \le m \le k \le 1.$$

This interpolation operator is feasible for Lemma 2.17. We recall the Fortin operator

$$I_l^F = I_l^{F,1} + I_l^{F,2}(I - I_l^{F,1}),$$

where we use a Scott-Zhang Operator $I_l^{F,1}$ preserving quadratic polynomials for w and linear polynomials for β. The operator $I_l^{F,2}$ adjusts β-bubbles for elements with $h > t$, see Section 4.2.2. Then the Kirchhoff part is defined as

$$(w^1, \beta^1) := I_l^F(I_l^{H^2} I_t^w w_l, \nabla\, I_l^{H^2} I_t^w w_l). \tag{4.134}$$

The rest is

$$(w^2, \beta^2) := (w_l, \beta_l) - (w^1, \beta^1).$$

We recall the two-level approximation norm (4.30)

$$\|(w,\beta)\|_{V_l,loc}^2 = \frac{1}{h^2}\|\beta\|_0^2 + \frac{1}{h^2(h+t)^2}\|w\|_0^2.$$

By Lemma 2.17, we can use a constructive expression of the norm

$$\|(w_l,\beta_l)\|_{V^-}^2 \simeq \|I_t^w w_l\|_1^2 + t^{-2}\|w_l - I_t^w w_l\|_0^2.$$

Using this representation, one verifies that

$$\|(w^2,\beta^2)\|_{V_l,loc}^2 + t^{-2}\|\overline{\nabla w^2 - \beta^2}\|_0^2 \preceq h^{-2}\|(w_l,\beta_l)\|_{V^-}^2 + t^{-2}\|\overline{\nabla w_l - \beta_l}\|_0^2,$$

and (4.35) gives the estimate for the rest (w^2, β^2):

$$\|(w^2,\beta^2)\|_{D_l} \preceq \|(w_l,\beta_l)\|_{V_l,\tilde{0}}.$$

To bound the Kirchhoff part (w^1, β^1) we have to apply operator interpolation. We define $T : L_2 \to V_l$ by

$$\tilde{w} \to T\tilde{w} := I_l^F(I_l^{H^2}\tilde{w}, \nabla I_l^{H^2}\tilde{w}). \tag{4.135}$$

We use Theorem 2.13 to estimate

$$\begin{aligned}
\|T\tilde{w}\|_{A_l}^2 &\preceq \|(I_l^{H^2}\tilde{w}, \nabla I_l^{H^2}\tilde{w})\|_A^2 \\
&\simeq \|\nabla I_l^{H^2}\tilde{w}\|_1^2 + t^{-2}\|0\|_0^2 \\
&\preceq \|I_l^{H^2}\tilde{w}\|_2^2 \\
&\preceq \|\tilde{w}\|_2^2.
\end{aligned}$$

Using a partition of unity, the additive Schwarz lemma, and an inverse inequality for $I_l^{H^2}$, we obtain

$$\begin{aligned} \|T\tilde{w}\|_{D_l}^2 &\leq \sum \|T(\psi_i \tilde{w})\|_{A_l}^2 \\ &\preceq \sum \|I_l^{H^2}(\psi_i \tilde{w})\|_2^2 \\ &\preceq \sum h^{-4} \|(\psi_i \tilde{w})\|_0^2 \preceq h^{-4} \|\tilde{w}\|_0^2. \end{aligned}$$

Operator interpolation and $H_0^1 = [H_0^2, L_2]_{1/2}$ gives

$$\|T\tilde{w}\|_{[A_l, D_l]_{1/2}} \preceq h^{-1} \|\tilde{w}\|_1 \qquad \forall\, \tilde{w} \in H_0^1.$$

We apply this to $\tilde{w} = I_t^w w_l$ and get the result

$$\begin{aligned} \|(w^2, \beta^2)\|_{[A_l, D_l]_{1/2}} &= \|T I_t^w w_l\|_{[A_l, D_l]_{1/2}} \preceq h^{-1} \|I_t^w w_l\|_1 \\ &\preceq h^{-1} \|(w_l, \beta_l)\|_{V^-} \preceq \|(w_l, \beta_l)\|_{V_l, \tilde{0}}. \end{aligned}$$

□

Chapter 5

Numerical results

Several versions of the multigrid methods in primal variables with local smoothers have been tested numerically. The following problems have been investigated within the finite element code FEPP on a SUN Ultra 1 / 166 MHz workstation with 320 MB RAM.

Problem A: Driven Cavity example.
We consider a Stokes flow on the unit square $\Omega = (0,1)^2$. The incompressibility is approximated by a penalty term. The initial triangulation $\mathcal{T}_1$ is given by two triangles, further meshes are obtained by successive refinement. We have used the finite element space based on P_2 elements. The bilinear form $A_h(.,.)$ on the finest level is defined by averaging the divergence on the finest mesh, i.e.

$$A_h(u,v) = \int_\Omega \nabla u \nabla v \,\mathrm{dx} + \varepsilon^{-1} \int_\Omega \overline{\operatorname{div} u}^h \, \overline{\operatorname{div} v}^h \,\mathrm{dx} \tag{5.1}$$

The source term is set to $f = 0$. Dirichlet boundary conditions are specified as

$$u_h = \begin{cases} (1,0)^T & \text{at nodes } \in [0,1] \times \{1\}, \\ (0,0)^T & \text{at nodes } \notin [0,1] \times \{1\}, \end{cases}$$

and incorporated by homogenization of the FE system. A plot of the solution at level 5 is given in Figure 5.1.

