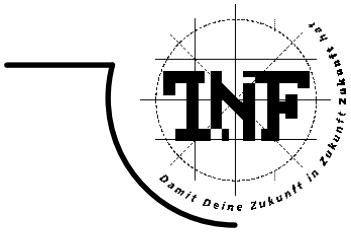
JOHANNES KEPLER
UNIVERSITÄT LINZ
Netzwerk für Forschung, Lehre und Praxis



Multigrid Solvers for Saddle Point Problems in PDE-Constrained Optimization

Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

Angefertigt am *SpezialForschungsBereich F013 - Numerical and Symbolic Scientific Computing*

Begutachter:

Prof. Dr. Walter Zulehner
Prof. Dr. Ronald H.W. Hoppe

Eingereicht von:

Dipl.-Math.techn. René Simon

Linz, Juni 2008

Abstract

This works deals with efficient solution methods for solving large scale discretized optimization problems with constraints in form of partial differential equations (PDEs). Using the simultaneous approach, i.e. the state constraint is not eliminated, one has to solve the corresponding optimality system, also called the Karush-Kuhn-Tucker (KKT) system. This leads to large scale saddle point problems.

A special class of efficient iterative solvers are multigrid methods. In this work the so-called one-shot multigrid-strategy is chosen, i.e. the multigrid method is directly applied to the whole discretized problem and not only as a preconditioner for the involved PDE and the Schur complement of the KKT system. One of the most important ingredients of a multigrid iteration is an appropriate smoother. We consider here the construction of additive Schwarz-type patch smoothers, where the computational domain is divided into small patches. On each patch local problems are solved one-by-one in a Jacobi-type manner. Under suitable conditions on the choice of the patches and the local problems, a complete multigrid convergence analysis is given. Numerical experiments are shown which confirm the theoretical results.

The theoretical and numerical results are presented for a model problem of a typical class of elliptic optimal control problems. However, the method itself carries over to more general problems, since the construction of the smoother does not rely on structural information of the problem.

In a second approach the special structure of the optimal control problem with distributed control is used to reduce the KKT system. For this reduced linear system the construction of a similar patch smoother is presented, which leads to a highly efficient multigrid method with convergence rates comparable to convergence rates known from classical scalar elliptic problems. Again a rigorous convergence analysis is given.

Zusammenfassung

Diese Arbeit beschäftigt sich mit effizienten Lösungsverfahren von großdimensionierten Optimierungsproblemen in denen die Nebenbedingungen in Form von partiellen Differentialgleichungen beschrieben sind. Ein möglicher Lösungsansatz ist es, die Nebenbedingungen, beschrieben durch die sogenannte Zustandsgleichung, nicht zu eliminieren, sondern die Optimalitätsbedingungen für das vollständige System zu formulieren. Dieses Optimalitätssystem, auch Karush-Kuhn-Tucker-System (KKT-System) genannt, ist ein großdimensioniertes Sattelpunktsproblem.

Eine spezielle Klasse von effizienten iterativen Lösungsmethoden ist die Klasse der Mehrgittermethoden. Während die Mehrgitteridee zum Beispiel zur näherungsweisen und effizienten Lösung der im Optimalitätssystem auftretenden partiellen Differentialgleichungen und des zugehörigen Schur-Komplements genutzt werden kann, konzentrieren wir hier uns auf die Anwendung des Mehrgitterverfahrens auf das gesamte Optimalitätssystem. Eines der wichtigsten Bestandteile des Mehrgitterverfahrens ist die geeignete Wahl eines Glätters. Wir beschreiben die Konstruktion von additiven Patchglättern. Dabei wird das Gebiet in kleine Teilgebiete zerlegt und auf jedem dieser kleinen Gebiete wird dann ein zugehöriges lokales Problem gelöst. Eine Glättungsschritt besteht dann darin, der Reihe nach auf jedem Teilgebiet die lokalen Probleme zu lösen und additiv zur Gesamtlösung zusammenzusetzen. Für eine geeignete Wahl der Aufteilung in Teilgebiete und der zugehörigen lokalen Probleme ist es möglich eine vollständige Konvergenzanalyse des Mehrgitterverfahrens durchzuführen. Es werden experimentelle Rechnungen präsentiert, die die theoretischen Resultate bestätigen.

Die theoretischen und numerischen Ergebnisse werden anhand eines typischen Modellproblems aus der Klasse der Optimalsteuerungsprobleme aufgeführt. Nichtsdestotrotz ist die Methode selbst auf eine allgemeinere Klasse von gemischten Problemen anwendbar, da die Konstruktion des Glätters keine tieferen strukturellen Kenntnisse des Problems voraussetzt.

In einem zweiten Ansatz wird die spezielle Struktur des Optimalsteuerungsproblems mit verteilter Steuerung ausgenutzt, um eine Variable (die Steuerungsvariable) zu eliminieren und damit das zugehörige Optimalitätssystem zu verkleinern. Für das reduzierte System wird ein analoger Glätter konstruiert. Dies führt zu einer höchst effizienten Mehrgittermethode, in der die Konvergenzraten vergleichbar mit den niedrigen Konvergenzraten der klassischen, skalaren elliptischen Probleme sind. Auch hier wird ein vollständiger Konvergenzbeweis des Mehrgitterverfahrens gegeben.

Acknowledgement

First of all I want to express my deep gratitude to Professor W. Zulehner for advising my thesis and guiding my work. Discussions with him have always been very stimulating. I want to thank Professor U. Langer for arousing my interest in computational mathematics and for offering me a doctorate position within the research project SFB F013 'Numerical and Symbolic Scientific Computing'. Both are responsible for deepening my understanding in this scientific field. At the same time, I am grateful to Professor R.H.W. Hoppe for co-refereeing this thesis.

Moreover, I want to thank all my colleagues of the SFB F013 for the nice working climate. Special thanks are devoted to Dr. J. Valdman for a lot of fruitful and inspiring discussions.

Last but not least, I want to acknowledge the financial support of the Austrian Science Fund 'Fonds zur Förderung der wissenschaftlichen Forschung' FWF under the grant SFB F013/F1309.

Contents

Chapter 1

Introduction

During recent years, the importance of solving optimization problems with constraints in form of partial differential equations (PDEs) has been growing. Typical areas are topology optimization (c.f. Bendsøe and Sigmund [15]), shape optimization (c.f. Pironneau [73]), and optimal control problems (c.f. Lions [69], Tröltzsch [89]).

There are basically two approaches to solve such optimization problems. Under proper conditions, the constraining PDE, mostly called the state equation, can be eliminated and formally hidden in the objective functional. This so-called nested or black-box approach is commonly used in practical applications including commercial applications since it allows a simple integration of the PDE solver into the optimization algorithms. Since the dimension of the problem is reduced, this approach can be advantageous, especially if the corresponding control (or design) space is small, see, e.g., Mühlhuber [70]. At every iteration step of the optimization algorithm the state constraint is fulfilled. Therefore, the nested approach can be regarded as a feasible path method. But on the other hand it is necessary to solve the state equation for each evaluation of the objective functional and of the gradient, respectively. Additionally the adjoint problem to the state equation has to be solved. Moreover the reduction of variables will possibly change the original properties of the objective functional like linearity or convexity.

In comparison to this classical nested formulation, there exists the simultaneous formulation, where the state equation is treated as constraint, see, e.g., Hoppe, Linsenmann and Petrova [64], Hoppe and Petrova [65], Burger and Mühlhuber [40, 41]. Using this approach, one can solve the optimization problem by solving the corresponding system of optimality conditions, called the Karush-Kuhn-Tucker system (KKT system). This leads to large scale symmetric and indefinite problems, which are usually solved iteratively. Since this approach does not follow the feasible path a significant speed-up is expected. But in order to exploit the potential speed-up, efficient solution techniques to solve the KKT systems are needed, which typically rely on efficient solvers for linear saddle point problems.

Saddle point systems also arise in a lot of other fields of computational science and engineering, e.g., in computational fluid dynamics (c.f. Glowinski [54], Wesseling [96]),

linear elasticity (c.f. Braess [24], Ciarlet [44]), and mixed finite element approximation of elliptic PDEs (c.f. Brezzi [38], Brezzi and Fortin [39]).

A lot of methods for solving linear saddle point problems are proposed in the literature. In Chapter 2 we will discuss a few iterative solvers for saddle point problems. For more details and a nice overview we refer to the survey article Benzi, Golub and Liesen [16].

A special class of iterative techniques are multigrid methods. Multigrid methods are highly efficient. They are optimal in the sense that the number of arithmetic operations needed to solve a discrete problem is proportional to the number of unknowns. In this thesis we discuss multigrid methods for solving large-scale systems of discretized mixed variational problems.

In particular, we will consider elliptic optimal control problems, see, e.g., Lions [69], Tröltzsch [89]. In such problems the primal unknown, say x, consists of two parts: a function y, the so-called state, and a function u, the so-called control. The problem is to find $x = (y, u)$ from appropriate function spaces that minimizes a given cost functional subject to a constraint, the so-called state equation, which, for each control u, is an elliptic boundary value problem in y. The corresponding KKT system involves another (dual) unknown, say p (the Lagrangian multiplier or the adjoint state), and consists of three components: the state equation (see above), the adjoint state equation, which, for each state y, is an elliptic boundary value problem in p, and the control equation, which is typically an algebraic relation between u and p.

In principle, there are two different approaches for mixed problems, such as KKT systems, to take advantage of the multigrid idea. One way is to use an outer iteration, typically a preconditioned Richardson method (possibly accelerated by a Krylov subspace method), applied to the discretized mixed problem. For typical preconditioners of KKT systems in elliptic optimal control, see, e.g., Battermann and Heinkenschloss [13], Battermann and Sachs [14], Biros and Ghattas [17], Hazra and Schulz [62] and the references cited there. These preconditioners usually rely on efficient solvers or preconditioners for the state equation (as a PDE in y) and the adjoint state equation (as a PDE in p) and on the construction of a good preconditioner for the corresponding Schur complement of the KKT system, which is the reduced Hessian of the Lagrangian. A preconditioner based on a different Schur complement is proposed in Schöberl and Zulehner [83]. Multigrid techniques (as an inner iteration or approximation) can be used for (some or all of) these components, see, e.g., Hackbusch [58], Dreyer, Maar and Schulz [48], Schöberl and Zulehner [83].

The other way is to use multigrid methods directly applied to the (discretized) mixed problem as an outer iteration based on appropriate smoothers (as a sort of inner iteration). For PDE-constrained optimization problems this approach is also known as one-shot multigrid strategy, see Ta'asan [88]. One of the most important ingredients of such a multigrid method is an appropriate smoother.

A first approach for constructing such smoothers is to combine standard smoothers applied to the components elliptic state and adjoint equations complemented with a special relaxation method for the control equation, see, e.g., Arian and Ta'asan [2].

A second class of smoothers are point smoothers, where the variables, here y, u and p, are grouped pointwise (with respect to the points (nodes) of the underlying mesh) and one or several sweeps of point-block Jacobi or point-block Gauß-Seidel iterations with respect to this grouping are performed, see, e.g, Borzi, Kunisch and Kwak [20]. Recently this approach of solving local optimality systems on the grid-point level was used to solve optimal control problems with pointwise control and state constraints, see Borzi [18].

A natural extension of point smoothers are patch smoothers: The computational domain is divided into small (overlapping or non-overlapping) patches. One iteration step of the smoothing process consists of solving local mixed problems on each patch one-by-one either in a Jacobi-type or Gauss-Seidel-type manner. This results in an additive or multiplicative Schwarz-type smoother. The technique was successfully used for the Navier-Stokes equations, see Vanka [92]. The general construction and the analysis of patch smoothers for mixed problems was discussed in Schöberl and Zulehner [82], where a particular patch smoother was proposed for the Stokes problem. An essential feature exploited in the multigrid convergence analysis of the Stokes problem was (in the terminology introduced here) that the adjoint state equation is an elliptic problem in y, where in elliptic optimal control problems the adjoint state equation is typically elliptic in p but not necessarily in y. Therefore, a straight forward application of this construction to KKT systems for elliptic optimal control problems fails.

Another well-known class of smoothers for mixed problems are Braess-Sarazin smoothers, see, Braess and Sarazin [26], Zulehner [101], which are well suited for Stokes-like problems but are too expensive if applied to KKT systems of elliptic optimal control problems.

A last approach discussed here for constructing smoothers for mixed problems leads to so-called transforming smoothers, see Wittum [97, 98], which were successfully analyzed for the Stokes problem and the Navier-Stokes problem. Recently, the analysis of transforming smoothers has been extended to the class of PDE constrained optimization problems, see Schulz and Wittum [84].

So far, the multigrid convergence analysis for KKT systems of PDE-constrained optimization problems is not as developed as for elliptic PDEs. One line of argument exploits the fact that the KKT system reduced to y and p by eliminating u using the control equation is a compact perturbation of an elliptic problem. This guarantees the convergence of the multigrid method if the coarse grid is sufficiently fine, see Borzi, Kunisch and Kwak [20]. A second strategy is based on a Fourier analysis, which, strictly speaking, covers only the case of uniform meshes with special boundary conditions (and small perturbations of this situation), see, e.g., Borzi, Kunisch and Kwak [20], Arian and Ta'asan [2].

The aim of this thesis is to contribute to the one-shot multigrid approach for KKT systems. In order to keep the notations simple and the strategy transparent the material is presented for a model problem in optimal control only. However, since the construction of the method does not rely on structural information of some Schur complement, the method is easily applicable to more general problems.

This thesis is structured as follows. The second chapter contains the preliminaries which are needed to describe the multigrid method for mixed problems and for the convergence analysis. We briefly introduce variational problems (in primal and mixed form), the basics of the finite element method and error analysis, and finally discuss classical solution methods for the linear system.

In the third chapter we introduce the multigrid method and the basic convergence analysis. The second part of this chapter is devoted to the construction of appropriate smoothers with the emphasis on patch smoother. Finally a short introduction in the local mode analysis as a tool for quantitative analysis and design of smoothing methods is given.

The central part of this thesis is Chapter 4 and 5. In Chapter 4 we apply the general construction of the patch smoother to the KKT system arising from an optimal control problem and give a rigorous convergence analysis. At the end of this chapter a few numerical results are presented.

In Chapter 5 the same optimal control problem as in Chapter 4 is investigated. The so-called control equation is used to reduce the KKT system. The reduced system is then again solved by a multigrid method and the smoothing method is discussed in detail. We compare our theoretical and numerical results with known results from literature. The local mode analysis is used to tune the smoothing method in order to improve the computational performance of the multigrid method.

Finally in the last chapter we end with some concluding remarks and give an outlook on possible future work.

Chapter 2

Preliminaries

In this chapter we introduce variational problems and collect the most important analytical results, concerning existence and uniqueness of the solution and error analysis of the finite element approximation. For simplicity we restrict ourselves to a model Poisson problem with homogeneous Neumann boundary conditions. This model problem will be later play the role of the state constraints in an optimal control problem. In the first part on primal variational problems we mainly follow the introduction of finite element methods in Braess [23]. The second part of this chapter is devoted to mixed variational problems. Again, existence and uniqueness of the solution and the error analysis is discussed. Finally, we will give a short overview on approaches to solve the resulting linear system of equations.

2.1 Sobolev Spaces

The natural spaces for variational problems are Sobolev spaces. For a detailed description of Sobolev Spaces we refer to Adams [1]. Let the domain Ω be a bounded open set from $\mathbb{R}^d$ with $d \in \{1, 2, 3\}$. Here we are only interested in the subset of Hilbert spaces, which are build up on the $L^2(\Omega)$ space, consisting of all square-Lebesgue integrable functions. With the scalar product

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

$L^2(\Omega)$ is a Hilbert space with associated norm

$$\|u\|_{L^2(\Omega)} := (u, u)^{1/2}_{L^2(\Omega)}.$$

The function $z = D^\alpha u$ is said to be the α-th *weak* derivative of the function $u \in L^2(\Omega)$ if $z \in L^2(\Omega)$ and the following is satisfied:

$$\int_\Omega z(x)v(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u(x)D^\alpha v(x)\,\mathrm{d}x \quad \text{for all } v \in C_0^\infty(\Omega),$$

where $C_0^\infty(\Omega)$ denotes the set of $C^\infty(\Omega)$ functions with compact support in Ω.

Now we can define the Hilbert spaces $H^m(\Omega)$:

Definition 1. *Let m be a non-negative integer and suppose that the weak derivatives $D^\alpha u$ exist for all $|\alpha| \leq m$. The Hilbert space $H^m(\Omega)$ is the set of all such functions $u \in L^2(\Omega)$ equipped with the scalar product*

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

and the associated norm

$$\|u\|_{H^m(\Omega)} := (u, u)^{1/2}_{H^m(\Omega)} = \left(\sum_{|\alpha| \leq m} \|D^\alpha u\|^2_{L^2(\Omega)} \right)^{1/2}.$$

Additionally we define the semi-norm

$$|u|_{H^m(\Omega)} := \left(\sum_{|\alpha| = m} \|D^\alpha u\|^2_{L^2(\Omega)} \right)^{1/2}.$$

2.2 Primal Variational Problems

As a short introduction in variational problems, we start with primal problems. As an example we consider one of the most elementary boundary value problem: the Poisson equation with homogeneous Neumann boundary conditions.

$$\begin{aligned} -\Delta u(x) + u(x) &= f(x) \qquad x \in \Omega, \\ \frac{\partial u(x)}{\partial n(x)} &= 0 \qquad x \in \Gamma, \end{aligned} \tag{2.1}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded convex domain and Γ denotes the boundary $\partial\Omega$.

In order to get the variational formulation, we multiply the PDE with a testfunction v, integrate over the domain Ω and use Gauss' Theorem. We end up with the following variational problem: Find $u \in V = H^1(\Omega)$ such that

$$\int_\Omega \nabla u(x)^T \nabla v(x) \, \mathrm{d}x + \int_\Omega u(x) v(x) = \int_\Omega f(x) v(x) \, \mathrm{d}x, \qquad \forall v \in V. \tag{2.2}$$

For the abstract formulation, we define the bilinearform $a \colon V \times V \longrightarrow \mathbb{R}$ by

$$a(u, v) = \int_\Omega \nabla u(x)^T \nabla v(x) \, \mathrm{d}x + \int_\Omega u(x) v(x), \qquad u, v \in V,$$

and the linear functional $F\colon V \longrightarrow \mathbb{R}$ by

$$\langle F, v\rangle = \int_\Omega f(x)v(x)\,\mathrm{d}x, \qquad v \in V.$$

In the following discussion of primal variational problems, if we do not pose any additional conditions, we will always assume the following:

Assumption 1. *1. $(H, (\cdot,\cdot))$ is a Hilbert space.*

2. V is a closed subspace of H.

3. $a(\cdot,\cdot)$ is a bilinear form on V.

4. $a(\cdot,\cdot)$ is continuous (bounded) on V, i.e., there exists a constant $c_1 > 0$ such that

$$|a(w,v)| \le c_1\|w\|_V\|v\|_V, \quad \textit{for all } w, v \in V.$$

5. $a(\cdot,\cdot)$ is coercive on V, i.e., there exists a constant $c_2 > 0$ such that

$$a(w,w) \ge c_2\|w\|_V^2, \quad \textit{for all } w \in V.$$

By means of the Lax-Milgram theorem there exists a unique solution $u \in V$ for the variational problem: Find $u \in V$ such that

$$a(u,v) = \langle F, v\rangle, \qquad \text{for all } v \in V. \tag{2.3}$$

A function $u \in V$ is called *weak* solution to the problem (2.1) if it satisfies the corresponding variational problem (2.2).

The bilinear form $a(\cdot,\cdot)$ defines a continuous linear operator $A\colon V \longrightarrow V'$ by

$$\langle Au, v\rangle_{V'\times V} = a(u,v) \qquad \text{for all } u \in V,\ v \in V,$$

where V' denotes the dual space of V.

The variational problem (2.3) can be written as an operator equation:

$$Au = F.$$

Theorem 1 (Lax-Milgram)**.** *For any continuous linear functional $F \in V'$, the variational problem* (2.3) *has a unique solution $u \in V$ with*

$$\frac{1}{c_1}\|F\|_{V'} \le \|u\|_V \le \frac{1}{c_2}\|F\|_{V'}.$$

Proof. See e.g. Brenner and Scott [37] □

If additionally $a(\cdot,\cdot)$ is symmetric, i.e.,

$$a(u,v) = a(v,u) \qquad \text{for all } u \in V,\ v \in V,$$

and non-negative, i.e.,

$$a(v,v) \geq 0 \qquad \text{for all } v \in V,$$

then $u \in V$ is solution of the variational problem (2.3), if and only if $u \in V$ is a minimizer of the so-called Ritz functional J given by

$$J(v) = \frac{1}{2}a(v,v) - \langle F, v\rangle,$$

i.e., $J(u) = \min_{v\in V} J(v)$.

2.2.1 Finite Element Method

The main idea of Galerkin's method, in particular of the finite element method, is now to replace the infinite dimensional space V by some finite dimensional space V_h. Here h denotes the discretization parameter and indicates that there should be a convergence to the continuous problem for $h \longrightarrow 0$. In the standard approach (conforming method) the space V_h is a subspace of V and the corresponding finite dimensional problem is just the reduction of the variational problem to: Find $u_h \in V_h$ such that

$$a(u_h, v) = \langle F, v\rangle, \qquad \text{for all } v \in V_h. \tag{2.4}$$

For computing we have to choose a basis $(\varphi_i)_{i=1}^{N_h}$ for the finite element space V_h, where N_h denotes the dimension of V_h. Due to the linearity, it suffices to test only with the basis functions. By means of the basis we can define the system matrix $A \in \mathbb{R}^{N_h\times N_h}$, the load vector $\underline{f} \in \mathbb{R}^{N_h}$ and the solution vector $\underline{u} = (u_i)_{i=1}^{N_h}$ by

$$\begin{aligned} A &= (a(\varphi_i,\varphi_j))_{i,j=1}^{N_h}, \\ \underline{f} &= (\langle F, \varphi_i\rangle)_{i=1}^{N_h}, \\ \underline{u}_h &= \sum_{i=1}^{N_h} u_i\varphi_i. \end{aligned}$$

Using these definitions we can rewrite the problem (2.4) equivalently as the linear system

$$A\underline{u} = \underline{f}. \tag{2.5}$$

We have to analyze how good u_h, the solution of (2.4), approximates u, the solution of (2.3). Céa's Lemma is of fundamental importance for the error analysis.

Lemma 1 (Céa's Lemma)**.** *Given a Hilbert space $(V, (\cdot,\cdot))$, a continuous, coercive bilinear form $a\colon V \times V \longrightarrow \mathbb{R}$ and suppose that $u \in V$ solves the variational problem (2.3) in V. For the solution $u_h \in V_h$ of the finite element variational problem (2.4) we have*

$$\|u - u_h\|_V \leq \frac{c_1}{c_2} \min_{v \in V_h} \|u - v\|_V, \tag{2.6}$$

where c_1 is the continuity constant and c_2 is the coercivity constant of $a(\cdot,\cdot)$ on V.

Proof. A proof can be found in any finite element book, e.g. Brenner and Scott [37] or Braess [23]. □

Remark 1. *Céa's Lemma shows that u_h is quasi-optimal in the sense that the error $\|u - u_h\|_V$ is proportional to the best it can be using the subspace V_h.*

We estimated the discretization error of the variational problem by the best approximation error to the true solution u. The next step in the error analysis is the estimation of the approximation error, which depends on the properties of the space V_h and smoothness properties (regularity properties) of the solution u.

In order to construct the finite element spaces V_h the domain Ω is partitioned into a finite set of simple subdomains $T_{h,i}$. Together they form the triangulation $\mathcal{T}_h = \{T_{h,i}\}$. For simplification we assume that Ω is polygonal. Usual elements are triangles or quadrilaterals in 2D, tetrahedrons, hexahedrons and prisms in 3D. For more details on regular triangulations we refer to Ciarlet [45]. The finite elements are then introduced by defining shape functions, usually polynomials, with local support over the geometrical elements.

We will consider here only the case of continuous piecewise polynomial finite elements. For $t \geq 2$ we have $H^t(\Omega) \subset C^0(\Omega)$. That means, we can define an interpolation operator $I_h\colon H^t(\Omega) \longrightarrow V_h$, where $v_h = I_h v$ is the continuous piecewise polynomial function with

$$v_h(x_i) = v(x_i),$$

where x_i belong to a set of points, such that any function $v_h \in V_h$ is well defined by its values in the points x_i. Additionally, the distribution of these points have to guarantee the continuity of the function v_h, see Figure 2.1 for an example of distributing the nodes, on which the function is evaluated, for linear, quadratic and cubic elements on triangles.

Since we consider here $C^0(\Omega)$-finite elements, which belong only to $H^1(\Omega)$, the higher Sobolev norms are not declared. Therefore we need the so-called broken Sobolev norms based on a given triangulation, which are defined for any $m \in \mathbb{N}$ by

$$\|u\|_{m,h} := \left(\sum_{T \in \mathcal{T}} \|u\|^2_{H^m(T)} \right)^{1/2}.$$

The standard approximation result is presented in the following theorem.

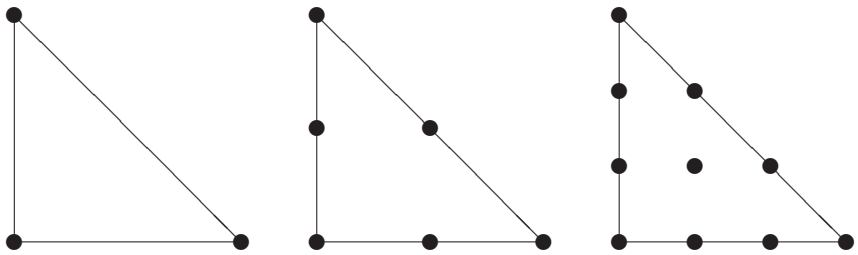


Figure 2.1: nodes of the nodal basis for linear, quadratic and cubic triangle elements

Theorem 2. *Given a quasi-uniform triangulation of Ω and suppose $t \geq 2$. For the interpolation with piecewise polynomials of degree $t-1$, we have*

$$\|u - I_h u\|_{m,h} \leq c \cdot h^{t-m} |u|_{H^t(\Omega)} \quad \text{for } u \in H^t(\Omega), \quad 0 \leq m \leq t. \tag{2.7}$$

Additionally we need regularity properties of the solution $u \in V$. There exist a lot of regularity theorems which require properties of the domain (convex, polygonal, smooth boundary), the right hand side as well as the boundary conditions. For our model problem we have so-called H^2-regularity:

Theorem 3. *If Ω is convex, or has a $C^{1,1}$-boundary, the variational problem (2.2) is H^2-regular, i.e., for every $f \in L^2(\Omega)$, there exists a solution $u \in H^2(\Omega)$ with*

$$\|u\|_{H^2(\Omega)} \leq c \|f\|_{L^2(\Omega)}. \tag{2.8}$$

Proof. A proof can be found in Braess [23]. □

We refer to Grisvard [57] for a whole bunch of similar regularity results, concerning other properties of the domain Ω and boundary conditions.

Both the approximation and regularity results lead then to error estimates in the energy norm, like the following theorem.

Theorem 4. *Given a family of quasi-uniform triangulations $\mathcal{T}_h$ of Ω. For the finite element solution $u_h \in V_h$ of our model problem holds:*

$$\|u - u_h\|_{H^1(\Omega)} \leq ch \|u\|_{H^2(\Omega)} \leq ch \|f\|_{L^2(\Omega)}. \tag{2.9}$$

Proof. The proof relies on the H^2-regularity and the interpolation error estimates for nodal linear elements together with Céa's Lemma, see, e.g. Braess [23]. □

Finally, we are also interested in error estimates in different norms, e.g. the L^2-norm. Due to Theorem 2 we gain one power of h, if we measure the approximation error in the L^2-norm. For proving that this carry over to the discretization error, we need again the H^2-regularity and duality arguments, which are called Aubin-Nitsche-trick.

Theorem 5. *Under the standard assumptions of Theorem 4 we get*

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

If additionally $f \in L^2(\Omega)$ and therefore $u \in H^2(\Omega)$, there holds

$$\|u - u_h\|_{L^2(\Omega)} \leq ch^2\|f\|_{L^2(\Omega)}. \tag{2.10}$$

Proof. See e.g. Braess [23]. □

2.3 Mixed Problems

In this section we consider mixed variational problems, which are variational problems stated in a product of spaces. Existence of solutions of primal variational problems relies on ellipticity of the bilinear form $a(\cdot, \cdot)$. For mixed problems we will need a more abstract result without the ellipticity condition. This abstract setting is given in the famous Theorem of Babuška and Aziz, see Babuška and Aziz [8]. Next we present the abstract setting of mixed variational problems and discuss the theory of Brezzi and the finite element analysis. At the end of this section one classical example is given, namely the Stokes problem.

2.3.1 An Abstract Existence Result

The following theorem is essential for the analysis.

Theorem 6 (Babuška and Aziz). *Let X and Y be real Hilbert spaces. The linear and continuous operator $\mathcal{A}\colon X \longrightarrow Y'$ is an isomorphism if and only if the corresponding bilinear form $a\colon X \times Y \longrightarrow \mathbb{R}$, given by $a(x, y) = \langle \mathcal{A}x, y\rangle_{Y' \times Y}$, satisfies the following conditions:*

1. *(**continuity**). There exists a constant $\mu_1 \geq 0$ with*

$$|a(x, y)| \leq \mu_1\|x\|_X\|y\|_Y \qquad \text{for all } x \in X,\ y \in Y.$$

2. *(**inf-sup-condition**). There exists a constant $\mu_2 > 0$ with*

$$\sup_{0 \neq y \in Y} \frac{a(x, y)}{\|y\|_Y} \geq \mu_2\|x\|_X \qquad \text{for all } x \in X.$$

3. *For each $y \in Y$ with $y \neq 0$ there exists an $x \in X$ with*

$$a(x, y) \neq 0.$$

Proof. The proof is based on the closed range theorem (Yosida [100]) and can be found for example in Braess [23] and Girault and Raviart [53]. □

Remark 2. *The well-known theorem of Lax and Milgram for elliptic problems follows directly from the theorem of Babuška-Aziz: Let $X = Y = V$. Since*

$$\sup_{v \in V} \frac{a(u,v)}{\|v\|_V} \geq \frac{a(u,u)}{\|u\|_V} \geq \mu \|u\|_V$$

(where $\mu > 0$ is the constant of the coercivity condition), the second condition is satisfied. The third condition also follows from the coerciveness: For each $v \neq 0$ we can take $u = v$ such that

$$a(u,v) = a(v,v) \geq \mu \|v\|_V^2 > 0.$$

2.3.2 The Theory of Brezzi

Let X and Q be real Hilbert spaces, $a\colon X \times X \longrightarrow \mathbb{R}$, $b\colon X \times Q \longrightarrow \mathbb{R}$ continuous bilinear forms, and $F\colon X \longrightarrow \mathbb{R}$, $G\colon Q \longrightarrow \mathbb{R}$ continuous linear functionals.

We consider the following mixed variational problem: Find $x \in X$ and $p \in Q$ such that

$$\begin{array}{lcll} a(x,w) + b(w,p) & = & \langle F, w \rangle_{X' \times X} & \text{for all } w \in X, \\ b(x,q) & = & \langle G, q \rangle_{Q' \times Q} & \text{for all } q \in Q. \end{array} \tag{2.11}$$

Here, $\langle F, w \rangle_{X' \times X}$ ($\langle G, q \rangle_{Q' \times Q}$) denotes the evaluation of the linear functional F (G) at the point w (q).

The bilinear forms define linear continuous operators $A\colon X \longrightarrow X'$ by

$$\langle Ax, w \rangle_{X' \times X} = a(x,w), \qquad \forall x \in X, \forall w \in X,$$

$B\colon X \longrightarrow Q'$ and its adjoint operator $B^*\colon Q \longrightarrow X'$ by

$$\langle Bx, q \rangle_{Q' \times Q} = \langle x, B^* q \rangle_{X \times X'} = b(x,q), \qquad \forall x \in X, \forall q \in Q.$$

With these operators we obtain the following system of operator equations:

$$Ax + B^* q = F \quad \text{in } X', \tag{2.12}$$

$$Bx = G \quad \text{in } Q'. \tag{2.13}$$

The mixed variational problem can also be formulated as a non-mixed variational problem on $X \times Q$:

Find $(x,p) \in X \times Q$ such that

$$\mathcal{B}((x,p),(w,q)) = \langle \mathcal{F}, (w,q) \rangle_{(X \times Q)' \times (X \times Q)} \quad \text{for all } (w,q) \in X \times Q \tag{2.14}$$

with the bilinear form

$$\mathcal{B}((x,p),(w,q)) = a(x,w) + b(w,p) + b(x,q)$$

and the linear functional

$$\langle \mathcal{F}, (w,q)\rangle_{(X\times Q)'\times(X\times Q)} = \langle F, w\rangle_{X'\times X} + \langle G, q\rangle_{Q'\times Q}.$$

Under the assumption that a is a symmetric and non-negative bilinear form, the equations (2.11) are the first-order optimality conditions of the saddle point problem:

Find $(x,p) \in X \times Q$ such that

$$\mathcal{L}(x,q) \leq \mathcal{L}(x,p) \leq \mathcal{L}(w,p)$$

with

$$\mathcal{L}(w,q) = \frac{1}{2}a(w,w) + b(w,q) - \langle F, w\rangle_{X'\times X} - \langle G, q\rangle_{Q'\times Q}. \tag{2.15}$$

This is just the Lagrangian for the constrained optimization problem:

Find $x \in X$ such that

$$J(x) = \inf_{Bw=G} J(w)$$

with

$$J(w) = \frac{1}{2}a(w,w) - \langle F, w\rangle.$$

In this context the variable p will be the Lagrange multiplier associated with the constraint $Bx = g$.

In the case of mixed problems existence and uniqueness of a solution follow from the theorem of Brezzi (see Babuška [9], Brezzi [38]).

Theorem 7 (Brezzi). *Let X and Q be real Hilbert spaces, $F\colon X \longrightarrow \mathbb{R}$, $G\colon Q \longrightarrow \mathbb{R}$ continuous linear functionals and $a\colon X\times X \longrightarrow \mathbb{R}$, $b\colon X\times Q \longrightarrow \mathbb{R}$ bilinear forms. Assume that the following conditions are fulfilled:*

1. *(**continuity**). There exist constants $\alpha_1 > 0, \beta_1 > 0$ with*

$$|a(x,w)| \leq \alpha_1 \|x\|_X \|w\|_X \qquad \text{for all } x, w \in X, \text{ and}$$

$$|b(w,q)| \leq \beta_1 \|w\|_X \|q\|_Q \qquad \text{for all } w \in X,\ q \in Q.$$

2. *(**kernel-coercivity**). There exists a constant $\alpha_2 > 0$ such that*

$$a(x,x) \geq \alpha_2 \|x\|_X^2 \quad \text{for all } x \in W = \operatorname{Ker} B = \{x \in X : b(x,q) = 0 \text{ for all } q \in Q\}.$$

3. *(**inf-sup-condition**). There exists a constant β_2 such that*

$$\inf_{0\neq q\in Q} \sup_{0\neq x\in X} \frac{b(x,q)}{\|x\|_X \|q\|_Q} \geq \beta_2 > 0.$$

Then the mixed variational problem:

Find $x \in X$ and $p \in Q$ such that

$$\begin{aligned} a(x,w) + b(w,p) &= \langle F, w\rangle_{X'\times X} \quad \text{for all } w \in X, \\ b(x,q) &= \langle G, q\rangle_{Q'\times Q} \quad \text{for all } q \in Q \end{aligned}$$

has a unique solution. Moreover, we have

$$\begin{aligned} \|x\|_X &\le \frac{1}{\alpha_2}\|F\|_{X'} + \frac{1}{\beta_2}\left(1+\frac{\alpha_1}{\alpha_2}\right)\|G\|_{Q'}, \\ \|p\|_Q &\le \frac{1}{\beta_2}\left(1+\frac{\alpha_1}{\alpha_2}\right)\|F\|_{X'} + \frac{\alpha_1}{\beta_2^2}\left(1+\frac{\alpha_1}{\alpha_2}\right)\|G\|_{Q'}. \end{aligned}$$

Proof. The proof relies again on the closed range theorem and on the Lax-Milgram theorem for the auxiliary variational problem:

Find $z \in W$ such that

$$a(z,w) = \langle F, w\rangle_{X'\times X} - a(x_0, w), \qquad \text{for all } w \in W,$$

where $Bx_0 = G$. The details can be found in Brezzi and Fortin [39]. □

2.4 Mixed Finite Element Methods

Like in the primal case, we replace the infinite dimensional spaces X and Q by some finite dimensional subspaces $X_h \subset X$ and $Q_h \subset Q$. By Galerkin's principle the approximate solutions $x_h \in X_h$ and $p_h \in Q_h$ solve the discrete variational problem

$$\begin{aligned} a(x_h, w) + b(w, p_h) &= \langle F, w\rangle_{X'\times X} \quad \text{for all } w \in X_h, &(2.16)\\ b(x_h, q) &= \langle G, q\rangle_{Q'\times Q} \quad \text{for all } q \in Q_h. &(2.17) \end{aligned}$$

If we choose a basis $\{\phi_1, \ldots, \phi_n\}$ of X_h and a basis $\{\psi_1, \ldots, \psi_m\}$ of Q_h, any element $w_h \in X_h$ and $q_h \in Q_h$ can be represented in the following form:

$$w_h = \sum_{i=1}^{n} w_i \phi_i \quad \text{and} \quad q_h = \sum_{j=1}^{m} q_j \psi_j.$$

From the discrete variational problem a linear system of equations is obtained:

$$\begin{pmatrix} A_h & B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \underline{x}_h \\ \underline{p}_h \end{pmatrix} = \begin{pmatrix} \underline{f}_h \\ \underline{g}_h \end{pmatrix}$$

with

$$\begin{aligned} A_h &= (a(\phi_i, \phi_j))_{i,j=1,\ldots,n}, \\ B_h &= (b(\phi_i, \psi_j))_{i=1,\ldots,n;j=1,\ldots,m}, \\ \underline{x}_h &= (x_i)_{i=1,\ldots,n}, \quad \underline{p}_h = (p_j)_{j=1,\ldots,m}, \\ \underline{f}_h &= (\langle F, \phi_i\rangle_{X'\times X})_{i=1,\ldots,n}, \quad \underline{g}_h = (\langle G, \psi_j\rangle_{Q'\times Q})_{j=1,\ldots,m}. \end{aligned}$$

The discrete problem (2.16) can be analyzed in the same way as the continuous problem (2.11). Since we have chosen $X_h \subset X$ and $Q_h \subset Q$, the continuity conditions of Brezzi's Theorem still hold. But observe that $\operatorname{Ker} B_h \not\subset \operatorname{Ker} B$, in general. That means the coercivity of a on $\operatorname{Ker} B$ does not necessarily imply the coercivity of a on $\operatorname{Ker} B_h$. Similarly, the inf-sup-condition of the continuous problem does not necessarily imply the inf-sup-condition of the discrete problem. These two conditions must be explicitly verified for the particular choice of subspaces X_h and Q_h.