Problem B: Flow through a pipe.
The geometry and the solution at level 4 are given in Figure 5.2. The boundary is split into the jacket Γ_1, inlet boundary Γ_2 and outlet boundary Γ_3. We specify homogeneous Dirichlet boundary conditions at Γ_1 and natural boundary conditions elsewhere. We solve the finite element problem find $u_h \in V_h$ such that

$$A_h(u_h, v_h) = (g, v_h)_{0,\Gamma_2} \qquad \forall\, v_h \in V_h.$$

The bilinear form $A_h(.,.)$ is the form defined in (5.1). The boundary stress is defined as $g = (0,1)^T$. The problem involves curved boundary approximation, a non-convex domain

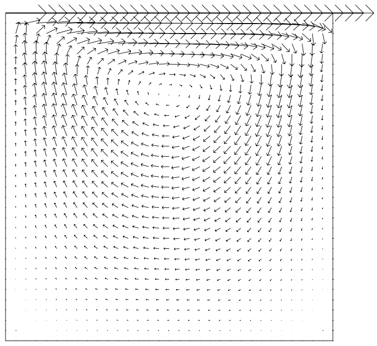


Figure 5.1: Solution of Problem A

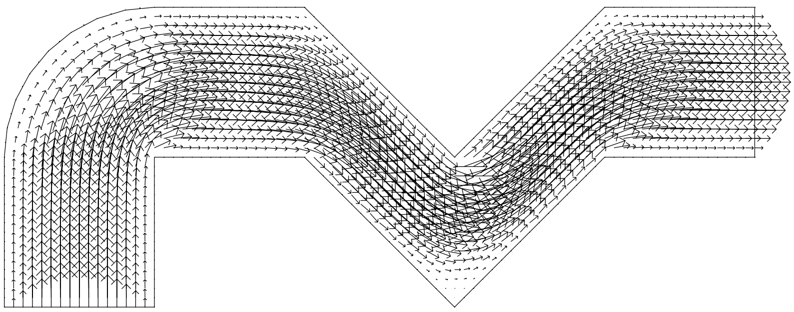


Figure 5.2: Solution of Problem B

and mixed boundary conditions.

Problem C: Nearly Incompressible Sub-domains.
We consider a problem of linear elasticity with two incompressible sub-domains. The geometry and the solution at level 4 are given in Figure 5.3. Dirichlet boundary conditions are introduced at the bottom, twisting volume forces are applied in Ω_2. The material data are $E = 100, \nu = 0.3$ in $\Omega_1 \cup \Omega_2$ and $E = 1, \nu = 0.499$ in $\Omega_3 \cup \Omega_4$.

At first, we investigate the behavior of the condition number $\kappa(C_h^{-1}A_h)$ in dependence of the number of levels L and the parameter ε. The preconditioner C_h is obtained by the application of a symmetric multigrid operator, either a W-2-2 cycle or a V-1-1 cycle. We used an additive smoother as well as a multiplicative smoother. The numerical results for the condition number $\kappa(C_h^{-1}A_h)$ for Problem A obtained by the Lanzcos method are given in Table 5.1 for a V-1-1 cycle and in Table 5.2 for a W-2-2 cycle. For the W-2-2 the calculated condition numbers neither depend on the level nor on the parameter, what is in correspondence with the analysis provided. We do not have optimal estimates for V-cycle convergence rate yet, but the numerical results seem to be very promising.

Next, we used the V-1-1 multigrid preconditioner in a preconditioned conjugate gradi-

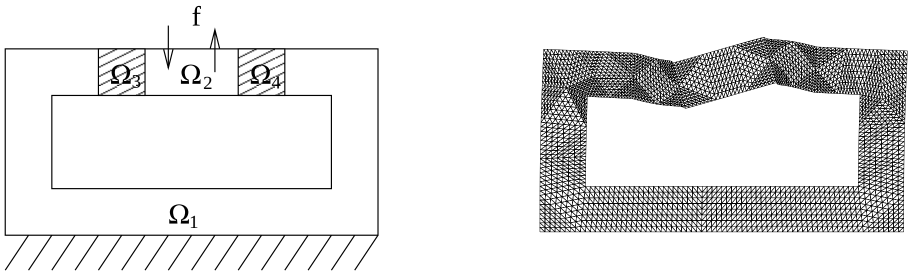


Figure 5.3: Geometry and solution of Problem C

l	Unknowns	additive smoother				multiplicative smoother			
	$\varepsilon =$	10^0	10^{-2}	10^{-4}	10^{-6}	10^0	10^{-2}	10^{-4}	10^{-6}
2	50	1.82	2.51	2.66	2.66	1.04	1.10	1.11	1.11
3	162	2.27	6.79	7.66	7.67	1.26	2.15	2.29	2.30
4	578	2.58	8.59	9.91	9.93	1.37	2.47	2.64	2.64
5	2178	2.72	9.79	11.60	11.62	1.39	2.56	2.73	2.73
6	8450	2.79	10.84	13.12	13.15	1.39	2.65	2.82	2.82
7	33282	2.73	11.66	14.41	14.45	1.39	2.72	2.90	2.91

Table 5.1: Condition numbers for V-1-1 cycle, Problem A

l	Unknowns	additive smoother				multiplicative smoother			
	$\varepsilon =$	10^0	10^{-2}	10^{-4}	10^{-6}	10^0	10^{-2}	10^{-4}	10^{-6}
2	50	1.05	1.08	1.10	1.10	1.000	1.00	1.00	1.00
3	162	1.15	1.65	1.74	1.74	1.002	1.05	1.06	1.06
4	578	1.19	1.76	1.73	1.73	1.002	1.05	1.05	1.06
5	2178	1.24	1.79	1.87	1.86	1.002	1.04	1.05	1.05
6	8450	1.26	1.86	1.92	1.91	1.002	1.05	1.05	1.05
7	33282	1.26	1.87	1.92	1.92	1.002	1.05	1.05	1.05

Table 5.2: Condition numbers for W-2-2 cycle, Problem A

Level	Unknowns	Iterations	Time[sec]
2	50	4	0.01
3	162	10	0.08
4	578	15	0.41
5	2178	15	1.88
6	8450	16	8.56
7	33282	16	37.06
8	132098	16	154.80

Table 5.3: Iteration numbers and CPU times for Problem A, PCG with V-1-1 cycle

Level	Unknowns	Iterations	Time[sec]
2	230	10	0.1
3	810	13	0.6
4	3026	15	2.7
5	11682	17	12.9
6	45890	18	58.2
7	181890	18	242.0

Table 5.4: Iteration numbers and CPU times for Problem B, PCG with V-1-1 cycle

ents solver for the solution of Problem A, Problem B, and Problem C. The small parameter is set to $\varepsilon = 10^{-6}$ in Problem A and Problem B. The iteration is terminated after a reduction of the error in energy norm by a factor of 10^8. The necessary iteration numbers and CPU times are shown in Table 5.3, Table 5.4, and Table 5.5, respectively.

Problem D: Timoshenko beam.
We consider the Timoshenko beam model. The discretization is done with P^1-elements for both variables, and element-wise averaging of the shear at the finest mesh. The considered

Level	Unknowns	Iterations	Time[sec]
2	1344	11	0.7
3	4928	14	4.1
4	18816	15	19.4
5	73472	16	87.3
6	290304	16	351.4

Table 5.5: Iteration numbers and CPU times for Problem C, PCG with V-1-1 cycle

Level	Nodes	$t = 1$	$t = 0.1$	$t = 0.01$	$t = 0.001$
3	5	1.000	1.00	1.00	1.00
4	9	1.002	1.11	1.22	1.22
5	17	1.002	1.13	1.30	1.30
6	33	1.002	1.14	1.34	1.35
7	65	1.002	1.14	1.36	1.38
8	129	1.002	1.14	1.36	1.40
9	257	1.002	1.14	1.36	1.42
10	513	1.002	1.14	1.36	1.42
11	1025	1.002	1.14	1.36	1.43
12	2049	1.002	1.14	1.36	1.43
13	4097	1.002	1.14	1.36	1.43
14	8193	1.002	1.14	1.36	1.43
15	16385	1.002	1.14	1.36	1.43

Table 5.6: Condition numbers for Timoshenko beam, V-1-1 cycle

bilinear form $A_h(.,.)$ is

$$A_h((w,\beta),(v,\eta)) = \int_0^1 w'v' \,\mathrm{dx} + t^{-2}\int_0^1 \overline{w'-\beta}^h \,\overline{v'-\eta}^h \,\mathrm{dx}.$$

Problem E: Reissner Mindlin plate.

We consider the Reissner Mindlin plate model on the unit square $\Omega = (0,1)^2$. The discretization is done by the stabilized method from [CS98]. The considered bilinear form $A_h(.,.)$ is

$$\begin{aligned} A_h((w,\beta),(v,\eta)) &= \int_\Omega e(\beta) : D : e(\eta)\,\mathrm{dx} + \frac{1}{h^2+t^2}\int_\Omega (\nabla w - \beta)^T(\nabla v - \eta)\,\mathrm{dx} \\ &\quad + \left\{1 - \frac{1}{h^2+t^2}\right\}\int_\Omega \overline{(\nabla w - \beta)}^{h,T}\,\overline{(\nabla v - \eta)}^h \,\mathrm{dx}. \end{aligned}$$

The tensor D corresponds to plane stress linear elasticity. We have chosen $E = 1$ and $\nu = 0.2$. We used P_2^+ elements for both variables, and condensed the bubble in the assembling procedure.