The next theorem is a generalization of Céa's lemma to mixed problems:

Theorem 8. *Let the assumptions of the theorem of Brezzi (7) be fulfilled. For finite dimensional subspaces $X_h \subset X$, $Q_h \subset Q$ assume that additionally the following conditions hold:*

1. *(***discrete coercivity***). There exists a constant $\tilde{\alpha}_2 > 0$ with*

$$a(w_h, w_h) \geq \tilde{\alpha}_2 \|w_h\|_X^2 \quad \text{for all } w_h \in W_h,$$

   where $W_h = \operatorname{Ker} B_h = \{w_h \in X_h \colon b(w_h, q_h) = 0 \text{ for all } q_h \in Q_h\}$.

2. *(***discrete inf-sup-condition***). There exists a constant $\tilde{\beta}_2 > 0$ such that*

$$\sup_{0 \neq w_h \in X_h} \frac{b(w_h, q_h)}{\|w_h\|_X} \geq \tilde{\beta}_2 \|q_h\|_Q \quad \text{for all } q_h \in Q_h.$$

Then the problem (2.16) has a unique solution $(x_h, p_h) \in X_h \times Q_h$ and:

$$\begin{aligned}
\|x - x_h\|_X &\leq \left(1 + \frac{\alpha_1}{\tilde{\alpha}_2}\right)\left(1 + \frac{\beta_1}{\tilde{\beta}_2}\right) \inf_{w_h \in X_h} \|x - w_h\|_X + \frac{\beta_1}{\tilde{\alpha}_2} \inf_{q_h \in Q_h} \|p - q_h\|_Q, \\
\|p - p_h\|_Q &\leq \left(1 + \frac{\alpha_1}{\tilde{\alpha}_2}\right)\left(1 + \frac{\beta_1}{\tilde{\beta}_2}\right) \frac{\alpha_1}{\tilde{\beta}_2} \inf_{w_h \in X_h} \|x - w_h\|_X + \\
&\quad + \left[1 + \frac{\beta_1}{\tilde{\beta}_2}\left(1 + \frac{\alpha_1}{\tilde{\alpha}_2}\right)\right] \inf_{q_h \in Q_h} \|p - q_h\|_Q,
\end{aligned}$$

where $(x, p) \in X \times Q$ denotes the unique solution of the continuous problem (2.11).

Proof. The existence and uniqueness of the solution $(x_h, p_h) \in X_h \times Q_h$ follow from the theorem of Brezzi.

For the proof of the error estimates, the Galerkin orthogonality is used to construct an auxiliary problem on which the theorem of Brezzi is applied to. Details can be found in Brezzi and Fortin [39]. □

Remark 3. *If the constants $\tilde{\alpha}_2$ and $\tilde{\beta}_2$ are independent of the discretization parameter h, the discretization error goes to 0 for $h \to 0$ of the same order as the approximation error does.*

2.4.1 The Stokes Problem

By means of the Stokes problem as an example of a mixed problem, we briefly discuss stability analysis and the mixed finite element method. More details on this topic can be found, e.g. in Pironneau [74] and Braess [23].

The Stokes equations of a stationary flow of an incompressible fluid in a domain Ω are given by:

$$\begin{aligned} -\nu\Delta u(x) + \nabla p(x) &= f(x) && x \in \Omega, \\ \operatorname{div} u(x) &= 0 && x \in \Omega, \\ u(x) &= 0 && x \in \partial\Omega, \end{aligned}$$

where u denotes the velocity field, p the static pressure, and f some external force.

Now let

$$\begin{aligned} V &:= (H_0^1(\Omega))^d \qquad \text{(where } d \text{ is the dimension of } \Omega \subset \mathbb{R}^d\text{)}, \\ Q &:= L_0^2(\Omega) = \left\{ q \in L^2(\Omega) : \int_\Omega q \,\mathrm{d}x = 0 \right\}. \end{aligned}$$

The variational formulation reads as follows: Find $u \in V$ and $p \in Q$ such that

$$\begin{aligned} a(u,v) + b(v,p) &= \langle F, v\rangle \quad \text{for all } v \in V, && (2.18) \\ b(u,q) &= 0 \quad \text{for all } q \in Q, && (2.19) \end{aligned}$$

with the bilinear forms

$$\begin{aligned} a(u,v) &= \nu \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x, \\ b(u,q) &= \int_\Omega q \operatorname{div} u \,\mathrm{d}x, \end{aligned}$$

and the linear form

$$\langle F, v\rangle = \int_\Omega f v \,\mathrm{d}x.$$

For stability we have to check, if the conditions of Brezzi's Theorem are fulfilled. The bilinear form a is even $H_0^1(\Omega)$-elliptic. It remains to prove, that the bilinear form b fulfills the inf-sup-condition. This was for instance done in Girault and Raviart [53].

Now we replace the infinite dimensional spaces V and Q by some finite dimensional subspaces $V_h \subset V$ and $Q_h \subset Q$. Then we can formulate the discrete version: Find $(u_h, p_h) \in V_h \times Q_h$ such that

$$\begin{aligned} a(u_h,v) + b(v,p_h) &= \langle F, v\rangle \quad \text{for all } v \in V_h, && (2.20) \\ b(u_h,q) &= 0 \quad \text{for all } q \in Q_h. && (2.21) \end{aligned}$$

The aim is to find appropriate spaces V_h and Q_h such that the discrete problem is stable, i.e., the discrete kernel-ellipticity and the discrete inf-sup-condition have to be satisfied.

In literature, e.g., GIRAULT AND RAVIART [53], BRAESS [23] or BREZZI AND FORTIN [39] there are examples of finite elements, which do not fulfill the inf-sup-condition and, therefore, lead to instabilities (e.g. checkerboard effects).

We will briefly present two stable finite elements for the Stokes equations. The first one is the Taylor-Hood element (or $P_2 - P_1$ element), specified by (for a given triangulation $\mathcal{T}_h$, based on triangular (2D) or tetrahedral (3D) elements):

$$\begin{aligned} V_h &= \left\{u \in C(\bar{\Omega})^d \cap H_0^1(\Omega)^d;\ u|_T \in \mathcal{P}_2 \quad \text{for all } T \in \mathcal{T}_h\right\} \\ Q_h &= \left\{q \in C(\Omega) \cap L_0^2(\Omega);\ q|_T \in \mathcal{P}_1 \quad \text{for all } T \in \mathcal{T}_h\right\}. \end{aligned}$$

An element $(v_h, q_h) \in V_h \times Q_h$ is uniquely determined by the values of the components of v_h on the nodes and on the midpoints of the edges of the elements and the values of q_h on the nodes of the elements. The inf-sup-condition is shown in BREZZI AND FORTIN [39] and one obtains the following approximation result:

$$\|u - u_h\|_{H^1(\Omega)} + \|p - p_h\|_{L^2(\Omega)} \leq Ch^m(|u|_{H^{m+1}(\Omega)} + |p|_{H^m(\Omega)}), \quad \text{for } m = 1, 2.$$

As an example of a nonconforming finite element, i.e. the finite dimensional spaces are not subsets of the continuous ones, we mention the Crouzeix-Raviart element, where nonconforming velocity components are used. We define the discrete spaces as follows:

$$\begin{aligned} V_h := \{&v :\ v|_T \in \mathcal{P}_1 \quad \text{for all } T \in \mathcal{T}_h \\ &v \text{ is continuous at the midpoints of the edges (2D) / faces (3D)}, \\ &v = 0 \text{ at all midpoints of boundary edges / faces}\} \end{aligned}$$

and

$$Q_h := \{q \in Q :\ q|_T \in \mathcal{P}_0 \quad \text{for all } T \in \mathcal{T}_h\}.$$

An element $(v_h, q_h) \in V_h \times Q_h$ is determined by the velocity values at the edge- / face-centers and the pressure values at the element centers. More details can be found in CROUZEIX AND RAVIART [47], including the following approximation result:

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch(|u|_{H^2(\Omega)} + |p|_{H^1(\Omega)})$$

for $(u, p) \in H^2(\Omega) \times H^1(\Omega)$.

The name "Crouzeix-Raviart element" is also used for the nonconforming divergence-free $\mathcal{P}_1$ element, where the finite element itself is already divergence free, such that the pressure can be eliminated from the equations, see e.g. CROUZEIX AND RAVIART [47] or BRAESS [23].

2.5 Iterative Solvers

In this section we give an introduction to iterative solution methods for linear systems arising from finite element methods. In the first part we consider elliptic problems like the

model problem and introduce the most basic iterative method, the Richardson method. For a detailed discussion of iterative methods we refer to e.g. AXELSSON [7], SAAD [79] and HACKBUSCH [60]. We also briefly discuss the concept of preconditioning.

The second part is then devoted to iterative solvers for saddle point systems. We explicitly exclude here multigrid methods, because they will be the subject of the next chapter. We will restrict ourselves to symmetric saddle point systems, since all problems we will consider in this work fit in this class of problems.

2.5.1 Richardson Method and Preconditioning

We will now present the Richardson method for solving a linear system of the form

$$Ax = b, \tag{2.22}$$

where $A \in \mathbb{R}^{N\times N}$ is symmetric and positive definite. Note that the solution of (2.22) is the minimizer of the function

$$\phi(x) = \frac{1}{2}x^T Ax - x^T b.$$

The Richardson method is defined by the following iteration:

$$x^{(k+1)} = x^{(k)} + \omega(f - Ax^{(k)}), \tag{2.23}$$

with a positive relaxation parameter ω. This is just one step in the direction of the negative gradient of $\phi(x)$, which is the direction of the steepest descent, with the fixed steplength ω.

Let us denote the smallest and largest eigenvalue of the matrix A by $\lambda_{\min}$ and $\lambda_{\max}$, respectively. The Richardson method (2.23) converges for all ω such that

$$0 < \omega < \frac{2}{\lambda_{\max}},$$

and the convergence rate is given by the spectral radius of the iteration matrix $M = I - \omega A$:

$$\rho(M) = \max\{|1 - \omega\lambda_{\min}|, |1 - \omega\lambda_{\max}|\}. \tag{2.24}$$

From (2.24) it is possible to determine the optimal relaxation parameter ω_{opt} such that the convergence rate is minimzed. The optimal convergence rate is achieved by choosing

$$\omega_{\text{opt}} = \frac{2}{\lambda_{\max} + \lambda_{\min}}$$

and is then given by

$$\rho_{\text{opt}}(M) = \frac{\lambda_{\max} - \lambda_{min}}{\lambda_{\max} + \lambda_{\min}}.$$

By means of the condition number

$$\kappa(A) := \frac{\lambda_{\max}}{\lambda_{\min}},$$

the optimal convergence rate is

$$\rho_{\text{opt}}(M) = \frac{\kappa(A) - 1}{\kappa(A) + 1}.$$

Typically, the condition number becomes very large as the discretization parameter h tends to zero, e.g., in our model problem, the condition number behaves like $\mathcal{O}(h^{-2})$. The convergence can be improved by using the preconditioned Richardson method:

$$x^{(k+1)} = x^{(k)} + \hat{A}^{-1}(f - Ax^{(k)}), \tag{2.25}$$

where $\hat{A}$ is the preconditioner of A. The performance of the preconditioner $\hat{A}$ depends on the numerical effort to solve a linear system of the form

$$\hat{A}w = r,$$

which has to be solved in every iteration step, and on the condition number of $\hat{A}^{-1}A$. If possible, the condition number should be independent of the mesh parameter h, i.e., $\kappa(\hat{A}^{-1}A) = \mathcal{O}(1)$.

The efficiency can be further improved by using the conjugate gradients method (cg), see Hestenes and Stiefel [63]. In comparison to the Richardson method the convergence rate is accelerated by the square root: The error in the k-th iteration in the A-energy norm $\|e^{(k)}\|_A$ is bounded by

$$\|e^{(k)}\|_A \leq c^k \frac{1}{1 + c^{2k}} \|e^{(0)}\|_A,$$

with

$$c := \frac{\sqrt{\kappa(\hat{A}^{-1}A)} - 1}{\sqrt{\kappa(\hat{A}^{-1}A)} + 1}.$$

For more details see e.g. Jung and Langer [66].

2.5.2 Solution Methods for Indefinite Systems

We assume now, that we already have a discretized system in saddle point form:

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \tag{2.26}$$

Throughout this section we will assume that A is symmetric and positive definite, C is symmetric and positive semidefinite. Since A is nonsingular, there exists for the saddle point matrix

$$\mathcal{K} = \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}$$

the following block triangular factorization:

$$\mathcal{K} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -S \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix}, \tag{2.27}$$

where

$$S = C + BA^{-1}B^T$$

is the so-called (negative) Schur complement. From the factorization (2.27) it directly follows that $\mathcal{K}$ is nonsingular if and only if S is non-singular. The Schur complement S is symmetric positive definite (hence, invertible) if and only if $\operatorname{Ker} C \cap \operatorname{Ker} B^T = \{0\}$. Obvious sufficient conditions therefore are that C is positive definite or that B^T has full column rank.

Note that the requirements on the block matrices can be relaxed. For more details, we refer to the nice survey paper Benzi, Golub and Liesen [16].

The iterative methods presented so far are based on positive definite systems, nevertheless these methods are also interesting for indefinite systems as long as it is possible to find preconditioner in such a way that the preconditioned system is positive definite with respect to a certain scalar product. Then the cg-method can be applied, see e.g., Bramble and Pasciak [28] and the more recent work Zulehner [102].

For indefinite systems there also exist other Krylov subspace methods, see Saad [79] for an overview. Suitable candidates are GMRES (c.f. Saad and Schultz [80]), QMR (c.f. Freund and Nachtigal [52]) and MINRES (c.f. Paige and Saunders [72]). As always, in order to obtain efficient methods, efficient preconditioning is needed. There exists a lot of literature concerning preconditioners for mixed problems, see e.g., Rusten and Winther [78], Vassilevski and Lazarov [93], Golub, Greif and Varah [56], and Golub and Greif [55], which is by far not a complete list. A well known class of iteration methods for saddle point problems are Uzawa methods, which are the topic of the next section.

2.5.3 (Inexact) Uzawa Methods

The first iterative schemes for the solution of saddle point problems were the ones developed by the mathematical economists Arrow, Hurwicz and Uzawa, see Arrow, Hurwicz and Uzawa [3]. We start with the classical Uzawa method, see Uzawa [91]. Given an initial guess $p^{(0)}$ for p, Uzawa's method consists of the following coupled iteration:

$$\begin{aligned} x^{(k+1)} &= A^{-1}(f - B^T p^{(k)}), \\ p^{(k+1)} &= p^{(k)} + \omega(Bx^{(k+1)} - Cp^{(k)} - g), \end{aligned} \tag{2.28}$$

where $\omega > 0$ is a relaxation parameter.

If we use the first equation in (2.28) to eliminate $x^{(k+1)}$ from the second one, we obtain

$$p^{(k+1)} = p^{(k)} + \omega(BA^{-1}f - g - (C + BA^{-1}B^T)p^{(k)}), \tag{2.29}$$

which is the classical Richardson method applied to the Schur complement system

$$(C + BA^{-1}B^T)p = BA^{-1}f - g. \tag{2.30}$$

The great disadvantage of the classical Uzawa method is the necessity of the exact solving of the linear system with coefficient matrix A. Often there is no need to solve the system exactly, but it can be replaced by an inner inexact solver, see e.g. Bramble, Pasciak and Vassilev [29], Chen [42], Cheng and Zou [43], Elman and Golub [50], Zulehner [102].

The Arrow-Hurwicz method, see Arrow and Hurwicz [4], is an alternative to the Uzawa method. The iterate $x^{(k+1)}$ given by the first step of (2.28) is the minimizer of the objective function

$$\phi(x) = \frac{1}{2}x^T Ax - x^T(f - B^T p^{(k)}).$$

The solving of the linear system with the matrix A is now replaced by taking one step in the direction of the negative gradient of $\phi(x)$ (which is the direction of the steepest descent) with a fixed step length τ. Given the initial guesses $x^{(0)}$ and $p^{(0)}$, the Arrow-Hurwicz method consists of the following iteration:

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \tau(f - Ax^{(k)} - B^T p^{(k)}), \\ p^{(k+1)} &= p^{(k)} + \omega(Bx^{(k+1)} - Cp^{(k)} - g). \end{aligned} \tag{2.31}$$

The convergence of this method depends on the two relaxation parameters, τ and ω. For convergence results and estimates for the optimal choice of the relaxation parameters, we refer to e.g. Astrakhantsev [6], Fortin and Glowinski [51], Langer and Queck [68], Queck [75].

The convergence of these methods can be improved by using appropriate preconditioners. If we use the classical Richardson method for the Schur complement system (2.30) with an appropriate preconditioner $\hat{S}$ for the Schur complement $S = C + BA^{-1}B^T$, the iterative method reads

$$p^{(k+1)} = p^{(k)} + \hat{S}^{-1}(BA^{-1}f - g - Sp^{(k)}),$$

or in the original form:

$$\begin{aligned} x^{(k+1)} &= A^{-1}(f - B^T p^{(k)}), \\ p^{(k+1)} &= p^{(k)} + \hat{S}^{-1}(Bu^{(k+1)} - Cp^{(k)} - g), \end{aligned} \tag{2.32}$$

which is then called the preconditioned Uzawa method.

In order to avoid the computation of the exact solution of the linear system with coefficient matrix A, we use one step of a preconditioned Richardson method for determining the next iterate $x^{(k+1)}$:

$$x^{(k+1)} = x^{(k)} + \hat{A}^{-1}(f - Ax^{(k)} - B^T p^{(k)}), \tag{2.33}$$

where $\hat{A}$ is an appropriate preconditioner for A. Combining (2.33) and the second step of (2.32), we obtain the so-called preconditioned Arrow-Hurwicz method:

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \hat{A}^{-1}(f - Ax^{(k)} - B^T p^{(k)}), \\ p^{(k+1)} &= p^{(k)} + \hat{S}^{-1}(Bx^{(k+1)} - Cp^{(k)} - g). \end{aligned} \tag{2.34}$$

Simple calculation shows that (2.34) can be rewritten in the following form:

$$\begin{aligned} \hat{A}(x^{(k+1)} - x^{(k)}) &= f - Ax^{(k)} - B^T p^{(k)}, \\ B(x^{(k+1)} - x^{(k)}) - \hat{S}(p^{(k+1)} - p^{(k)}) &= g - Bx^{(k)} + Cp^{(k)}, \end{aligned}$$

or equivalently,

$$\begin{pmatrix} x^{(k+1)} \\ p^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ p^{(k)} \end{pmatrix} + \hat{\mathcal{K}}^{-1} \left[\begin{pmatrix} f \\ g \end{pmatrix} - \mathcal{K} \begin{pmatrix} x^{(k)} \\ p^{(k)} \end{pmatrix} \right]$$

with

$$\hat{\mathcal{K}} = \begin{pmatrix} \hat{A} & 0 \\ B & -\hat{S} \end{pmatrix}.$$

That means, the preconditioned Arrow-Hurwicz method is nothing else but the preconditioned Richardson method for (2.26) with the block triangular preconditioner $\hat{\mathcal{K}}$.