The relative condition numbers of $C_h^{-1}A_h$ with a V-1-1 cycle multigrid with multiplicative smoother for the Timoshenko beam (Problem D) has been measured. Table 5.6 shows the results for varying mesh size and thickness parameter. For the Reissner Mindlin plate problem, we could observe diverging W-cycle schemes for $m \leq 3$. We measured the condition numbers of $C_h^{-1}A_h$ for the V-1-1-cycle multigrid with multiplicative smoother. The

Level	Elements	mesh-size	$t = 1$	$t = 0.1$	$t = 0.01$	$t = 0.001$
3	32	0.25	1.29	1.49	3.00	3.11
4	128	0.0125	1.36	1.31	3.94	4.60
5	512	0.0062	1.39	1.39	3.50	5.40
6	4096	0.0031	1.39	1.46	2.87	5.53
7	8192	0.0016	1.39	1.47	2.78	5.08
8	32768	0.0008	1.39	1.48	2.76	4.30

Table 5.7: Condition numbers for Reissner Mindlin plate, V-1-1 cycle

results are given in Table 5.7. We mention that the condition number decreases from a certain level of refinement. The reason may be that the forms are nearly nested for $h \leq t$.

Bibliography

[AAFM99] S. M. Alessandrini, D. N. Arnold, R. S. Falk, and A. L. Madureira. Derivation and justification of plate models by variational methods. In M. Fortin, editor, *Plates and Shells*, volume 21 of *CRM Proceeding and Lecture Notes*, pages 1–20. American Mathematical Society, 1999.

[AB93] D. N. Arnold and F. Brezzi. Some new elements for the Reissner-Mindlin plate model. In J.-L. Lions and C.Baiocchi, editors, *Boundary Value Problems for Partial Differential Equations and Applications*, pages 287–292. Masson, 1993.

[Ada76] R. A. Adams. *Sobolev Spaces.* Academic Press, New York, 1976.

[AF89] D. N. Arnold and R. S. Falk. A uniformly accurate finite element method for the Reissner-Mindlin plate model. *SIAM J.Numer.Anal.*, 26:1276–1290, 1989.

[AFW97a] D. N. Arnold, R. S. Falk, and R. Winther. Preconditioning discrete approximations of the Reissner-Mindlin plate model. *Math. Mod. Numer. Anal.*, 1997.

[AFW97b] D. N. Arnold, R. S. Falk, and R. Winther. Preconditioning in $H(div)$ and applications. *Math. Comp. 66*, 66:957–984, 1997.

[AHU58] K. Arrow, L. Hurwicz, and H. Uzawa. *Studies in Nonlinear Programming.* Stanford University Press, Stanford, CA, 1958.

[AR93] U. Andelfinger and E. Ramm. EAS-elements for two-dimensional, three-dimensional, plate and shell structures and their equivalence to HR-elements. *Int. J. Num. Meth. Eng.*, 36:1311–1337, 1993.

[Arn81] D. N. Arnold. Discretization by finite elements of a parameter dependent problem. *Numer. Math.*, 37:405–421, 1981.

[Axe79] O. Axelsson. Preconditioning of indefinite problems by regularization. *SIAM J. Numer. Anal.*, 18:58–69, 1979.

[Axe96] O. Axelsson. Stabilization of AMLI methods; additive methods. In O. Axelsson and B. Polman, editors, *Proc. Conf. on Algebraic Multilevel Iteration Methods with Applications*, pages 49–62, The Netherlands, 1996. University of Nijmegen.

[BA72] I. Babuška and A. K. Aziz. Survey lectures on the mathematical foundations of the finite elemenet method. In A. K. Aziz, editor, *The Mathematical Foundation of the Finite Element Method with Applications to Partial Differential Equations*, pages 3–359. Academic Press, New York, London, 1972.

[Bat96] K.-J. Bathe. *Finite Element Procedures.* Prentice Hall, Englwood Cliffs, New Jersey, 1996.

[BB90] D. Braess and C. Blömer. A multigrid method for a parameter dependent problem in solid mechanics. *Numer. Math.*, 57:747–761, 1990.

[BBF89] K. J. Bathe, F. Brezzi, and M. Fortin. Mixed-interpolated elements for Reissner-Mindlin plates. *Internat. J. Numer. Methods. Engrg.*, 28:1787–1801, 1989.

[BBF$^+$92] F. Brezzi, M.-O. Bristeau, L. P. Franca, M. Mallet, and G. Rogé. A relationship between stabilized finite element methods and the Galerkin method with bubble functions. *Comput. Methods Appl. Mech. Engrg.*, 96:117–129, 1992.

[BD81] R. E. Bank and T. Dupont. An optimal order process for solving finite element equations. *Math. Comp.*, 36:35–51, 1981.

[BD85] R. E. Bank and C. C. Douglas. Sharp estimates for multigrid rates of convergence with general smoothing and acceleration. *SIAM J. Numer. Anal.*, 22:617–633, 1985.

[BD88] F. Brezzi and J. Douglas. Stabilized mixed methods for the Stokes problem. *Numer. Math.*, 53:225–235, 1988.

[BD96] F. Bornemann and P. Deuflhard. The Cascadic multigrid method for elliptic problems. *Numer. Math.*, 75:135–152, 1996.

[BD99] D. Braess and W. Dahmen. A Cascadic multigrid algorithm for the Stokes equation. *Numer. Math.*, 82:179–191, 1999.

[Ber96] M. Bernadou. *Finite Element Methods for Thin Shell Problems.* Wiley, Chichester, 1996.

[BF91] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods.* Springer, Berlin, Heidelberg, New York, 1991.

[BFHR97] F. Brezzi, L. P. Franca, T. J. R. Hughes, and A. Russo. $b = \int g$. *Comp. Meth. Appl. Mech. Engrg.*, 145:329–339, 1997.

[BFS91] F. Brezzi, M. Fortin, and R. Stenberg. Error analysis of mixed-interpolated elements for Reissner-Mindlin plates. *Math. Models and Methods in Appl. Sci.*, 1:125–151, 1991.

[BH83] D. Braess and W. Hackbusch. A new convergence proof for the multigrid method including the V-cycle. *SIAM J. Numer. Anal.*, 20(5):967–975, 1983.

[Bis99] M. Bischoff. *Reliable Finite Elements for Nonlinear Analysis of Shells.* PhD thesis, University Stuttgart, Institute of Structural Mechanics, 1999.

[BL91] I. Babuška and L. Li. Hierarchic modelling of plates. *Computers and Structures*, 40:419–430, 1991.

[BL97] Daniele Boffi and Carlo Lovadina. Analysis of new augmented Lagrangian formulations for mixed finite element schemes. *Numer. Math.*, 75:405–419, 1997.

[BP88] J. H. Bramble and J. E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comp.*, 50(181):1–17, 1988.

[BP92] J. H. Bramble and J. E. Pasciak. The analysis of smoothers for multigrid algorithms. *Math. Comp.*, 58(198):467–488, April 1992.

[BP93] J. H. Bramble and J. E. Pasciak. New estimates for multilevel algorithms including the V-cycle. *Math. Comp.*, 60(202):447–471, April 1993.

[BPV97] J. Bramble, J. Pasciak, and A. Vassilev. Analysis of the inexact Uzawa algorithm for saddle point problems. *SIAM J. Numer. Anal.*, 34:1072–1092, 1997.

[BPWX91] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57(195):1–21, 1991.

[BPX91] J. H. Bramble, J. E. Pasciak, and J. Xu. The analysis of multigrid algorithms with non-imbedded or non-inherited quadratic forms. *Math. Comp.*, 55:1–34, 1991.

[BR80] H. Blum and R. Rannacher. On the boundary value problem of the biharmonic operator on domains with angular corners. *Math. Meth. in the Appl. Sci.*, 2:556–581, 1980.

[Bra93] J. H. Bramble. *Multigrid Methods.* Longman Scientific & Technical, Longman House, Essex, England, 1993.

[Bra95] J. Bramble. Interpolation between Sobolev spaces in Lipschitz domains with an application to multigrid theory. *Math. Comp.*, 64:1359–1365, 1995.

[Bra96] D. Braess. Stability of saddle point problems with penalty. M^2AN, 30:731–742, 1996.

[Bra97] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics.* Cambridge University Press, 1997.

[Bra98] D. Braess. Enhanced assumed strain elements and locking in membrane problems. *Comp. Meth. in Appl. Mech. Eng.*, 165:155–174, 1998.

[Bre74] F. Brezzi. On the existence, uniqueness and approximation of saddle point problems arising from Lagrangian multipliers. *RAIRO Modél Math. Anal. Numér*, 2:129–151, 1974.

[Bre89] S. C. Brenner. An optimal-order nonconforming multigrid method for the biharmonic equation. *SIAM J. Numer. Anal.*, 26(5):1124–1138, 1989.

[Bre90] S. C. Brenner. A nonconforming multigrid method for the stationary Stokes equations. *Math. Comp.*, 55(192):411–437, 1990.

[Bre96] S. C. Brenner. Multigrid methods for parameter dependent problems. *Math. Modelling Numer. Anal.*, 30:265–297, 1996.

[BS92a] I. Babuška and M. Suri. Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, 62:439–463, 1992.

[BS92b] I. Babuška and M. Suri. On locking and robustness in the finite element method. *SIAM J. Numer. Anal.*, 29:1261–1293, 1992.

[BS92c] S. C. Brenner and L.-Y. Sung. Linear finite element methods for planar linear elasticity. *Math. Comp.*, 59(200):321–338, 1992.

[BS94] S. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer, Berlin, Heidelberg, New York, 1994.