Remark 4. *The preconditioned Arrow-Hurwicz method* (2.34) *includes the classical Uzawa method* (2.28) *with the setting*

$$\hat{A} = A, \qquad \hat{S} = \frac{1}{\omega} I,$$

and also the classical Arrow-Hurwicz method (2.31) *with the setting*

$$\hat{A} = \frac{1}{\tau} I, \qquad \hat{S} = \frac{1}{\omega} I.$$

Observe that the preconditioner $\hat{\mathcal{K}}$ is formally obtained by replacing A by $\hat{A}$ and S by $\hat{S}$ in the first factor of (2.27). If we treat the second factor analogously, a different preconditioner $\hat{\mathcal{K}}$ is obtained:

$$\hat{\mathcal{K}} = \begin{pmatrix} \hat{A} & 0 \\ B & -\hat{S} \end{pmatrix} \begin{pmatrix} I & \hat{A}^{-1}B^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} \hat{A} & B^T \\ B & B\hat{A}^{-1}B^T - \hat{S} \end{pmatrix}. \tag{2.35}$$

The resulting preconditioned Richardson method requires the solution of the linear system with coefficient matrix $\hat{\mathcal{K}}$, which can be realized in the following three steps:

$$\begin{aligned} \hat{A}(\hat{x}^{(k+1)} - x^{(k)}) &= f - Ax^{(k)} - B^T p^{(k)}, \\ \hat{S}(p^{(k+1)} - p^{(k)}) &= B\hat{x}^{(k+1)} - Cp^{(k)} - g, \\ \hat{A}(x^{(k+1)} - x^{(k)}) &= f - Ax^{(k)} - B^T p^{(k+1)}. \end{aligned} \tag{2.36}$$

The preconditioner $\hat{\mathcal{K}}$ (2.35) is symmetric (and indefinite), therefore the method is often called the inexact symmetric Uzawa method.

For the convergence analysis of the preconditioned variants of the Uzawa method and the Arrow-Hurwicz method, we refer to e.g., Elman and Golub [50], Langer and Queck [67, 68] and Bank, Welfert and Yserentant [12] for the symmetric inexact Uzawa method.

There exist also Uzawa-type methods for nonsymmetric saddle point problem, see Bramble, Pasciak and Vassilev [30].

Chapter 3

Multigrid Methods

3.1 Motivation

Multigrid methods belong to the class of optimal order methods for solving linear systems arising from a discretization technique like the finite element method or the finite difference method. The convergence speed does not deteriorate when the discretization is refined, whereas classical iterative methods slow down for decreasing mesh size. For a wide overview about multigrid methods we recommend the books Hackbusch and Trottenberg [61], Hackbusch [59], Bramble [27] and Trottenberg, Oosterlee and Schüller [90].

The starting point of the multigrid idea is the observation that classical iteration methods (e.g. Jacobi and Gauss-Seidel schemes) have smoothing properties. Although these methods are characterized by poor global convergence rates, for errors whose length scales are comparable to the mesh size, they provide rapid damping, leaving behind smooth, longer wave-length errors. These smooth parts of the error are responsible for the poor convergence. A (geometric) multigrid method involves a hierarchy of meshes and related discretizations. A quantity that is smooth on a certain grid can, without any essential loss of information, also be approximated on a coarser grid. So the low-frequency error components can be effectively reduced by a coarse-grid correction procedure. Because the action of a smoothing iteration leaves only smooth error components, it is possible to represent them as the solution of an appropriate coarser system. Once this coarser problem is solved, its solution is interpolated back to the fine grid to correct the fine grid approximation for its low-frequency errors.

We will consider in this work only linear multigrid methods, since our applications are linear ones. However, multigrid algorithms are also suited for nonlinear boundary value problems. There are two different approaches. The first one is based on a global linearization method like Newton's iteration to the nonlinear problem. In each iteration step a linear problem has to be solved, which can be done by applying the linear multigrid solver. In the second approach the construction of a smoothing iteration and a coarse grid iteration is directly adapted to the nonlinear case. This leads to the so-called full approximation scheme (FAS), introduced by Brandt, see Brandt [32], Hackbusch [59]

for more details.

We should also mention that there exist situations, where the geometric multigrid method can not be applied, e.g., if the discretization provides no hierarchy of finite element meshes (which is the case for many finite element codes, especially commercial ones), or, if the coarsest grid is still too large to be solved efficiently by a direct or classical iterative solver. The algebraic multigrid method (AMG) provides an attractive multilevel variant whenever geometric multigrid is either too difficult to apply or can not be used at all. The most important conceptual difference between geometric and algebraic multigrid is the following: Geometric approaches employ fixed grid hierarchies and an efficient interplay between smoothing and coarse grid correction has to be ensured by constructing appropriate smoothing iterations. In contrast to this, AMG fixed the smoother to a simple relaxation scheme and enforces an efficient interplay with the coarse grid correction by adjusting the coarsening, i.e., by choosing the coarser levels and interpolation appropriately. For more details we refer to STÜBEN [86] and REITZINGER [76].

Finally, an initial approximation for iterative methods like multigrid can be obtained by nested iteration. The idea is to provide an initial guess on the finest grid by computing and interpolating approximations on coarser grids. The efficiency of multigrid can be improved if it is combined with the nested iteration idea. This combination is called full multigrid (FMG), see BRANDT [32], and is the most efficient multigrid method.

3.2 The Multigrid Algorithm

In this section we present the basic concepts of the multigrid method for solving a linear system of equations

$$\mathcal{K}_h \underline{u}_h = \underline{f}_h, \tag{3.1}$$

arising from a finite element discretization of a boundary value problem. Here, h denotes the discretization parameter.

The most essential ingredient of multigrid methods is the smoothing operator $\mathcal{S}$, which reduces the high frequency error components. Most of the computational effort is usually spent in the smoothing steps. Moreover, using the wrong smoother will destroy the efficiency of the entire multigrid algorithm. Typically, a proper smoother takes the special structure of the underlying problem into account. For finite element discretizations with Lagrange finite elements for scalar elliptic PDEs of second order, the classical damped Jacobi or Gauss-Seidel iteration is an appropriate choice. Nevertheless, for more advanced problems, e.g., indefinite problems, the definition of an efficient smoother is far from obvious. We will discuss several choices of smoother for saddle point problems in more detail in section 3.4.

Let $\underline{u}_h^{\text{old}}$ be some given approximation to the exact solution $\underline{u}_h = \mathcal{K}_h^{-1} \underline{f}_h$. Applying m steps of the smoothing method will result in an approximation

$$\bar{\underline{u}}_h = \mathcal{S}_h^m(\underline{u}_h^{\text{old}}, \underline{f}_h).$$

If the smoothing iteration is well chosen, the error

$$\underline{v}_h = \bar{\underline{u}}_h - \underline{u}_h$$

represents a smooth grid function. Obviously the original equation (3.1) is equivalent to the defect equation

$$\mathcal{K}_h \underline{v}_h = \underline{d}_h, \tag{3.2}$$

with $\underline{d}_h = \mathcal{K}_h \bar{\underline{u}}_h - \underline{f}_h$. The important difference is that the solution $\underline{v}_h$ can be approximated better by means of a coarser grid than $\underline{u}_h$, since $\underline{v}_h$ represents a smooth function.

The goal is to approximate the problem (3.2) by a coarse grid equation

$$\mathcal{K}_H \underline{v}_H = \underline{d}_H. \tag{3.3}$$

In geometric multigrid the matrix $\mathcal{K}_H$ is already defined by the underlying bilinearform and the basis functions of the coarse finite dimensional space V_H. We define $\underline{d}_H$ as the restriction of the fine grid defect to the coarse grid, that is $\underline{d}_H := I_h^H \underline{d}_h$, by means of a suitable linear restriction operator

$$I_h^H \colon \mathbb{R}^{n_h} \longrightarrow \mathbb{R}^{n_H},$$

where n_h and n_H denote the dimension of the fine grid space V_h and of the coarse grid space V_H, respectively. Reasonably one expects that the exact solution $\underline{v}_H$ of the coarse grid equation (3.3) is an approximation of the fine grid solution $\underline{v}_h$. We interpolate the coarse grid function by

$$\tilde{\underline{v}}_h = I_H^h \underline{v}_H,$$

where I_H^h is a linear operator $I_H^h \colon \mathbb{R}^{n_H} \longrightarrow \mathbb{R}^{n_h}$, called the prolongation or coarse-to-fine operator.

Since $\underline{u}_h = \bar{\underline{u}}_h - \underline{v}_h$ is the exact solution and $\tilde{\underline{v}}_h$ is supposed to approximate $\underline{v}_h$, we update the value $\bar{\underline{u}}_h$ by

$$\underline{u}_h^{\text{new}} = \bar{\underline{u}}_h - \tilde{\underline{v}}_h.$$

All these steps can be condensed into the so-called coarse grid correction step

$$\underline{u}_h^{\text{new}} = \bar{\underline{u}}_h - I_H^h \mathcal{K}_H^{-1} I_h^H \left(\mathcal{K}_h \bar{\underline{u}}_h - \underline{f}_h \right). \tag{3.4}$$

For symmetry reasons and due to the fact, that the interpolation may introduce high frequent errors, it is often convenient to complete the two-grid iteration by applying m_2 postsmoothing steps after the coarse grid correction.

After these preliminaries we are able to state the two-grid iteration for solving (3.1), as in Algorithm 3.1. The two-grid iteration is a combination of a smoothing iteration, which is quite efficient in reducing the high-frequency components, and the coarse-grid iteration, which reduces the smooth error very well. Although both components by themselves converge slowly or not at all, the combination leads to a rapidly convergent method.

Algorithm 3.1 Two-grid method for solving $\mathcal{K}_h \underline{u}_h = \underline{f}_h$

Given iterate $\underline{u}_h^{(j)}$
Given m_1: number of presmoothing steps, m_2: number of postsmoothing steps

while (not converged) **do**
// presmoothing step
$\bar{\underline{u}}_h := \mathcal{S}_h^{m_1}(\underline{u}_j^{(j)}, \underline{f}_h)$
// coarse grid correction
$\underline{d}_h := \mathcal{K}_h \bar{\underline{u}}_h - \underline{f}_h$ // calculation of the defect
$\underline{d}_H := I_h^H \underline{d}_h$ // restriction of the defect
$\underline{v}_H := \mathcal{K}_H^{-1} \underline{d}_H$ // solution of the coarse grid equation
$\tilde{\underline{u}}_h := \bar{\underline{u}}_h - I_H^h \underline{v}_H$ // correction of $\bar{\underline{u}}_h$
// postsmoothing step
$\underline{u}_h^{(j+1)} := \mathcal{S}_h^{m_2}(\tilde{\underline{u}}_h)$
end while

Nevertheless, the two-grid iteration is impracticable because the exact solution of the coarse-grid equation is required. The system on the coarse grid is easier to solve than the one on the fine grid. By choosing a coarse-grid meshsize $H = 2h$ and uniform refinement, the number of unknowns decreases about to a quarter (in two dimensions). In practical applications we are interested in solving fine grid equations with millions of unknowns, therefore, the coarse-grid equation has still a too large complexity, such that the overall two-grid method is not optimal with respect to the computational effort. Since $\tilde{\underline{v}}_h = I_H^h \underline{v}_H$ is only an approximation of the correction $\underline{v}_h$, it is also not necessary to solve the coarse grid equation exactly. So we have to replace the exact solve by an iterative process. Since the coarse-grid system is of the same form as the original problem, the two-grid iteration can be used as an iterative solver on the coarse grid (with meshsize H) by introducing a next coarser grid. This can be repeated recursively until we reach a coarsest grid, where the problem size is small enough and on which the problem has to be solved exactly.

Let us consider a hierarchy of $L + 1$ meshes with corresponding finite element spaces $V_0 \subset \cdots \subset V_L$, mesh sizes $h_0 > \cdots > h_L$ and number of unknowns $n_0 < \cdots < n_L$. That means, we want to solve the problem (3.1) on the level $l = L$ and level $l = 0$ corresponds to the coarsest grid. The recursive multigrid procedure (MGM) on level l is then given by Algorithm 3.2. One iteration step of the multigrid method consists of a single call of the MGM procedure on level L.

It is sufficient to choose a small number of iteration steps on the coarser level $l - 1$ in order to get a good approximation of the coarse-grid solution. The sequence of operations during one step of the multigrid iteration for $\gamma = 1$ and $\gamma = 2$ is depicted in Figure 3.1. The arrows indicate the inter grid transfer, the $'\circ'$ represent the smoothing steps and $'\bullet'$ the exact solve on the coarsest level. Due to the form of Figure 3.1 the iteration with $\gamma = 1$ is also called a V-cycle and with $\gamma = 2$ a W-cycle, respectively.

Algorithm 3.2 Multigrid iteration for solving $\mathcal{K}_h \underline{u}_h = \underline{f}_h$

Given μ: number of MGM calls on level l
Given m_1 and m_2: number of pre- and post-smoothing steps

if $l == 0$ **then**
  $u := \mathcal{K}_0^{-1} f$ // exact solve on coarsest level
  return
else
  $u := \mathcal{S}_l^{m_1}(u, f)$ // presmoothing step
  $d := I_l^{l-1}(\mathcal{K}_l u - f)$ // restriction of the defect
  $v := 0$ // starting value
  for $j := 1, \ldots, \mu$ **do**
    MGM$(l-1, v, d)$ // recursive calls of MGM for coarse grid correction
  end for
  $u := u - I_{l-1}^{l} v$ // update on level l
  $u := \mathcal{S}_l^{m_2}(u, f)$ // postsmoothing step
end if

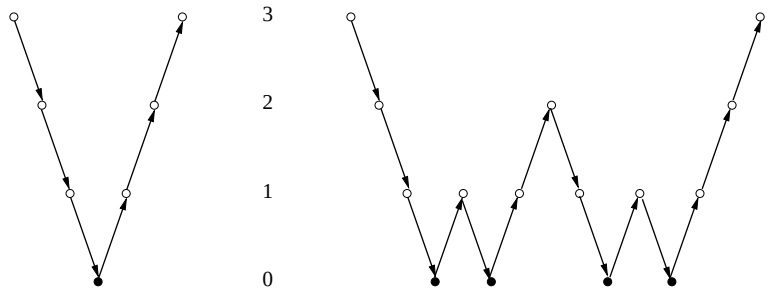


Figure 3.1: V-cycle and W-cycle

3.3 Basic Convergence Analysis

The standard techniques for proving the convergence of multigrid methods rely on convergence results of the corresponding two-grid method. Firstly, we present two approaches for proving the two-grid convergence based on different splittings of the corresponding iteration matrix. The multigrid convergence follows then by a perturbation argument, at least for the W-cycle.

3.3.1 Two-Grid Convergence

The two-grid iteration is a linear iteration having a representation

$$\underline{u}_h^{(j+1)} = M_h^H \underline{u}_h^{(j)} + \mathcal{N}_l \underline{f}_l$$

where the iteration matrix M_h^H depends on the number of (pre-)smoothing steps m (we consider here only the case $m_2 = 0$, i.e., only presmoothing) and is given by

$$M_h^H = (I - I_H^h \mathcal{K}_H^{-1} I_h^H) S_h^m.$$

The goal is then to estimate the contraction number

$$\sigma(M_h^H) = \|M_h^H\| = \sup_{\underline{v}_h \neq 0} \frac{\|M_h^H \underline{v}_h\|}{\|\underline{v}_h\|}.$$

The choice of the appropriate norms is a crucial point in order to get optimal convergence results. Since this choice will be problem-dependent, we keep this general notation $\|\cdot\|$ with the knowledge, that the appropriate norms have still to be specified. There exist two classical approaches to assure two-grid convergence by different splittings of the two-grid iteration operator.

The first technique, mainly developed by the Russian school, is based on a sum splitting, see Bachvalov [10], Astrakhantsev [5]. Here the identity, decomposed into the orthogonal sum $I = P_h^{\text{low}} \oplus P_h^{\text{high}}$ of projections on the subspaces spanned by the low and high frequency eigenvectors of the system matrix, is inserted into M_H^h

$$M_H^h = (I - I_H^h \mathcal{K}_H^{-1} I_h^H)(P_h^{\text{low}} \oplus P_h^{\text{high}}) S_h^m.$$

Two-grid convergence is then proven by showing that

$$\underbrace{\|(I - I_H^h \mathcal{K}_H^{-1} I_h^H) P_h^{\text{low}}\|}_{\text{small}} \underbrace{\|S_h\|^m}_{\leq 1}$$

is small (approximation of the low frequency parts on grid size H), and

$$\underbrace{\|I - I_H^h \mathcal{K}_H^{-1} I_h^H\|}_{\leq C} \underbrace{\|P_h^{\text{high}} S_h^m\|}_{\longrightarrow 0,\ (m \to \infty)}$$

is arbitrary small for sufficiently many smoothing steps m (smoothing of the high frequency parts).

The second classical technique uses a product splitting of the iteration matrix and can be found in Hackbusch [59]. By inserting $I = \mathcal{K}_h^{-1} \mathcal{K}_h$, the iteration matrix M_H^h is splitted into the factors

$$M_H^h = \left(\mathcal{K}_h^{-1} - I_H^h \mathcal{K}_H^{-1} I_h^H\right)\left(\mathcal{K}_h S_h^m\right).$$

Again two properties have to be shown. The first one is called the approximation property:

$$\|\mathcal{K}_h^{-1} - I_H^h \mathcal{K}_H^{-1} I_h^H\| \leq c h^{\delta}. \tag{3.5}$$

The smaller the factor, the better the coarse-grid solutions approximate $\underline{u}_h$. The second estimate

$$\|\mathcal{K}_h S_h^m\| \leq \eta(m) h^{-\delta}, \tag{3.6}$$

called the smoothing property, describes the efficiency of the smoothing iteration. Here $\eta(m)$ denotes the so-called smoothing function. This should be a function, independent of the discretization parameter h and additionally

$$\eta(m) \longrightarrow 0 \quad \text{for} \quad m \longrightarrow \infty.$$

Both estimates together imply the two-grid convergence for a sufficiently large number of smoothing steps. Here, again, the involved norms have to be chosen in such a way, that we achieve the same exponent δ with different sign, such that the contraction number of the two-grid method is independent of the discretization parameter h.

Remark 5. *Instead of* (3.5) *and* (3.6) *the approximation and smoothing properties are often formulated using* $\|\mathcal{K}_h\|$ *instead of* $h^{-\delta}$. *Then the approximation property reads as*

$$\|\mathcal{K}_h^{-1} - I_H^h \mathcal{K}_H^{-1} I_h^H\| \leq c/\|\mathcal{K}_h\|, \tag{3.7}$$

and the smoothing property changes to

$$\|\mathcal{K}_h S_h^m\| \leq \eta(m)\|\mathcal{K}_h\|. \tag{3.8}$$

3.3.2 Multigrid Convergence

The technique for proving multigrid convergence is based on the two-grid convergence. Again, the goal is to estimate the iteration matrix in a proper norm. Since the multigrid iteration is a recursive procedure, the iteration matrix M_L on the finest level L is rescursively described:

$$\begin{aligned} M_0 &= 0 \\ M_k &= S_k^{m_2}(I_k - I_{k-1}^k(I_{k-1} - M_{k-1}^\gamma)\mathcal{K}_{k-1}^{-1} I_k^{k-1}\mathcal{K}_k)S_k^{m_1} \qquad (k = 1, \ldots, L). \end{aligned}$$

The multigrid iteration matrix can be regarded as the two-grid iteration matrix M_k^{k-1} plus a perturbation term:

$$\begin{aligned} M_k &= S_k^{m_2}(I_k - I_{k-1}^k \mathcal{K}_{k-1}^{-1} I_k^{k-1}\mathcal{K}_k)S_k^{m_1} + B_{k-1}^k M_{k-1}^\gamma B_k^{k-1} \\ &= M_k^{k-1} + B_{k-1}^k M_{k-1}^\gamma B_k^{k-1} \end{aligned}$$

with

$$B_{k-1}^k := S_k^{m_2} I_{k-1}^k$$

and

$$B_k^{k-1} := \mathcal{K}_{k-1}^{-1} I_k^{k-1} \mathcal{K}_k S_k^{m_1}.$$

For a given two-grid contraction, i.e., $\|M_k^{k-1}\| \leq \sigma^* < 1$, and the additional assumptions

$$\|B_{k-1}^k\| \cdot \|B_k^{k-1}\| \leq C^*,$$

with a constant C^* independent of h_k, the contraction number can be bounded, i.e., $\|M_L\| \leq \rho_L$, and ρ_L is given by the following recursion

$$\begin{aligned} \rho_1 &:= \sigma^*, \\ \rho_{k+1} &:= \sigma^* + C^* \rho_k^\gamma \qquad (k = 1, \ldots, L-1). \end{aligned}$$

This recursion can now be used to prove contraction rates for the multigrid method, if the two-grid method converges for all level with a sufficiently small contraction number σ^* and, additionally, the cycle-type on each level is properly chosen.

Usually this technique is used to prove the convergence of the W-cycle, i.e., $\gamma = 2$ on each level. If σ^* is small enough, such that $4C^*\sigma^* \leq 1$, the contraction number $\|M_L\|$ of the W-cycle is estimated by

$$\|M_L\| \leq \rho^* = \frac{1 - \sqrt{1 - 4C^*\sigma^*}}{2C^*} \leq 2\sigma^*.$$

In this respect, for the construction of multigrid methods, it is usually sufficient to analyze only the corresponding two-grid method. Furthermore, it is usually not necessary to work with $\gamma > 2$ if σ^* is sufficiently small.

The technique presented above is not restricted to the W-cycle. For instance, it also works for the so-called V-W-cycle, where we alternate the cycle type between V-cycle and W-cycle on the coarser level, and also for the variable V-cycle, where the number of smoothing steps is increased on the coarser levels. The corresponding analysis then yields conditions, how many smoothing steps have to be applied.

This convergence analysis fails for the standard V-cycle. The convergence for the V-cycle was first proved in Braess [21, 22]. For general symmetric problems with sufficient regularity, a V-cycle proof can be found in Braess and Hackbusch [25] and for more general smoothing techniques, see Bank and Douglas [11]. A convergence proof for the V-cycle without full elliptic regularity is given in Brenner [34].

A different qualitative multigrid theory is based on subspace decompositions due to Bramble, Pasciak, Oswald and Xu. This is the basis of the multigrid subspace correction theory, see e.g., Bramble, Pasciak and Xu [31], Xu [99] and for a wide overview Bramble [27], Oswald [71].

3.4 Smoothing Techniques for Saddle Point Problems

While the construction and convergence theory for scalar, symmetric and elliptic variational problems is well understood, the case of mixed variational problems is still challenging. The most crucial part is the proper choice of a smoothing technique. Usually, the well-known smoothing iterations for the scalar problems (damped Jacobi, or Gauss-Seidel relaxation) are not appropriate for saddle point problems or are even not defined, e.g. in saddle point systems like (2.26), where $C = 0$, the diagonal matrix is singular. There are natural ways

to generalize scalar smoothing schemes to systems of PDEs. An immediate generalization of standard scalar relaxation is smoothing by collective relaxation. That is, all unknowns at each single grid point are relaxed simultaneously. We will call this type of smoother in the following the point smoother. The collective relaxation is used in finite difference multigrid schemes in optimal control problems, see, e.g., Borzi, Kunisch and Kwak [20] and Borzi and Kunisch [19].

Sometimes, however, collective point relaxation is not sufficient and more complex relaxation schemes have to be employed. The next generalization step leads to patch smoother (or sometimes called box relaxation). The computational domain is divided into (overlapping or non-overlapping) small patches. One smoothing iteration step consists of solving local problems on each patch. This is done patch-by-patch either in a Jacobi-type or Gauss-Seidel-type manner. The technique was successfully used for the Navier-Stokes equations, see Vanka [92]. Of course, the point smoother is a special case of the patch smoother.

In this section we give a general introduction in these two smoothing schemes and discuss the known theoretical results, whereas in the next chapter we will apply these smoothing principles to a certain class of optimal control problems.

3.4.1 General Construction

First, we will construct a general iterative method by setting up local problems. This will be done in a rather general setting and includes both the point and the patch smoothing technique. Next, the choice of patches and local problems are presented in more details and we show some theoretical results, which are necessary to prove the smoothing property for this kind of methods.

The smoothing procedure involves only one level k of the hierarchy of spaces. Therefore, we will simplify the notation by dropping the subscript k and, additionally, omitting underlining the vectors. So, in this section, we discuss iterative methods (as smoothers) for linear systems of equations of the form:

$$\mathcal{K}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \quad \text{with} \quad \mathcal{K} = \begin{pmatrix} A & B^T \\ B & -C \end{pmatrix}, \tag{3.9}$$

where $x \in \mathbb{R}^n$, $p \in \mathbb{R}^m$, under the assumption that A is a symmetric positive semi-definite $n \times n$ matrix, B is an $m \times n$ matrix, and C is a symmetric positive semi-definite $m \times m$ matrix, and that $\mathcal{K}$ is nonsingular.

For setting up local sub-problems a set of linear operators is introduced:

$$P_i \colon \mathbb{R}^{n_i} \longrightarrow \mathbb{R}^n, \quad Q_i \colon \mathbb{R}^{m_i} \longrightarrow \mathbb{R}^m, \quad \text{for } i = 1, \ldots, N,$$

where the dimensions n_i and m_i are typically much smaller than the dimensions n and m of the original spaces, respectively. The operators P_i and Q_i are interpreted as prolongation

operators with associated restriction operators P_i^T and Q_i^T. We assume that

$$\sum_{i=1}^{N} Q_i Q_i^T \text{ is nonsingular} \quad \text{and} \quad \sum_{i=1}^{N} P_i P_i^T \text{ is nonsingular.} \tag{3.10}$$

These conditions guarantee that we have complete space decompositions

$$\mathbb{R}^n = \sum_{i=1}^{N} P_i(\mathbb{R}^{n_i}) \quad \text{and} \quad \mathbb{R}^m = \sum_{i=1}^{N} Q_i(\mathbb{R}^{m_i}).$$

For instance, in case of the collective smoothing iteration we have a non-overlapping decomposition in one-dimensional subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and the prolongations P_i and Q_i are the corresponding canonical embeddings into $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Starting from some approximations $x^{(j)}$ and $p^{(j)}$ of the exact solutions x and p of (3.9) we consider iterative methods of form:

$$x^{(j+1)} = x^{(j)} + \sum_{i=1}^{N} P_i\, s_i^{(j)}, \quad p^{(j+1)} = p^{(j)} + \sum_{i=1}^{N} Q_i\, r_i^{(j)}, \tag{3.11}$$

where $(s_i^{(j)}, r_i^{(j)}) \in \mathbb{R}^{n_i} \times \mathbb{R}^{m_i}$ solves a local saddle point problem of the form

$$\hat{\mathcal{K}}_i \begin{pmatrix} s_i^{(j)} \\ r_i^{(j)} \end{pmatrix} = \begin{pmatrix} P_i^T \left[f - Ax^{(j)} - B^T p^{(j)} \right] \\ Q_i^T \left[g - Bx^{(j)} + Cp^{(j)} \right] \end{pmatrix} \quad \text{for all } i = 1, \ldots, N. \tag{3.12}$$

That means, that the residuals of the approximations are first restricted to the smaller spaces, then a series of small saddle point problems must be solved, and, finally, the solutions are prolongated and determine the next iterate. This process can be viewed as an additive Schwarz method. Here, the local problems are solved independently of each other. There is a multiplicative version of this method at hand, where the overall iteration through all patches is done in a Gauss-Seidel-type manner, i.e., the residual and therefore the right hand side of the local problems is updated after each local patch solve. This can be viewed as a multiplicative Schwarz-type iteration.

3.4.2 Symmetric Inexact Uzawa Method

Under suitable conditions on the prolongations and the involved local matrices, the additive Schwarz iteration can be viewed as a symmetric inexact Uzawa method. In Schöberl and Zulehner [82], new convergence and smoothing properties were derived, which are necessary for the smoothing analysis. We give a short review of the general results on Schwarz-type smoothers from Schöberl and Zulehner [82].

In addition to (3.10) we assume that

$$\sum_{i=1}^{N} P_i P_i^T = I,$$

where I denotes the identity matrix. The prolongations P_i determine then a special partition of unity: For each $x \in \mathbb{R}^n$ we have

$$u = \sum_{i=1}^{N} P_i u_i \quad \text{with} \quad u_i = P_i^T u \quad \text{and} \quad \|u\|_{\ell^2}^2 = \sum_{i=1}^{N} \|P_i u_i\|_{\ell^2}^2.$$

Here $\|\cdot\|_{\ell^2}$ denotes the Euclidean norm.

For each index $i \in 1, \ldots, N$, local matrices $\hat{A}_i, B_i$ and $\hat{S}_i$ of size $n_i \times n_i$, $m_i \times n_i$ and $m_i \times m_i$, respectively, have to be chosen, which determine local matrices $\hat{\mathcal{K}}_i$ of the form

$$\hat{\mathcal{K}}_i = \begin{pmatrix} \hat{A}_i & B_i^T \\ B_i & B_i \hat{A}_i^{-1} B_i^T - \hat{S}_i \end{pmatrix}.$$

A natural candidate for $\hat{S}_i$ would be the exact Schur complement $\hat{C}_i + B_i \hat{A}_i^{-1} B_i^T$ with some properly chosen matrix $\hat{C}_i$, resulting in a lower right block $-\hat{C}_i$ in $\hat{\mathcal{K}}_i$. The general form of the lower right block of $\hat{\mathcal{K}}_i$ is more flexible and also allows to choose an approximate Schur complement for $\hat{S}_i$ instead.

We assume that the local matrices B_i are related to the (global) matrix B by the following commutativity condition

$$Q_i^T B = B_i P_i^T \quad \text{for all } i = 1, 2, \ldots, N. \tag{3.13}$$

A similar condition is needed for the local matrices $\hat{A}_i$: We assume that there exists a (global) matrix $\hat{A}$ (typically not equal to A) such that:

$$P_i^T \hat{A} = \hat{A}_i P_i^T \quad \text{for all } i = 1, 2, \ldots, N. \tag{3.14}$$

Both commutativity conditions are important for the analysis. The second condition is actually very restrictive. It essentially implies that $\hat{A}$ can only be a (block) diagonal matrix.

From the local matrices $\hat{S}_i$ the following (global) matrix $\hat{S}$ is constructed:

$$\hat{S} = \left(\sum_{i=1}^{N} Q_i \hat{S}_i^{-1} Q_i^T \right)^{-1}. \tag{3.15}$$

The conditions (3.13) and (3.14) can also be written as commutative diagrams:

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\hat{A}} & \mathbb{R}^n \\ {\scriptstyle P_i^T}\downarrow & & \downarrow{\scriptstyle P_i^T} \\ \mathbb{R}^{n_i} & \xrightarrow{\hat{A}_i} & \mathbb{R}^{n_i} \end{array} \qquad \begin{array}{ccc} \mathbb{R}^n & \xrightarrow{B} & \mathbb{R}^m \\ {\scriptstyle P_i^T}\downarrow & & \downarrow{\scriptstyle Q_i^T} \\ \mathbb{R}^{n_i} & \xrightarrow{B_i} & \mathbb{R}^{m_i} \end{array}$$

Theorem 9. *Under the assumptions (3.10), (3.13) and (3.14) and with the construction (3.15), the iterative method* (3.11) *can be written equivalently as the following preconditioned Richardson method:*

$$x^{(j+1)} = x^{(j)} + s^{(j)}, \quad p^{(j+1)} = p^{(j)} + r^{(j)}, \tag{3.16}$$

where $(s^{(j)}, r^{(j)})$ *solve the equation*

$$\hat{\mathcal{K}} \begin{pmatrix} s^{(j)} \\ r^{(j)} \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} - \mathcal{K} \begin{pmatrix} x^{(j)} \\ p^{(j)} \end{pmatrix} \quad \textit{with} \quad \hat{\mathcal{K}} = \begin{pmatrix} \hat{A} & B^T \\ B & B\hat{A}^{-1}B^T - \hat{S} \end{pmatrix}. \tag{3.17}$$

Proof. The proof is given in Schöberl and Zulehner [82]. □

So the additive Schwarz-type iterative method can be represented as a symmetric inexact Uzawa method, see Section 2.5.3. Let $\mathcal{M}$ denote the associated iteration matrix, given by

$$\mathcal{M} = I - \hat{\mathcal{K}}^{-1}\mathcal{K},$$

which controls the error propagation for the iterative method.

In the next theorem an important estimate is formulated which is needed in the forthcoming smoothing property analysis. Here and in the sequel the following notations are used: $M < N$ ($N > M$) iff $N - M$ is positive definite, and $M \leq N$ ($N \geq M$) iff $N - M$ is positive semi-definite, for symmetric matrices M and N. Furthermore, for a symmetric and positive definite matrix S, the norms $\|v\|_S$ and $\|M\|_S$ of a vector v and a matrix M (as a representation of a bilinear form) are given by

$$\|v\|_S = \sqrt{(Sv, v)_{\ell^2}} \quad \text{and} \quad \|M\|_S = \sup_{v,w \neq 0} \frac{|(Mv, w)_{\ell^2}|}{\|v\|_S \, \|w\|_S}.$$

Theorem 10. *Let* A *be a symmetric and positive semi-definite* $n \times n$ *matrix,* B *an* $m \times n$ *matrix, and* C *be a symmetric and positive semi-definite* $m \times m$ *matrix. Let* $\hat{A}$ *be a symmetric and positive definite* $n \times n$ *matrix, and* $\hat{S}$ *a symmetric positive definite* $m \times m$ *matrix, satisfying*

$$\hat{A} \geq A \quad \textit{and} \quad \hat{S} \geq C + B\hat{A}^{-1}B^T. \tag{3.18}$$

Then

$$\|\mathcal{K}\mathcal{M}^m\|_{\mathcal{L}} \leq \eta_0(m) \, \|\mathcal{Q}\|_{\mathcal{L}},$$

where $\mathcal{K}$ *is given by (3.9),* $\hat{\mathcal{K}}$ *is given by (3.17),* $\mathcal{Q}$ *is given by*

$$\mathcal{Q} = \begin{pmatrix} \hat{A} - A & 0 \\ 0 & \hat{S} - C - B\hat{A}^{-1}B^T \end{pmatrix},$$

$\mathcal{L}$ *is an arbitrary symmetric and positive definite matrix, and*

$$\eta_0(m) = \frac{1}{2^{m-1}} \binom{m-1}{[m/2]} \leq \begin{cases} \sqrt{\dfrac{2}{\pi(m-1)}} & \textit{for even } m, \\ \sqrt{\dfrac{2}{\pi m}} & \textit{for odd } m. \end{cases}$$

Here $\binom{n}{k}$ *denotes the binomial coefficient and* $[x]$ *denotes the largest integer smaller than or equal to* $x \in \mathbb{R}$.

Proof. For the special case of the Euclidean norm ($\mathcal{L} = I$) see SCHÖBERL AND ZULEHNER [82], the proof for more general norms is completely analogous. □

Comparing the smoothing property (3.8) with Theorem 10 it is immediately clear that the smoothing property is satisfied for the additive Schwarz-type method (3.16), if the local problems are constructed in such a way that the associated global matrices $\hat{A}$, B and $\hat{S}$ satisfy the conditions (3.18) and if, additionally, the following scaling condition holds:

$$\|\mathcal{Q}\|_{\mathcal{L}} \leq c_R \|\mathcal{K}\|_{\mathcal{L}} \quad \text{with} \quad \mathcal{Q} = \begin{pmatrix} \hat{A} - A & 0 \\ 0 & \hat{S} - C - B\hat{A}^{-1}B^T \end{pmatrix} \tag{3.19}$$

for some constant c_R independent of the discretization parameter. The smoothing rate is then given by

$$\eta(m) = c_R \eta_0(m) = \mathcal{O}\left(\frac{1}{\sqrt{m}}\right).$$

3.4.3 Other Smoothing Techniques

For the sake of completeness we will give a short overview of some other possibilities to construct a smoothing iteration for mixed problems. Saddle point problems are indefinite, so we can not apply the standard smoothing techniques like Jacobi, Richardson or Gauss-Seidel. In order to keep the possibility of applying standard smoothers, the squared system

$$\mathcal{K}^T\mathcal{K}\begin{pmatrix} x \\ p \end{pmatrix} = \mathcal{K}^T\begin{pmatrix} f \\ g \end{pmatrix}$$

can be considered, which is symmetric and positive definite. In VERFÜRTH [94] and BRENNER [36] Richardson iterations on the squared system are used, which lead to a smoothing rate of order $1/\sqrt{m}$.

Another class of smoothers for mixed problems are Braess-Sarazin smoothers, which are connected to inexact symmetric Uzawa methods. The smoothing rate of order $1/m$ was shown in BRAESS AND SARAZIN [26] with $\hat{S} = C + B\hat{A}^{-1}B^T$ in (2.36). In ZULEHNER [101] the same smoothing behavior is shown under some weaker conditions.

The last class of smoothers we want to mention here are the so-called transforming smoothers, introduced in WITTUM [97, 98]. The idea is to transform the matrix $\mathcal{K}$ by multiplication from the left and (or) the right into a block-triangular matrix and then find appropriate smoothers. In WITTUM [98] it is shown, that they only have to fulfill the smoothing properties for the diagonal block matrices. The concept of transforming smoothers was recently applied to the more general class of PDE-constrained optimization problems, see SCHULZ AND WITTUM [84].

3.5 Local Fourier Analysis

We present here the local Fourier analysis (LFA) as a tool for the quantitative analysis of the two-grid method and the design of smoothing methods. The local Fourier analysis, also called the local mode analysis, was introduced in BRANDT [32] and later extended and refined in BRANDT [33]. Details can also be found in STÜBEN AND TROTTENBERG [87] or WESSELING [95].

In describing the fundamental ideas of LFA, we confine ourselves to the 2D case and to standard coarsening ($H = 2h$). But it can be immediately carried over to d dimensions. The local Fourier analysis is based on certain idealized assumptions and simplifications: the boundary conditions are neglected and the problem is considered on regular infinite grids $G_h = \{(x_1, x_2) : x_1 = ih_1,\ x_2 = jh_2 \colon i, j \in \mathbb{Z}\}$. The fundamental quantities are the grid functions

$$\phi(\theta, x) = e^{i\theta \cdot x/h} := e^{i\theta_1 x_1/h_1} e^{i\theta_2 x_2/h_2} \quad \text{for } x \in G_h.$$

We assume that θ varies continuously in $\mathbb{R}^2$. It is sufficient to consider

$$\phi(\theta, x) \quad \text{with } \theta \in [-\pi, \pi)^2.$$

Since the smoother should reduce the high-oscillating parts of the error, for the analysis of a smoothing iteration, we have to distinguish high and low frequency components on G_h with respect to an (infinite) coarse grid G_H, in our case $H = 2h$. The definition is based on the fact that only those frequency components

$$\phi(\theta, x) \quad \text{with } -\frac{\pi}{2} \le \theta \le \frac{\pi}{2}$$

are distinguishable on G_H.