[BS97] D. Braess and R. Sarazin. An efficient smoother for the Stokes problem. *Appl. Num. Math.*, 23:3–19, 1997.

[BV90] D. Braess and R. Verfürth. Multigrid methods for nonconforming finite element methods. *SIAM J. Numer Anal.*, 27:979–986, 1990.

[BWY90] R. E. Bank, B. D. Welfert, and H. Yserentant. A class of iterative methods for solving saddle point problems. *Numer. Math.*, 56:645–666, 1990.

[CB98] D. Chapelle and K. J. Bathe. Fundamental considerations for the finite element analysis of shell structures. *Comp. Struct.*, 66:19–36, 1998.

[Cia78] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam, 1978.

[Cia90] P. G. Ciarlet. *Plates and Junctions in Elastic Multi-Structures.* Springer, Berlin Heidelberg New-York, 1990.

[Clé75] P. Clément. Approximation by finite element functions using local regularization. *R.A.I.R.O. Anal. Numer.*, pages 77–84, 1975.

[CP94] D. Chenais and J.-C. Paumier. On the locking phenomenon for a class of elliptic problems. *Numer. Math.*, 67:427–440, 1994.

[CS98] D. Chapelle and R. Stenberg. An optimal low-order locking-free finite element method for Reissner-Mindlin plates. *Math. Models. Meth. Appl. Sciences*, 8(3):407–430, 1998.

[CTV95] P. G. Ciarlet, L. Trabucho, and J. M. Viaño, editors. *Asymptotic Methods for Elastic Structures.* de Gruyter, Berlin New York, 1995.

[DL76] G. Duvaut and J. L. Lions. *Les Inéquations en Mécanique et en Physique.* Dunod, Paris, 1976.

[DS96] P. Destuynder and M. Salaun. *Mathematical Analysis of Thin Plate Models.* Springer, Paris Berlin Heidelberg New York, 1996.

[DW89] M. Dryja and O. B. Widlund. Some domain decomposition algorithms for elliptic problems. In L. Hayes and D. Kincaid, editors, *Iterative Methods for Large Linear Systems*, pages 273–291, San Diego, CA, 1989. Academic Press.

[DW90] M. Dryja and O. B. Widlund. Towards a unified theory of domain decomposition algorithms for elliptic problems. In T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 3–21, Philadelphia, 1990. SIAM.

[EW92] R. Ewing and J. Wang. Analysis of the Schwarz algorithm for mixed finite element methods. *M^2AN Math. Modelling and Numer. Anal.*, 26:739–756, 1992.

[Fal91] Richard S. Falk. Nonconforming finite element methods for the equations of linear elasticity. *Math. Comp.*, 57:529–550, 1991.

[FG83] M. Fortin and R. Glowinski. *Augmented Lagrangian Methods.* North-Holland, Amsterdam, 1983.

[GO95] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative schwarz algorithms. *Numer. Math.*, 70:163–180, 1995.

[GR86] V. Girault and P. A. Raviart. *Finite Element Methods for Navier-Stokes Equations.* Springer, Berlin, Heidelberg, New York, 1986.

[Gri94] M. Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen.* Teubner Skripten zur Numerik. Teubner Verlag, 1994.

[Haa97] G. Haase. Hierarchical extension operators plus smoothing in domain decomposition preconditioners. *Appl. Num. Math.*, 23(3):327–346, 1997.

[Hac82] W. Hackbusch. Multi-grid convergence theory. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods*. Springer, Berlin, Heidelberg, New York, 1982.

[Hac85] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer, Berlin, Heidelberg, New York, 1985.

[Hac91] W. Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme*. Studienbücher Mathematik. B. G. Teubner, Stuttgart, 1991.

[Hip99] R. Hiptmair. Multigrid method for Maxwell's equations. *SIAM J. Numer. Anal.*, 1999.

[HLMN94] G. Haase, U. Langer, A. Meyer, and S. V. Nepomnyaschikh. Hierarchical extension operators and logcal multigrid methods in domain decomposition preconditioners. *East-West J. Numer. Math.*, 2(3):173–193, 1994.

[HS52] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards*, 49:409–436, 1952.

[Hua90] Z. Huang. A multi-grid algorithm for mixed problems with penalty. *Numer. Math.*, 57:227–247, 1990.

[Hug95] T. J. R. Hughes. Multiscale phenomena: Greens functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origin of stabilized methods. *Comput. Methods Appl. Mech. Engrg.*, 127:387–401, 1995.

[JLM+89] M. Jung, U. Langer, A. Meyer, W. Queck, and M. Schneider. Multigrid preconditioners and their applications. In G. Telschow, editor, *Third Multigrid Seminar, Biesenthal 1988*, Report R-MATH-03/89, pages 11–52, 1989.

[JP82] C. Johnson and J. Pitkäranta. Analysis of some mixed finite element methods related to reduced integration. *Math. Comp.*, 38:375–400, 1982.

[KL84] V. G. Korneev and U. Langer. *Approximate solution of plastic flow theory problems*. Teubner-Verlag, Leipzig, 1984.