One defines

$$\begin{aligned} \phi \text{ low frequency component} &\iff \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right)^2, \\ \phi \text{ high frequency component} &\iff \theta \in [-\pi, \pi)^2 \setminus \left[-\frac{\pi}{2}, \frac{\pi}{2}\right)^2. \end{aligned}$$

The relation between the error before and after one smoothing iteration can be described by the corresponding iteration matrix M:

$$e^{(j+1)} = M e^{(j)}.$$

To analyze this iteration we define the errors in terms of their frequency components:

$$e^{(j)} = \sum_{\theta} \mathcal{E}_\theta^{(j)} e^{i\theta x/h} \quad \text{and} \quad e^{(j+1)} = \sum_{\theta} \mathcal{E}_\theta^{(j+1)} e^{i\theta x/h},$$

where $\mathcal{E}_\theta^{(j)}$ and $\mathcal{E}_\theta^{(j+1)}$ denote the error amplitudes of the θ-frequency component before and after the smoothing sweep, respectively.

The smoothing factor (in the framework of LFA) is then defined by

$$\mu := \sup \left\{ \left| \frac{\mathcal{E}_\theta^{(j+1)}}{\mathcal{E}_\theta^{(j)}} \right| \ , \ \theta \in [-\pi, \pi)^2 \setminus \left[-\frac{\pi}{2}, \frac{\pi}{2} \right)^2 \right\}.$$

Whenever the smoothing iteration includes some (relaxation) parameters, the LFA can be used to optimize the choice of these parameters, such that the smoothing factor is minimized. In Section 5.5 LFA will be applied to motivate the numerical results for different choices of a relaxation parameter.

A similar technique can be used to investigate the convergence properties of the whole two-grid method. For any low frequency $\theta = (\theta_1, \theta_2) \in [-\pi/2, \pi/2)^2$ consider the frequencies

$$\begin{aligned} \theta^{(0,0)} &:= (\theta_1, \theta_2), & \theta^{(1,1)} &:= (\bar{\theta}_1, \bar{\theta}_2), \\ \theta^{(1,0)} &:= (\bar{\theta}_1, \theta_2), & \theta^{(0,1)} &:= (\theta_1, \bar{\theta}_2), \end{aligned}$$

where

$$\bar{\theta}_i := \begin{cases} \theta_i + \pi & \text{if } \theta_i < 0 \\ \theta_i - \pi & \text{if } \theta_i \geq 0 \end{cases}.$$

For any low frequency the quadruple of Fourier components with the above defined frequencies coincide on the coarse grid. These Fourier components are called harmonics of each other and define the four-dimensional space of harmonics

$$E_h^\theta := \text{span } [\phi(\theta^\alpha, \cdot) : \alpha = (\alpha_1, \alpha_2) \in \{(0,0), (0,1), (1,0), (1,1)\}].$$

These spaces turned out to be invariant under the two-grid operator M_h^H under general assumptions.

Let $\bar{M}_h^H(\theta)$ denotes the representation (also called the symbol) of the two grid operator on E_h^θ, which is a (4×4)-matrix. The asymptotic convergence factor ρ_{loc} is then defined as follows:

$$\rho_{\text{loc}}(M_h^H) = \sup\{\rho(\bar{M}_h^H(\theta)) : \theta \in [-\pi/2, \pi/2)^2\},$$

where $\rho(\bar{M}_h^H(\theta))$ is the spectral radius of the (4×4)-matrix $\bar{M}_h^H(\theta)$. Similar, one can introduce an error reduction factor σ_{loc} with respect to an appropriate norm:

$$\sigma_{\text{loc}}(M_h^H) = \sup\{\|\bar{M}_h^H(\theta)\| : \theta \in [-\pi/2, \pi/2)^2\}.$$

Chapter 4

Application to an Optimal Control Problem

In this chapter we want to apply the multigrid method to a certain class of optimal control problems. We present the construction of a patch smoother, related to the one presented in Section 3.4.2. In Schöberl and Zulehner [82] this patch smoother was applied to the Stokes problem, discretized by the Crouzeix-Raviart mixed finite element method. A straight forward application leads to a violation of the important scaling condition (3.19). It will be shown how the construction must be modified to keep the right scaling without losing any of the other requirements, such that we can develop a rigorous convergence analysis of the corresponding multigrid method. The content of this chapter is summarized in Simon and Zulehner [85].

Let Ω be a bounded convex polygonal domain in $\mathbb{R}^2$. Let $L^2(\Omega)$ and $H^1(\Omega)$ denote the usual Lebesgue space and Sobolev space, respectively. We consider the following elliptic optimal control problem: Find the state $y \in H^1(\Omega)$ and the control $u \in L^2(\Omega)$ such that

$$J(y,u) = \min_{(z,v)\in H^1(\Omega)\times L^2(\Omega)} J(z,v)$$

with cost functional

$$J(z,v) = \frac{1}{2}\|z - y_d\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\|v\|^2_{L^2(\Omega)}$$

subject to the state equations

$$\begin{aligned} -\Delta y + y &= u \qquad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= 0 \qquad \text{on } \Gamma, \end{aligned}$$

where Γ denotes the boundary of Ω, $y_d \in L^2(\Omega)$ is the desired state and $\gamma > 0$ is the weight of the cost of the control (or simply a regularization parameter).

By introducing the adjoint state $p \in H^1(\Omega)$ we get the following optimality system, see e.g., Tröltzsch [89]:

1. The adjoint state equation:

$$\begin{aligned} -\Delta p + p &= -(y - y_d) \quad \text{in } \Omega, \\ \frac{\partial p}{\partial n} &= 0 \qquad\qquad\quad \text{on } \Gamma. \end{aligned} \tag{4.1}$$

2. The control equation:

$$\gamma u - p = 0 \quad \text{in } \Omega. \tag{4.2}$$

3. The state equation:

$$\begin{aligned} -\Delta y + y &= u \quad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= 0 \quad \text{on } \Gamma. \end{aligned} \tag{4.3}$$

The weak formulation of this problem leads to a mixed variational problem: Find $x = (y, u) \in X = Y \times U$ with $Y = H^1(\Omega)$, $U = L^2(\Omega)$ and $p \in Q = H^1(\Omega)$ such that

$$\begin{aligned} a(x, w) + b(w, p) &= \langle F, w \rangle \qquad \text{for all } w \in X, \\ b(x, q) &= 0 \qquad\qquad\quad \text{for all } q \in Q \end{aligned} \tag{4.4}$$

with

$$\begin{aligned} a(x, w) &= (y, z)_{L^2(\Omega)} + \gamma (u, v)_{L^2(\Omega)}, \\ b(w, q) &= (z, q)_{H^1(\Omega)} - (v, q)_{L^2(\Omega)}, \\ \langle F, w \rangle &= (y_d, z)_{L^2(\Omega),} \end{aligned}$$

where $w = (z, v)$ with $z \in Y$, $v \in U$, and $(.,.)_H$ is the standard scalar product in a Hilbert space H, whose norm is denoted by $\|.\|_H$.

The mixed variational problem can also be written as a variational problem on $X \times Q$: Find $(x, p) \in X \times Q$ such that

$$\mathcal{B}((x, p), (w, q)) = \langle \mathcal{F}, (w, q) \rangle \quad \text{for all } (w, q) \in X \times Q$$

with the bilinear form

$$\mathcal{B}((x, p), (w, q)) = a(x, w) + b(w, p) + b(x, q) \tag{4.5}$$

and the linear functional

$$\langle \mathcal{F}(w, q) \rangle = \langle F, w \rangle.$$

It is trivial that $\mathcal{F}$ is a bounded and linear functional on $X \times Q$. The next lemma guarantees that the problem is well-posed:

Lemma 2. *The bilinear form $\mathcal{B}$ is stable and bounded on $X \times Q$, i.e., there are positive constants c and C such that*

$$c \, \|(s,r)\|_{X\times Q} \leq \sup_{0\neq(w,q)\in X\times Q} \frac{\mathcal{B}((s,r),(w,q))}{\|(w,q)\|_{X\times Q}} \leq C \, \|(s,r)\|_{X\times Q}$$

for all $(s,r) \in X \times Q$, where the norm on $X \times Q$ is given by

$$\|(w,q)\|^2_{X\times Q} = \|w\|^2_X + \|q\|^2_Q$$

with

$$\|w\|^2_X = \|z\|^2_{H^1(\Omega)} + \|v\|^2_{L^2(\Omega)} \quad \textit{for } w = (z,v) \quad \textit{and} \quad \|q\|_Q = \|q\|_{H^1(\Omega)}.$$

Proof. Boundedness follows easily from Cauchy's inequality, the stability from Brezzi's Theorem, see Brezzi and Fortin [39] and Section 2.3.2, if the bilinear form a is coercive on the set

$$\ker B = \{w \in X : b(w,q) = 0 \text{ for all } q \in Q\},$$

i.e.: there exists a constant $\alpha > 0$ such that

$$a(w,w) \geq \alpha \, \|w\|^2_X \quad \text{for all } w \in \ker B, \tag{4.6}$$

and the inf-sup condition for b is satisfied, i.e.: there exists a constant $\beta > 0$ such that

$$\sup_{0\neq w\in X} \frac{b(w,q)}{\|w\|_X} \geq \beta \, \|q\|_Q \quad \text{for all } q \in Q. \tag{4.7}$$

So we have to verify these two conditions (4.6) and (4.7).

To show (4.6) let $w = (z,v) \in \ker B$. Remember that $Y = Q$, so we have in particular $b(w,z) = 0$, i.e.

$$(z,z)_{H^1(\Omega)} = (v,z)_{L^2(\Omega)},$$

which implies

$$\|z\|^2_{H^1(\Omega)} \leq \|v\|_{L^2(\Omega)} \, \|z\|_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)} \, \|z\|_{H^1(\Omega)}.$$

Hence: $\|z\|_{H^1(\Omega)} \leq \|v\|_{L^2(\Omega)}$. Then

$$\begin{aligned} a(w,w) &= \|z\|^2_{L^2(\Omega)} + \gamma \, \|v\|^2_{L^2(\Omega)} \geq \|z\|^2_{L^2(\Omega)} + \frac{\gamma}{2} \, \|v\|_{L^2(\Omega)} + \frac{\gamma}{2} \, \|z\|^2_{H^1(\Omega)} \\ &\geq \frac{\gamma}{2} \left(\|v\|^2_{L^2(\Omega)} + \|z\|^2_{H^1(\Omega)} \right), \end{aligned}$$

which proves (4.6) with $\alpha = \gamma/2$.

Next we have

$$\sup_{0\neq w\in X} \frac{b(w,q)}{\|w\|_X} \geq \frac{b((q,0),q)}{\|(q,0)\|_X} = \frac{(q,q)_Q}{\|q\|_Q} = \|q\|_Q,$$

which shows (4.7) with $\beta = 1$, where we again used the fact that $Y = Q$. □

Remark 6. *The coercivity constant α deteriorates for small values γ. So the analysis presented here is certainly not designed for small values of γ. A different numerical approach addressing the question of robustness with respect to γ more rigorously is given in* Schöberl and Zulehner *[83]. However, the approach taken there is limited to a particular class of optimal control problems with distributed control.*

Let $(\mathcal{T}_k)$ be a sequence of triangulations of Ω, where $\mathcal{T}_{k+1}$ is obtained by dividing each triangle into four smaller triangles by connecting the midpoints of the edges of the triangles in $\mathcal{T}_k$. The quantity $\max\{\operatorname{diam} T : T \in \mathcal{T}_k\}$ is denoted by h_k.

We consider the following discretization by continuous and piecewise linear finite elements:

$$X_k = Y_k \times U_k = \{(z,v) \in C(\bar\Omega) \times C(\bar\Omega) : z|_T, v|_T \in P_1 \text{ for all } T \in \mathcal{T}_k\},$$
$$Q_k = \{q \in C(\bar\Omega) : q|_T \in P_1 \text{ for all } T \in \mathcal{T}_k\},$$

where P_1 denotes the set of polynomials of total degree less or equal to 1. Then we obtain the following discrete variational problem: Find $x_k \in X_k$ and $p_k \in Q_k$ such that

$$\begin{aligned} a(x_k, w_k) + b(w_k, p_k) &= \langle F, w_k\rangle && \text{for all } w_k \in X_k, \\ b(x_k, q_k) &= 0 && \text{for all } q_k \in Q_k. \end{aligned} \tag{4.8}$$

The discrete mixed variational problem can also be written as a discrete variational problem on $X_k \times Q_k$: Find $(x_k, p_k) \in X_k \times Q_k$ such that

$$\mathcal{B}((x_k,p_k),(w_k,q_k)) = \langle \mathcal{F}, (w_k,q_k)\rangle \quad \text{for all } (w_k,q_k) \in X_k \times Q_k. \tag{4.9}$$

The boundedness of the discrete mixed problem follows directly from the boundedness of the continuous problem, since the discrete spaces are nested. This does not carry over to the stability, but with exactly the same arguments as above the stability of the discrete mixed problem can be shown:

Lemma 3. *The bilinear form $\mathcal{B}$ is uniformly stable and bounded on $X_k \times Q_k$, i.e., there are positive constants c and C such that*

$$c\,\|(s,r)\|_{X\times Q} \le \sup_{0\neq(w,q)\in X_k\times Q_k} \frac{\mathcal{B}((s,r),(w,q))}{\|(w,q)\|_{X\times Q}} \le C\,\|(s,r)\|_{X\times Q}$$

for all $(s,r) \in X_k \times Q_k$ and all k, where the constants c and C are independent of the grid level k.

Proof. The proof is completely analog to the proof of Lemma 2:

Boundedness follows from Cauchy's inequality and for the stability we apply Brezzi's Theorem. We have to check, if the bilinear form a is coercive, now on the set

$$\ker B_k = \{w \in X_k : b(w,q) = 0 \text{ for all } q \in Q_k\},$$

i.e.: there exists a constant $\tilde{\alpha} > 0$ such that

$$a(w,w) \geq \tilde{\alpha}\|w\|_X \quad \text{for all } w \in \ker B_k, \tag{4.10}$$

and the discrete inf-sup condition for b is satisfied, i.e.: there exists a constant $\tilde{\beta} > 0$ such that

$$\sup_{0\neq w\in X_k} \frac{b(w,q)}{\|w\|_X} \geq \tilde{\beta}\,\|q\|_Q \quad \text{for all } q \in Q_k. \tag{4.11}$$

Let $w = (z,v) \in \ker B_k$. Then, in particular (by using $Y_k = Q_k$), we have $b(w,z) = 0$, i.e.

$$(z,z)_{H^1(\Omega)} = (v,z)_{L^2(\Omega)},$$

which implies

$$\|z\|^2_{H^1(\Omega)} \leq \|v\|_{L^2(\Omega)}\,\|z\|_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)}\,\|z\|_{H^1(\Omega)}.$$

Hence: $\|z\|_{H^1(\Omega)} \leq \|v\|_{L^2(\Omega)}$. Then

$$\begin{aligned} a(w,w) &= \|z\|^2_{L^2(\Omega)} + \gamma\,\|v\|^2_{L^2(\Omega)} \geq \|z\|^2_{L^2(\Omega)} + \frac{\gamma}{2}\,\|v\|_{L^2(\Omega)} + \frac{\gamma}{2}\,\|z\|^2_{H^1(\Omega)} \\ &\geq \frac{\gamma}{2}\left(\|v\|^2_{L^2(\Omega)} + \|z\|^2_{H^1(\Omega)}\right), \end{aligned}$$

which proves (4.10) with $\tilde{\alpha} = \gamma/2$.

Next we have (by using $Y_k = Q_k$)

$$\sup_{0\neq w\in X_k} \frac{b(w,q)}{\|w\|_X} \geq \frac{b((q,0),q)}{\|(q,0)\|_X} = \frac{(q,q)_Q}{\|q\|_Q} = \|q\|_Q,$$

which shows (4.11) with $\tilde{\beta} = 1$. □

By introducing the standard nodal basis, we finally obtain the following saddle point problem in matrix-vector notation:

$$\mathcal{K}_k \begin{pmatrix} \underline{x}_k \\ \underline{p}_k \end{pmatrix} = \begin{pmatrix} \underline{f}_k \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{K}_k = \begin{pmatrix} A_k & B_k^T \\ B_k & 0 \end{pmatrix},$$

where

$$A_k = \begin{pmatrix} M_k & 0 \\ 0 & \gamma\,M_k \end{pmatrix} \quad \text{and} \quad B_k = \begin{pmatrix} K_k & -M_k \end{pmatrix}.$$

Here M_k denotes the mass matrix representing the $L^2(\Omega)$ scalar product on Y_k, and K_k denotes the stiffness matrix representing the $H^1(\Omega)$ scalar product on Y_k.

4.1 Multigrid Convergence Analysis

We consider the following multigrid algorithm as an efficient solver of the discrete problems (4.9). One iteration step at level k is given in the following form:

Let $(x_k^{(0)}, p_k^{(0)}) \in X_k \times Q_k$ be a given approximation of the exact solution $(x_k, p_k) \in X_k \times Q_k$ to (4.9). Then the iteration proceeds in two stages:

1. Smoothing: For $j = 0, 1, \ldots, m-1$ compute $(x_k^{(j+1)}, p_k^{(j+1)}) \in X_k \times Q_k$ by an iterative procedure of the form
$$(x_k^{(j+1)}, p_k^{(j+1)}) = \mathcal{S}_k\,(x_k^{(j)}, p_k^{(j)}).$$

2. Coarse grid correction: Set
$$\langle \tilde{\mathcal{F}}, (w,q)\rangle = \langle \mathcal{F}, (w,q)\rangle - \mathcal{B}\left((x_k^{(m)}, p_k^{(m)}), (w,q)\right)$$
for $(w,q) \in X_{k-1} \times Q_{k-1}$ and let $(\tilde{s}_{k-1}, \tilde{r}_{k-1}) \in X_{k-1} \times Q_{k-1}$ satisfy
$$\mathcal{B}((\tilde{s}_{k-1}, \tilde{r}_{k-1}), (v,q)) = \langle \tilde{\mathcal{F}}, (w,q)\rangle \quad \text{for all } (w,q) \in X_{k-1} \times Q_{k-1}. \tag{4.12}$$
If $k = 1$, compute the exact solution of (4.12) and set $(s_{k-1}, r_{k-1}) = (\tilde{s}_{k-1}, \tilde{r}_{k-1})$.
If $k > 1$, compute approximations (s_{k-1}, r_{k-1}) by applying $\mu \geq 2$ iteration steps of the multigrid algorithm applied to (4.12) on level $k-1$ with zero starting values.
Set
$$(x_k^{(m+1)}, p_k^{(m+1)}) = (x_k^{(m)}, p_k^{(m)}) + (s_{k-1}, r_{k-1}).$$

For analyzing the convergence of the multigrid method, we have to introduce a pair of norms in which we can show the approximation property and the smoothing property. First we need an appropriate (L^2-like) mesh-dependent norm $|||(w,q)|||_{0,k}$ on $X_k \times Q_k$. We introduce a second discrete norm on $X_k \times Q_k$ by
$$|||(s,r)|||_{2,k} = \sup_{0 \neq (w,q) \in X_k \times Q_k} \frac{|\mathcal{B}((s,r),(w,q))|}{|||(w,q)|||_{0,k}}.$$

Now, we can formulate the approximation property and the smoothing property: Consider the two-grid algorithm (i.e. exact solution of the coarse grid correction equation (4.12) at level $k-1$). The approximation property measures the effect of the coarse grid correction: It is assumed that there is a constant c_A which is independent of k such that
$$|||\mathcal{B}|||_{0,k}\, |||(x_k^{(m+1)} - x_k, p_k^{(m+1)} - p_k)|||_{0,k} \leq c_A\, |||(x_k^{(m)} - x_k, p_k^{(m)} - p_k)|||_{2,k}, \tag{4.13}$$
where $(x_k, p_k) \in X_k \times Q_k$ solves (4.9) and the norm $|||\mathcal{B}|||_{0,k}$ of the bilinear form $\mathcal{B}$ is given by
$$|||\mathcal{B}|||_{0,k} = \sup_{0 \neq (s,r),(w,q) \in X_k \times Q_k} \frac{|\mathcal{B}((s,r),(w,q))|}{|||(s,r)|||_{0,k}\, |||(w,q)|||_{0,k}}.$$

The remaining part to complete the proof of the two-grid convergence is the smoothing property, which measures the effect of the smoothing procedure: It is assumed that

$$|||(x_k^{(m)} - x_k, p_k^{(m)} - p_k)|||_{2,k} \leq \eta(m)\, |||\mathcal{B}|||_{0,k}\, |||(x_k^{(0)} - x_k, p_k^{(0)} - p_k)|||_{0,k}$$

for some function $\eta(m)$ which is independent of k, and

$$\eta(m) \to 0 \quad \text{for } m \to \infty.$$

The convergence of the two-grid method for a sufficiently large number m of smoothing steps easily follows by combining the approximation property and the smoothing property. From this the convergence of the multigrid method can be derived by standard arguments, see, e.g., Hackbusch [59] and Section 3.3.

Let $\mathcal{L}_k$ be the symmetric and positive definite matrix on $\mathbb{R}^{n_k} \times \mathbb{R}^{m_k}$ which represents the mesh dependent norm $|||(w,q)|||_{0,k}$:

$$|||(w,q)|||_{0,k} = \left(\left(\mathcal{L}_k \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix}, \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix} \right)_{\ell^2} \right)^{1/2} = \left\| \begin{pmatrix} \underline{w} \\ \underline{q} \end{pmatrix} \right\|_{\mathcal{L}_k} \tag{4.14}$$

for $w \in X_k$, $q \in Q_k$ with vector representations $\underline{w} \in \mathbb{R}^{n_k}$, $\underline{q} \in \mathbb{R}^{m_k}$.

It easily follows that the smoothing property translates to the following condition in matrix-notation:

$$\|\mathcal{K}_k \mathcal{M}_k^m\|_{\mathcal{L}_k} \leq \eta(m)\, \|\mathcal{K}_k\|_{\mathcal{L}_k}. \tag{4.15}$$

4.2 Approximation Property

Here we follow the general technique presented in Brenner [35] for proving the approximation property (4.13) for the special mesh dependent norm on $X_k \times Q_k$, given by

$$|||(w,q)|||_{0,k} = h_k \left(\|\underline{z}\|_{\ell^2}^2 + h_k^2 \, \|\underline{v}\|_{\ell^2}^2 + \|\underline{q}\|_{\ell^2}^2 \right)^{1/2} \tag{4.16}$$

with $w = (z,v) \in X_k$ and $q \in Q_k$ and their vector representations $\underline{z}$, $\underline{v}$ and $\underline{q}$.

Nine assumptions (A.1) - (A.9) are formulated in Brenner [35] which imply the approximation property. We cite them here, grouped in a few lemmata and translated into our notation. The constants in the nine assumptions have to be independent of the mesh level k. In our case the constants may be dependent of the regularization parameter γ, which means, that we are not able to prove robustness of the method with respect to this parameter.

Assumptions (A.1) and (A.2) are the (uniform) boundedness and stability of the continuous variational problem which were already shown in Lemma 2.

The verification of Assumption (A.3), which is a elliptic regularity result, is the content of the next lemma:

Lemma 4. *For all $f, g \in L^2(\Omega)$ the variational problem*

$$\mathcal{B}((x,p),(w,q)) = (f,z)_{L^2(\Omega)} + (g,q)_{L^2(\Omega)} \tag{4.17}$$

for all $w = (z,v) \in X = Y \times U$, $q \in Q$, has a solution $x = (y,u) \in H^2(\Omega) \times H^1(\Omega)$ and $q \in H^2(\Omega)$. There exists a constant C such that

$$\|y\|_{H^2(\Omega)} + \|u\|_{H^1(\Omega)} + \|p\|_{H^2(\Omega)} \leq C\,(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

for all $f, g \in L^2(\Omega)$.

Proof. From (4.17) with $v = 0$ and $q = 0$ we obtain the elliptic variational problem for p:

$$(p,z)_{H^1(\Omega)} = (f,z)_{L^2(\Omega)} - (y,z)_{L^2(\Omega)} \quad \text{for all } z \in H^1(\Omega).$$

Since we have assumed that $\Omega \subset \mathbb{R}^2$ is a convex polygonal domain, elliptic regularity follows:

$$\|p\|_{H^2(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)} + \|y\|_{L^2(\Omega)}\right).$$

Then Lemma 2 implies

$$\|p\|_{H^2(\Omega)} \leq C\,(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

(Throughout this proof, C denotes a generic constant.) From (4.17) with $z = 0$ and $v = 0$ we obtain the elliptic variational problem for y:

$$(y,q)_{H^1(\Omega)} = (g,q)_{L^2(\Omega)} + (u,q)_{L^2(\Omega)} \quad \text{for all } q \in H^1(\Omega).$$

Then elliptic regularity and Lemma 2 imply

$$\|y\|_{H^2(\Omega)} \leq C\,(\|g\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}) \leq C\,(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

From (4.17) with $z = 0$ and $q = 0$ we obtain

$$\gamma\, u = p.$$

Then Lemma 2 implies

$$\|u\|_{H^1(\Omega)} = \frac{1}{\gamma}\,\|p\|_{H^1(\Omega)} \leq C\,(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

□

Assumptions (A.4) and (A.5) are the (uniform) boundedness and stability of the discrete variational problem and were already shown in Lemma 3.

Assumption (A.6) requires standard L^2 estimates of the approximation error of the finite element space, which is guaranteed by the following lemma.

Lemma 5. *The interpolation operator*

$$\Pi_k \colon (X \times Q) \cap (H^2(\Omega) \times H^1(\Omega) \times H^2(\Omega)) \longrightarrow X_k \times Q_k$$

has the property that

$$\|((z,v),q) - \Pi_k((z,v),q)\|_{L^2(\Omega)} \leq C(h_k^2|z|_{H^2(\Omega)} + h_k|v|_{H^1(\Omega)} + h_k^2|q|_{H^2(\Omega)}), \tag{4.18}$$

and

$$\|((z,v),q) - \Pi_k((z,v),q)\|_{X\times Q} \leq Ch_k(|z|_{H^2(\Omega)} + |v|_{H^1(\Omega)} + |q|_{H^2(\Omega)}) \tag{4.19}$$

for all $(w,q) \in X \times Q$.

Proof. For $z \in H^2(\Omega)$ and $q \in H^2(\Omega)$ we can use the standard Lagrange interpolation results. Since v lies only in $H^1(\Omega) \not\subset C(\Omega)$, we have to use a different interpolation operator, e.g., the Clément-interpolation, see Clément [46], to get the desired approximation result. □

Assumption (A.7) is a discretization error estimate:

Lemma 6. *Let* $(x,p) \in X \times Q$ *satisfies*

$$\mathcal{B}((x,p),(w,q)) = (f,z)_{L^2(\Omega)} + (g,q)_{L^2(\Omega)} \quad \forall (w,q) \in X \times Q$$

and $(x_k,p_k) \in X_k \times Q_k$ *satisfies*

$$\mathcal{B}((x_k,p_k),(w,q)) = (f,z)_{L^2(\Omega)} + (g,q)_{L^2(\Omega)} \quad \forall (w,q) \in X_k \times Q_k,$$

then the following estimate holds:

$$\|y - y_k\|_{L^2(\Omega)} + h_k\|u - u_k\|_{L^2(\Omega)} + \|p - p_k\|_{L^2(\Omega)} \leq Ch_k^2(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

Proof. First we need an error estimate in the energy norm. From the extension of Céa's Lemma to mixed problems, see Theorem 8, we have:

$$\|x - x_k, p - p_k\|_{X\times Q} \leq C \inf_{(w_k,q_k)\in X_k\times Q_k} \|(x - w_k, p - q_k)\|_{X\times Q}. \tag{4.20}$$

Together with the interpolation properties (4.19) and regularity result (4.17) we obtain:

$$\begin{aligned} \|(x - x_k, p - p_k)\|_{X\times Q} &\leq Ch_k(\|y\|_{H^2(\Omega)} + \|u\|_{H^1(\Omega)} + \|p\|_{H^2(\Omega)}) \\ &\leq Ch_k(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}). \end{aligned} \tag{4.21}$$

The error estimate in the L^2-norm is now obtained by standard duality arguments: Consider the following auxiliary problem: Find $(r,s) \in X \times Q$ such that

$$\mathcal{B}((w,q),(r,s)) = (y - y_k, z)_{L^2(\Omega)} + (p - p_k, q)_{L^2(\Omega)}$$

for all $(w,q) \in X \times Q$.

Using Galerkin orthogonality, the regularity results and the error estimates in the energy norm we obtain:

$$\begin{aligned}
\|y-y_k\|^2_{L^2(\Omega)} + \|p-p_k\|^2_{L^2(\Omega)} &= \mathcal{B}((x-x_k,p-p_k),(r,s)) \\
&= \mathcal{B}((x-x_k,p-p_k),(r,s)-\Pi_k(r,s)) \\
&\leq C\|(x-x_k,p-p_k)\|_{X\times Q}\|(r,s)-\Pi_k(r,s)\|_{X\times Q} \\
&\leq Ch_k^2(\|f\|_{L^2(\Omega)}+\|g\|_{L^2(\Omega)})(\|y-y_k\|_{L^2(\Omega)}+\|p-p_k\|_{L^2(\Omega)}).
\end{aligned}$$

The estimate for $h_k\|u-u_k\|_{L^2(\Omega)}$ follows immediately from (4.20), which completes the proof. □

Assumption (A.8) requires the equivalence of the mesh-dependent norm with an L^2-norm.

Lemma 7. *For all $(w,q) \in X_k \times Q_k$ we have*

$$c|||(w,q)|||_{0,k} \leq \left(\|z\|^2_{L^2(\Omega)} + h_k^2\|v\|^2_{L^2(\Omega)} + \|q\|^2_{L^2(\Omega)}\right)^{1/2} \leq C|||(w,q)|||_{0,k}.$$

Proof. The equivalence of the two norms follows immediately from

$$\underline{c}h_k^2(\underline{u},\underline{u}) \leq \|u\|^2_{L^2(\Omega)} \leq \bar{c}h_k^2(\underline{u},\underline{u}),$$

for quasi-uniform triangulations and nodal finite elements, where $\underline{u}$ is the vector representation of u. This can be shown by standard scaling arguments. □

Finally Assumption (A.9) deals with the intergrid transfer operator $I^k_{k-1}\colon X_{k-1} \times Q_{k-1} \longrightarrow X_k \times Q_k$ from the coarse to the fine grid. Here the finite element spaces are nested, so we can take the natural injection operator and the assumptions (A.9a) and (A.9b) are trivial. It remains to show the last assumption (A.9c), which reads in our case:

Lemma 8. *For all $(w,q) \in (H^2(\Omega)\times H^1(\Omega)\times H^2(\Omega)) \cap (X \times Q)$ we have*

$$|||\Pi_{k-1}(w,q)-\Pi_k(w,q)|||_{0,k} \leq Ch_k^2(\|z\|_{H^2(\Omega)}+\|v\|_{H^1(\Omega)}+\|q\|_{H^2(\Omega)}),$$

with $w=(z,v)$.

Proof. Let us denote $\Pi_k(w,q)=(w_k,q_k)$ and $\Pi_{k-1}(w,q)=(w_{k-1},q_{k-1})$. Using the triangle inequality and the interpolation error estimate (4.18), we get:

$$\begin{aligned}
\|z_{k-1}-z_k\|_{L^2(\Omega)} &\leq \|z_{k-1}-z\|_{L^2(\Omega)} + \|z-z_k\|_{L^2(\Omega)} \\
&\leq ch_k^2\|z\|_{H^2(\Omega)}, \qquad (4.22)
\end{aligned}$$

With the same line of arguments we show:

$$\begin{aligned}
h_k\|v_{k-1}-v_k\|_{L^2(\Omega)} &\leq h_k(\|v_{k-1}-v\|_{L^2(\Omega)} + \|v-v_k\|_{L^2(\Omega)}) \\
&\leq Ch_k^2\|v\|_{H^1(\Omega)}, \qquad (4.23)
\end{aligned}$$

and

$$\begin{aligned} \|q_{k-1} - q_k\|_{L^2(\Omega)} &\leq \|q_{k-1} - q\|_{L^2(\Omega)} + \|q - q_k\|_{L^2(\Omega)} \\ &\leq Ch_k^2 \|q\|_{H^2(\Omega)}. \end{aligned} \tag{4.24}$$

Finally, (4.22),(4.23),(4.24), together with the equivalence of the mesh-dependent norm with the L^2-norm (Lemma 7), completes the proof. □

So, in summary, we have:

Theorem 11. *The approximation property (4.13) is satisfied for the mesh-dependent norm, given by (4.16).*

4.3 Smoothing Property

For the smoothing procedure we have to define appropriate local sub-problems at grid level k. So the prolongation operators and the matrices of the local sub-problems must be specified for the optimal control problem.

Let N_k be the number of nodes of the triangulation $\mathcal{T}_k$. Then we have $n_k = 2N_k$ degrees of freedom for the primal variable $x_k = (y_k, u_k) \in X_k$ and $m_k = N_k$ degrees of freedom for the dual variable $p_k \in Q_k$.

We will now define a space decomposition of $\mathbb{R}^{n_k} \times \mathbb{R}^{m_k}$ into N_k subspaces: For each $i \in \{1, \ldots, N_k\}$ representing a node of the triangulation, let $\mathcal{N}_{k,i}$ be the set of all indices consisting of i and the indices of all neighboring nodes (all nodes which are connected to the node with index i by an edge of the triangulation). Then, for each $i \in \{1, \ldots, N_k\}$, the associated local patch consists of all unknowns of y_k and u_k which are associated to nodes with indices from $\mathcal{N}_{k,i}$ and the unknown of p_k which is associated to the node with index i, see Figure 4.1 for an illustration of a local patch. The corresponding prolongations are the canonical embeddings into $\mathbb{R}^{n_k}$ and $\mathbb{R}^{m_k}$ and are denoted by $\hat{P}_{k,i}$ and $Q_{k,i}$, respectively.

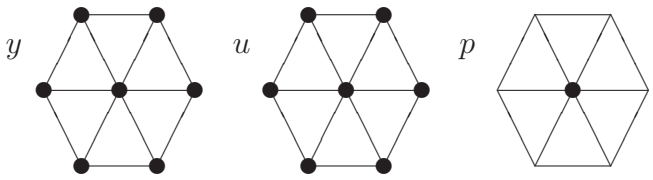


Figure 4.1: Local patches

Observe that all entries in $\hat{P}_{k,i}$ and $Q_{k,i}$ are either 0 or 1. A single component of p_k belongs to exactly one patch, while a single component of y_k or u_k belongs, in general, to more than one patch. It is easy to see that

$$\sum_{i=1}^{N_k} Q_{k,i} Q_{k,i}^T = I \quad \text{and} \quad \sum_{i=1}^{N_k} \hat{P}_{k,i} \hat{P}_{k,i}^T = \mathcal{D}_k,$$

where

$$\mathcal{D}_k = \begin{pmatrix} D_k & 0 \\ 0 & D_k \end{pmatrix}$$

with the $N_k \times N_k$ diagonal matrix D_k whose diagonal entries $d_{k,j}$ are the local overlap depth at the node with index j, i.e., the number of all indices l with $j \in \mathcal{N}_{k,l}$, for $j = 1, \ldots, N_k$. Observe that

$$1 \leq d_{k,j} \leq d_{\max} \quad \text{for all } j = 1, \ldots, N_k \tag{4.25}$$

with a constant $d_{\max}$ independent of the grid level k.

In order to guarantee Condition (3.10), we have to scale the prolongations $\hat{P}_{k,i}$ accordingly:

$$P_{k,i} = \mathcal{D}_k^{-1/2} \hat{P}_{k,i}.$$

Next we have to choose a matrix $\hat{A}_k$, needed on the left-hand side in Condition (3.14). It seems to be natural to choose

$$\hat{A}_k = \frac{1}{\sigma} \operatorname{diag} A_k = \frac{1}{\sigma} \begin{pmatrix} \operatorname{diag} M_k & 0 \\ 0 & \gamma \operatorname{diag} M_k \end{pmatrix} \tag{4.26}$$

with a suitable parameter $\sigma > 0$ as it was done in Schöberl and Zulehner [82] for the Stokes problem. But in order to prove the smoothing property, we have to check, if the estimate (3.19) is fulfilled. This is not the case with this definition of $\hat{A}_k$ for parameters $\sigma = O(1)$, since the mass matrix, representing the $L^2(\Omega)$ inner product, does not scale properly in $H^1(\Omega)$. We have to choose the diagonal part of the discretization matrix corresponding to a positive bilinear form, which is coercive and bounded on the whole space X instead:

$$\hat{A}_k = \frac{1}{\sigma} \begin{pmatrix} \operatorname{diag} K_k & 0 \\ 0 & \gamma \operatorname{diag} M_k \end{pmatrix} \tag{4.27}$$

with σ small enough to ensure

$$\hat{A}_k \geq A_k. \tag{4.28}$$

Since $K_k \geq M_k$ and the maximal number of non-zero entries per row in M_k and K_k is bounded by a constant, say nnz_x, independently of k, it suffices to choose $\sigma = 1/nnz_x$. Here we use the estimate $M \leq nnz(M) \operatorname{diag} M$ for all symmetric and positive definite matrices M, where $nnz(M)$ denotes the maximum number of non-zero entries per row in M.

For the local sub-problems we choose just the restriction of $\hat{A}_k$ to those components of $\underline{x}_k$ whose indices are in $\mathcal{N}_{k,i}$:

$$\hat{A}_{k,i} = \hat{P}_{k,i}^T \hat{A}_k \hat{P}_{k,i}. \tag{4.29}$$

Since the matrices $\hat{A}_k$ and $\hat{A}_{k,i}$ are diagonal the condition (3.14) is satisfied.

The other matrices of the local sub-problems are specified similarly: For

$$B_{k,i} = Q_{k,i}^T B_k \mathcal{D}_k^{1/2} \hat{P}_{k,i} \tag{4.30}$$

one can verify the relation (3.13), since

$$B_{k,i}P_{k,i}^T = Q_{k,i}^T B_k \mathcal{D}_k^{1/2}\, \hat{P}_{k,i}\hat{P}_{k,i}^T \mathcal{D}_k^{-1/2}$$

and the i-th component of $B_k v$ depends only on the degrees of freedom on the surrounding triangles, whose indices are collected in the set $\mathcal{N}_{k,i}$. On this index set $\hat{P}_{k,i}\hat{P}_{k,i}^T$ acts like the identity.