[Kla98] A. Klawonn. An optimal preconditioner for a class of saddle point problems with a penalty term. *SIAM J. Sci. Comput.*, 19:540–552, 1998.

[KM99] T. V. Kolev and S. D. Margenov. Two-level preconditioning of elasticity nonconforming fem systems. *Num. Lin. Alg. with Appl.*, 1999. (to appear).

[KO76] R. B. Kellog and J. E. Osborn. A regularity result for the Stokes prbolem on a convex polygon. *J. Funct. Anal.*, 21:397–431, 1976.

[Lio88] P. L. Lions. On the Schwarz alternating method I. In R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 1–42. SIAM, 1988.

[Lov96] C. Lovadina. A new class of mixed finite element methods for Reissner-Mindlin plates. *SIAM J. Numer. Math.*, 6:2457–2467, 1996.

[LP64] J.-L. Lions and J. Peetre. Sur une classe d'espaces d'interpolation. *Institut des Hautes Etudes Scientifique, Publ.Math.*, 19:5–68, 1964.

[LQ86] U. Langer and W. Queck. On the convergence factor of Uzawa's algorithm. *J. Comput. Appl. Math.*, 15:191–202, 1986.

[LQ87] U. Langer and W. Queck. Preconditioned Uzawa-type iterative methods for solving mixed finite element equations. Wissenschaftliche Schriftenreihe 3/1987, Technische Universität Karl-Marx-Stadt, 1987.

[LS99] M. Lyly and R. Stenberg. Stabilized finite element methods for Reissner-Mindlin plates, 1999. (in preparation).

[MB96] J. M. Melenk and I. Babuska. The partition of unity finite element method: Basic theory and applications. *Comput. Meths. Appl. Mech. Engrg*, 139:289–314, 1996.

[MH78] D. S. Malkus and T. J. R. Hughes. Mixed finite element methods. Reduced and selective integration techniques : A unification of concepts. *Comp. Methods Appl. Mech. Eng.*, 15:63–81, 1978.

[MMB87] J. Mandel, S. McCormick, and R. Bank. Variational multigrid theory. In S. McCormick, editor, *Multigrid methods*. SIAM, Philadelphia, 1987.

[MN85] A. M. Matsokin and S. V. Nepomnyaschikh. Schwarz alternating method in a subspace. *Sov. Math.*, 29:78–84, 1985.

[Nep86] S. V. Nepomnyaschikh. *Domain Decomposition and Schwarz Methods in a Subspace for the Approximate Solution of Elliptic Boundary Value Problems.* PhD thesis, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR, 1986.

[Nep92] S. V. Nepomnyaschikh. Decomposition and fictitious domain methods for elliptic boundary value problems. In T. F. Chan, D. E. Keyes, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, editors, *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*. SIAM, 1992.

[Neu98] N. Neuss. V-cycle convergence with unsymmetric smoothers and application to an anisotropic model probelm. *SIAM J. Numer. Anal.*, 35:1201–1212, 1998.

[Osw94] P. Oswald. *Multilevel Finite Element Approximation.* Teubner Stuttgart, 1994.

[Pad97] A. Padiy. Multilevel iterative method for anisotropic elliptic problems. In *Proc. of the Conf. on Preconditioned Iterative Solution Methods for Large Scale Problems in Scientific Computations PRISM-97*, pages 145–169, University of Nijmegen, 1997.

[PB87] P. Peisker and D. Braess. A conjugate gradient method and a multigrid algorithm for Morley's finite element approximation of the Biharmonic equation. *Numer. Math.*, 50:567–586, 1987.

[PB92] P. Peisker and D. Braess. Uniform convergence of mixed-interpolated elements of Reissner-Mindlin plates. *Math. Mod. and Num. Anal. (M^2AN)*, 26:557–574, 1992.

[Pei91] P. Peisker. A multigrid method for Reissner-Mindlin plates. *Numer. Math.*, 59:511–528, 1991.

[Pit92] J. Pitkäranta. The problem of membrane locking in finite element analysis of cylindrical shells. *Numer. Math.*, 61:523–542, 1992.

[PS96] J. Pitkäranta and M. Suri. Design principles and error analysis for reduced-shear plate-bending finite elements. *Numer. Math.*, 75:223–266, 1996.

[Que89] W. Queck. The convergence factor of preconditioned algorithms of the Arrow-Hurwicz type. *SIAM J. Numer. Anal.*, 26:1016–1030, 1989.

[RT91] J. E. Roberts and J. M. Thomas. Mixed and hybrid methods. In P. G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis II.* North-Holland, Amsterdam, 1991.

[RW92] T. Rusten and R. Winther. A preconditioned iterative method for saddle-point problems. *SIAM J. Matrix Anal. Appl.*, 13:887–904, 1992.

[SBG96] B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain Decomposition. Parallel Multilevel Methods for Elliptic Partial Differential Equations.* Cambridge University Press, 1996.

[SBS95] M. Suri, I. Babuška, and C. Schwab. Locking effects in the finite element approximation of plate models. *Math. Comp.*, 64:461–482, 1995.

[Sch69] H. A. Schwarz. Über einige Abbildungsaufgaben. *J. Reine Angew. Math.*, 70:105–120, 1869.

[Sch89] C. Schwab. *Dimensional reduction for elliptic boundary value problems.* PhD thesis, Dept. of Mathematics, Univ. of Maryland College Park, MD 20742 USA, 1989.

[Sch96] C. Schwab. A-posteriori modeling error estimation for hierarchic plate models. *Numer. Math.*, 74:221–259, 1996.

[Sch99a] J. Schöberl. Multigrid methods for a class of parameter dependent problems in primal variables. Technical Report 99-03, Spezialforschungsbereich F013, Johannes Kepler University, 4040-Linz, Austria, 1999.

[Sch99b] J. Schöberl. Multigrid methods for a parameter dependent problem in primal variables. *Numer. Math.*, (1999). (to appear), Available as Institutsbericht 539, Johannes Kepler Universtät Linz, Austria, Fachbereich für Mathematik.

[Sch99c] J. Schöberl. Multigrid methods for parameter dependent problems I: The Stokes-type case. In W. Hackbusch and G. Wittum, editors, *Proceedings of the Fifth European Multigrid Conference 1996*. Springer, 1999.

[Sha96] V. V. Shaidurov. Some estimates of the rate of convergecne for the Cascadic conjugate gradient method. *Comp. Math. Appl.*, 31:161–171, 1996.

[Sin78] K. Singh. Reduced integration for improved accuracy of finite element approximations. *Comput. Meth. Appl. Mech. Engrg.*, 14:23–37, 1978.

[SR90] J.C. Simo and S. Rifai. A class of mixed assumed strain methods and the method of incompatible modes. *Int. J. Num. Meth. Eng.*, 29:1595–1638, 1990.

[Ste84] R. Stenberg. Analysis of mixed finite element methods for the Stokes problem: a unified approach. *Math. Comp.*, 42:9–23, 1984.

[Ste88] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53:513–538, 88.

[Ste93] R. Stevenson. Robustness of multi-grid applied to anisotropic equations on convex domains with reentrant corners. *Numer. Math.*, 66:373–398, 93.

[SZ90] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[Tur94] S. Turek. Multigrid techniques for a divergence-free finite element discretization. *East-West J. Numer. Math.*, 2(3):229–255, 1994.

[Van86] S. P. Vanka. A calculation procedure for three-dimensional steady recirculating flows using multigrid methods. *Comp. Meth. Appl. Mech. Eng.*, 55:321–338, 1986.

[Ver84] R. Verfürth. A multilevel algorithm for mixed problems. *SIAM J. Numer. Anal.*, 21:264–271, 1984.

[VL96] P. S. Vassilevski and R. D. Lazarov. Preconditioning mixed finite element saddle-point elliptic-problems. *Num. Lin. Alg. with Appl.*, 3:1–20, 1996.

[VW92] P. S. Vassilevski and J. Wang. Multilevel iterative methods for mixed finite element discretizations of elliptic problems. *Numer. Math.*, 63:503–520, 1992.

[Wie99] Ch. Wieners. Robust multigrid methods for nearly incompressible linear elasticity. *Computing*, Special edition of 10^{th} GAMM Seminar on Multigrid Method, 1999. (submitted).

[Wit89a] G. Wittum. On the convergence of multigrid methods with transforming smoothers. Theory with application to Navier-Stokes equations. *Numer. Math.*, 54:543–563, 1989.

[Wit89b] G. Wittum. On the robustness of ILU smoothing. *SIAM J. Sci. Stat. Comp.*, 10:699–717, 1989.

[Xu92] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34(4):581–613, 1992.

[Yse86] H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math*, 49:379–412, 1986.

[Zha91] X. Zhang. *Studies in Domain Decomposition: Multilevel Methods and the Biharmonic Dirichlet Problem.* PhD thesis, Courant Institute, New York, 1991.

[Zha92] X. Zhang. Multilevel Schwarz methods. *Numer. Math.*, 63:521–539, 1992.

[Zul98a] W. Zulehner. Analysis of iterative methods for saddle point problems: A unified approach. Technical Report 538, Department of Mathematics, University of Linz, Austria, 1998. (submitted).

[Zul98b] W. Zulehner. A class of smoothers for saddle point problems. Technical Report 546, Department of Mathematics, University of Linz, Austria, 1998. (submitted).

Ich, Dipl.-Ing. Joachim Schöberl, geboren am 28. Mai 1972 in Steyr, versichere daß ich die vorliegende Dissertation selbständig verfaßt habe und keine anderen Quellen als die im Literaturverzeichnis angegebenen verwendet habe. Weiters erkläre ich, daß ich die vorliegende Dissertation bisher an keiner inländischen oder ausländischen Universität vorgelegt habe.

Linz, im Juni 1999

(Dipl.-Ing. Joachim Schöberl)