Finally, we set

$$\hat{S}_{k,i} = \frac{1}{\tau}\, B_{k,i}\hat{A}_{k,i}^{-1}B_{k,i}^T$$

with τ small enough to guarantee

$$\hat{S}_k \geq B_k \hat{A}_k^{-1} B_k^T. \tag{4.31}$$

Using (4.29), (4.30), and (4.25), we get

$$\hat{S}_{k,i} = \frac{1}{\tau}\, Q_{k,i}^T B_k \mathcal{D}_k^{1/2} \hat{A}_k^{-1} \mathcal{D}_k^{1/2} B_k^T Q_{k,i} \geq \frac{1}{\tau}\, Q_{k,i}^T B_k \hat{A}_k^{-1} B_k^T Q_{k,i}$$

and with (3.15)

$$\hat{S}_k \geq \frac{1}{\tau}\, \mathrm{diag}(B_k \hat{A}_k^{-1} B_k^T). \tag{4.32}$$

So, in order to guarantee (4.31) it suffices to choose τ such that

$$\frac{1}{\tau}\, \mathrm{diag}(B_k \hat{A}_k^{-1} B_k^T) \geq B_k \hat{A}_k^{-1} B_k^T.$$

Since the maximal number of non-zero entries per row in $B_k \hat{A}_k^{-1} B_k^T$ is bounded by a constant, say nnz_p, independent of k it suffices to choose $\tau = 1/nnz_p$.

The matrix $\mathcal{L}_k$ representing the mesh-dependent norm $|||.|||_{0,k}$ is given by

$$\mathcal{L}_k = h_k^2 \begin{pmatrix} I & 0 & 0 \\ 0 & h_k^2 I & 0 \\ 0 & 0 & I \end{pmatrix}.$$

The last missing part for the smoothing property (4.15) is the estimate (3.19), which will be shown in the next lemma:

Lemma 9. *With the setting from above there is a constant c_R, independent of k, such that*

$$\|\mathcal{Q}_k\|_{\mathcal{L}_k} \leq c_R \, \|\mathcal{K}_k\|_{\mathcal{L}_k} \quad \textit{with} \quad \mathcal{Q}_k = \begin{pmatrix} \hat{A}_k - A_k & 0 \\ 0 & \hat{S}_k - B_k \hat{A}_k^{-1} B_k^T \end{pmatrix}.$$

Proof. Since

$$0 \le \mathcal{Q}_k = \begin{pmatrix} \hat{A}_k - A_k & 0 \\ 0 & \hat{S}_k - B_k \hat{A}_k^{-1} B_k^T \end{pmatrix} \le \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix}$$

we have

$$\|\mathcal{Q}_k\|_{\mathcal{L}_k} = \left\| \begin{pmatrix} \hat{A}_k - A_k & 0 \\ 0 & \hat{S}_k - B_k \hat{A}_k^{-1} B_k^T \end{pmatrix} \right\| \le \left\| \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix} \right\|_{\mathcal{L}_k}.$$

Using the block diagonal form of $\hat{A}_k$ and $\mathcal{L}_k$ we obtain

$$\left\| \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix} \right\|_{\mathcal{L}_k} = \frac{1}{h_k^2} \max\left(\frac{1}{\sigma} \| \operatorname{diag} K_k \|_{\ell^2}, \frac{\gamma}{\sigma\, h_k^2} \| \operatorname{diag} M_k \|_{\ell^2}, \|\hat{S}_k\|_{\ell^2} \right).$$

It follows from (4.25) that

$$\hat{S}_{k,i} = \frac{1}{\tau} Q_{k,i}^T B_k \mathcal{D}_k^{1/2} \hat{A}_k^{-1} \mathcal{D}_k^{1/2} B_k^T Q_{k,i} \le \frac{d_{\max}}{\tau} Q_{k,i}^T B_k \hat{A}_k^{-1} B_k^T Q_{k,i}$$

and, therefore,

$$\hat{S}_k \le \frac{d_{\max}}{\tau} \operatorname{diag}(B_k \hat{A}_k^{-1} B_k^T).$$

Using this estimate and the simple general estimates $\| \operatorname{diag} M\|_{\ell^2} \le \|M\|_{\ell^2}$ for any matrix M and $\|M\|_{\ell^2} \le \|N\|_{\ell^2}$ for symmetric matrices M, N with $0 \le M \le N$ we obtain

$$\left\| \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix} \right\|_{\mathcal{L}_k} \le \frac{1}{h_k^2} \max\left(\frac{1}{\sigma} \|K_k\|_{\ell^2}, \frac{\gamma}{\sigma\, h_k^2} \|M_k\|_{\ell^2}, \frac{d_{\max}}{\tau} \|B_k \hat{A}_k^{-1} B_k^T\|_{\ell^2} \right).$$

Furthermore,

$$\begin{aligned} \|B_k \hat{A}_k^{-1} B_k^T\|_{\ell^2} &= \|\sigma K_k (\operatorname{diag} K_k)^{-1} K_k + \frac{\sigma}{\gamma} M_k (\operatorname{diag} M_k)^{-1} M_k\|_{\ell^2} \\ &\le \sigma \|K_k (\operatorname{diag} K_k)^{-1} K_k\|_{\ell^2} + \frac{\sigma}{\gamma} \|M_k (\operatorname{diag} M_k)^{-1} M_k\|_{\ell^2} \\ &\le \sigma\, nnz_x \|K_k\|_{\ell^2} + \frac{\sigma\, nnz_x}{\gamma} \|M_k\|_{\ell^2} \\ &= \|K_k\|_{\ell^2} + \frac{1}{\gamma} \|M_k\|_{\ell^2} \le \left(1 + \frac{1}{\gamma}\right) \|K_k\|_{\ell^2}. \end{aligned}$$

Summarizing the estimates from above we obtain

$$\|\mathcal{Q}_k\|_{\mathcal{L}_k} \le \frac{c_R}{h_k^2} \max\left(\|K_k\|_{\ell^2}, \frac{\gamma}{h_k^2} \|M_k\|_{\ell^2} \right)$$

with $c_R = \max(1/\sigma, (1 + 1/\gamma) d_{\max}/\tau)$. Finally, using the fact that the spectral norm of a block matrix is greater or equal to the spectral norm of each of its sub-blocks, it follows that

$$\|\mathcal{K}_k\|_{\mathcal{L}_k} \ge \frac{1}{h_k^2} \max\left(\|K_k\|_{\ell^2}, \frac{\gamma}{h_k^2} \|M_k\|_{\ell^2} \right),$$

which completes the proof. □

So, in summary, we have

Theorem 12. *For the additive Schwarz smoother constructed in this section the smoothing property (4.15) holds with a smoothing rate* $\eta(m) = O(1/\sqrt{m})$.

The convergence of the two-grid algorithm is obtained by combining the smoothing and approximation properties.

Theorem 13 (Convergence of the two-grid method)**.** *There exists a constant* $C > 0$ *such that*

$$|||(x_k^{(m+1)} - x_k, p_k^{(m+1)} - p_k)|||_{0,k} \leq Cm^{-1/2}|||(x_k^{(0)} - x_k, p_k^{(0)} - p_k)|||_{0,k},$$

where (x_k, p_k) *is the solution of the discrete problem* (4.9), $(x_k^{(0)}, p_k^{(0)})$ *is the starting value, and* $(x_k^{(m+1)}, p_k^{(m+1)})$ *is the output of the two-grid iteration. Therefore the two-grid method is a contraction with contraction number bounded away from one, independent of the mesh level, if the number of smoothing steps is sufficiently large.*

As already mentioned in Section 3.3, a standard perturbation argument then yield the convergence of the multigrid method.

Theorem 14 (Convergence of the multigrid method, W-cycle)**.** *There exists a constant* $C > 0$ *such that*

$$|||(x_k^{(m+1)} - x_k, p_k^{(m+1)} - p_k)|||_{0,k} \leq Cm^{-1/2}|||(x_k^{(0)} - x_k, p_k^{(0)} - p_k)|||_{0,k},$$

where (x_k, p_k) *is the solution of the discrete problem* (4.9), $(x_k^{(0)}, p_k^{(0)})$ *is the starting value, and* $(x_k^{(m+1)}, p_k^{(m+1)})$ *is the output of the multigrid iteration. Therefore the W-cycle multigrid method is a contraction with contraction number bounded away from one, independent of the mesh level, if the number of smoothing steps is sufficiently large.*

4.4 Numerical Experiments

Next we present some numerical results for the domain $\Omega = (0,1) \times (0,1)$ and homogeneous data $y_d = 0$. The initial grid consists of two triangles by connecting the nodes $(0,0)$ and $(1,1)$. For the first series of experiments the regularization parameter γ was set equal to 1. The dependence of the convergence rate on the regularization parameter γ was investigated subsequently.

Randomly chosen starting values for $x_k^{(0)}$ and $p_k^{(0)}$ for the exact solution $x_k = 0$ and $p_k = 0$ were used. The discretized problem was solved by a multigrid iteration with a W-cycle ($\mu = 2$) and $m/2$ pre- and $m/2$ post-smoothing steps. The multigrid iteration was performed until the Euclidean norm of the solution was reduced by a factor $\varepsilon = 10^{-8}$.

Next we have to choose the parameter σ and τ such that the conditions (4.28) and (4.31) are fulfilled. For this purpose we could estimate the largest eigenvalues of the following eigenvalue problems:

$$M_k u = \lambda \operatorname{diag}(M_k) u \tag{4.33}$$

and

$$B_k \hat{A}_k^{-1} B_k^T p = \lambda \, \mathrm{diag}(B_k \mathcal{D}_k \hat{A}_k^{-1} B_k^T) p. \tag{4.34}$$

These eigenvalues are easily estimated by a few steps of a simple power iteration on a low level, say $k = 3$. In our case we get $\lambda_1 = 0.5$ and $\lambda_2 = 1.87$, where λ_1 and λ_2 denotes the largest eigenvalue of the eigenvalue problem (4.33) and (4.34), respectively. This leads to the choice

$$\sigma = 0.5 \qquad \text{and} \qquad \tau = 1.87.$$

Table 4.1 contains the total number of unknowns n_k+m_k, the number of iterations it and the (average) convergence rates q depending on the level k and the number m of smoothing steps. It shows a typical multigrid convergence behavior, namely the independence of the grid level and the expected improvement of the rates with an increasing number of smoothing steps.

		smoothing steps							
level	$n_k + m_k$	5+5		7+7		10+10		15+15	
5	3 267	49	0.68	33	0.57	23	0.45	18	0.35
6	12 675	52	0.70	37	0.60	25	0.48	18	0.35
7	49 923	52	0.70	38	0.61	26	0.49	18	0.35
8	198 147	53	0.70	38	0.61	26	0.49	18	0.35
9	789 507	53	0.70	38	0.61	26	0.49	18	0.35

Table 4.1: Convergence rates for the additive Schwarz smoother ($\sigma = 0.5,\ \tau = 1.87$)

Numerical experiments show a slightly better convergence rate for a slightly increased relaxation parameter τ. The following tests were performed with $\sigma = 0.5$ and $\tau = 2$. Table 4.2 shows the convergence rates for the additive smoother .

		smoothing steps							
level	$n_k + m_k$	5+5		7+7		10+10		15+15	
5	3 267	46	0.67	30	0.54	21	0.41	17	0.33
6	12 675	48	0.68	34	0.58	24	0.46	17	0.33
7	49 923	49	0.69	35	0.59	25	0.47	17	0.33
8	198 147	49	0.69	35	0.59	25	0.47	17	0.33
9	789 507	49	0.69	35	0.59	25	0.47	17	0.33

Table 4.2: Convergence rates for the additive Schwarz smoother ($\sigma = 0.5,\ \tau = 2$)

Table 4.3 shows the convergence rates of the multiplicative version of the smoother. As expected, the rates are significantly better than the rates for the additive smoother. The number of smoothing steps which are necessary to achieve convergence on all levels is much smaller than in the additive version.

		smoothing steps							
level	$n_k + m_k$	5+5		7+7		10+10		15+15	
5	3 267	21	0.41	19	0.37	16	0.31	12	0.21
6	12 675	22	0.42	20	0.38	16	0.31	12	0.21
7	49 923	22	0.42	19	0.38	16	0.31	12	0.21
8	198 147	22	0.42	20	0.38	16	0.31	12	0.21
9	789 507	22	0.42	20	0.38	16	0.31	12	0.21

Table 4.3: Convergence rates for the multiplicative Schwarz smoother

For comparison, convergence rates are shown for the additive Schwarz smoother based on (4.26) instead of (4.27) in Table 4.4. There is an increase in the number of iterations and there is no clear indication of level-independent convergence rates. This underlines the significance of the modification in the construction of the smoother compared to the original construction in Schöberl and Zulehner [82].

		smoothing steps							
level	$n_k + m_k$	5+5		7+7		10+10		15+15	
5	3 267	45	0.66	38	0.62	31	0.54	23	0.44
6	12 675	47	0.67	40	0.63	32	0.56	24	0.46
7	49 923	51	0.70	43	0.65	34	0.58	26	0.48
8	198 147	54	0.71	45	0.66	36	0.60	27	0.50
9	789 507	57	0.72	48	0.68	36	0.61	29	0.52

Table 4.4: Convergence rates for the unmodified additive Schwarz smoother

All numerical experiments shown so far were performed for the regularization parameter $\gamma = 1$. Observe that the analysis presented here does not predict convergence rates that are robust in γ. Table 4.5 shows the results of numerical experiments obtained at grid level $k = 7$ with 10 pre- and 10 post-smoothing steps for values of γ ranging from 1 down to 10^{-5}. For values of γ below 10^{-5} the number of iterations dramatically increases. Similar results were observed for level $k = 8$ and $k = 9$. The only difference is that the range of values for γ with a reasonable number of iterations increases with increasing level of mesh refinement, i.e., with decreasing mesh size h: The number of iterations does not exceed 25 at level $k = 7$ for $\gamma \in [2 \times 10^{-5}, 1]$, at level $k = 8$ for $\gamma \in [6 \times 10^{-6}, 1]$, and at level $k = 9$ for $\gamma \in [2 \times 10^{-6}, 1]$.

In summary, the numerical experiments confirm the theoretical results of a level-independent convergence rate for the multigrid method with the additive Schwarz smoother. The multiplicative smoother leads to better rates, however, a theoretical analysis for the convergence and smoothing properties is still missing. The modification of the local problems compared to previous work, see Schöberl and Zulehner [82], leads to an

γ	iter.	rate
1	25	0.47
10^{-1}	24	0.46
10^{-2}	24	0.46
10^{-3}	23	0.44
10^{-4}	15	0.28
10^{-5}	35	0.59

Table 4.5: Dependence on the regularization parameter γ

improvement. The range of values for the regularization parameter γ with a reasonable number of iterations increases with decreasing mesh size.

V-cycle convergence of the method proposed in this paper is not covered by the theoretical considerations. Nevertheless, the numerical experiments showed practically identical results for the multigrid iteration with a V-cycle, see Table 4.6.

		smoothing steps							
level	$n_k + m_k$	5+5		7+7		10+10		15+15	
5	3 267	44	0.66	31	0.55	22	0.43	17	0.33
6	12 675	47	0.68	34	0.58	24	0.46	17	0.33
7	49 923	49	0.68	35	0.59	24	0.46	17	0.33
8	198 147	49	0.68	35	0.59	25	0.47	17	0.33
9	789 507	49	0.69	35	0.59	25	0.47	17	0.33

Table 4.6: V-cycle convergence rates for the additive Schwarz smoother ($\sigma = 0.5,\ \tau = 2$)

4.5 Some Remarks

The basic idea of constructing a Schwarz-type smoother carries over to a much wider class of mixed variational problems: An essential step in the construction of the local problems is the choice of the matrix $\hat{A}_k$. Instead of using for $\hat{A}_k$ a multiple of the diagonal part of the discretization matrix A_k corresponding to the bilinear form $a : X \times X \longrightarrow \mathbb{R}$ one should choose the diagonal part of the discretization matrix corresponding to a positive bilinear form on X which is coercive and bounded on the whole space X instead. For the Stokes problem the bilinear form a is already positive, coercive and bounded on X, so the diagonal of A_k will do the job. For the presented optimal control problem the bilinear form a is coercive only on $\ker B$, therefore, $\hat{A}_k$ was chosen differently.

The local matrices are then constructed by restricting the global matrices to the local patches. The requirements of the analysis in terms of certain commutativity relations of

the involved global and local matrices determine the size of the patches and the overlap. Additionally, an appropriate scaling is required which takes into account the local overlap depth of the components of the primal variables. The size of the relaxation parameters can be easily computed in a preprocessing step by a few iterations of a simple power method. This can be done on a much coarser grid such that the computational effort is negligible. In this sense the proposed method can be automatized.

4.6 Other Multigrid Strategies

In Schöberl and Zulehner [83] an iteration method is proposed, which relies on multigrid methods as an inner iteration. The outer iteration is a preconditioned cg-method and the multigrid method is used for approximating the Schur complement. The theory presented there shows that the convergence rates of the iteration method are robust with respect to the discretization parameter and, additionally, also robust with respect to the regularization parameter γ. However, the construction of the preconditioners relies on very accurate approximations of the constants in Brezzi's Theorem, which are rarely available, in general.

There exist also solution methods, which rely on a reduction of the KKT system. In Hackbusch [58], the optimality system is reduced to a single integral equation for the control u and then the multigrid method is applied to this integral equation. The optimality of this method is proven under the condition that the coarse grid is fine enough.

In some sense complementary to this reduction, the control equation can be used to eliminate the control u. Then the resulting optimality conditions consist of two elliptic PDEs in the state y and the adjoint state p. The treatment of this reduced KKT system will be the topic of the next chapter.

Chapter 5

A Reduced KKT System

In the case of the elliptic optimal control problem with distributed control, we can use the control equation in the optimality system to eliminate the control variable. The resulting system can either be written as an elliptic, but nonsymmetric system or again as a symmetric saddle point problem. Both systems are investigated in this chapter. For the elliptic system, known theoretical results are on one hand a quantitative analysis using the local mode analysis. On the other hand, the elliptic system can be seen as a compact perturbation of a decoupled elliptic system. Using the known results of the scalar elliptic case, a multigrid convergence can be proven under the condition that the coarsest grid is fine enough. Both results can be found in Borzi, Kunisch and Kwak [20].

If the KKT system is written in a symmetric saddle point form, the idea of constructing a patch smoother, presented in Section 3.4.2 can be applied. The smoothing property is shown and together with the approximation property this leads to a rigorous multigrid convergence result. By introducing another relaxation parameter in the smoothing iteration, a better numerical behavior is achieved. Using the local mode analysis, the choice of this relaxation parameter can be motivated.

5.1 The Framework

Let us consider again the optimal control problem of the previous chapter, but now we use the control equation (4.2) in the first-order optimality system (4.1)-(4.3) to eliminate the control u. The reduced system consists of the two PDEs, the adjoint state equation and the state equation, where u is replaced by $\gamma^{-1}p$. We consider two equivalent systems. The first one is an elliptic, but nonsymmetric problem (already in weak form): Find $(y,p) \in Y \times Q = H^1(\Omega) \times H^1(\Omega)$ such that

$$\begin{array}{rcll} \gamma(y,z)_{H^1(\Omega)} & - & (p,z)_{L^2(\Omega)} = 0 & \text{for all } z \in Y \\ (y,q)_{L^2(\Omega)} & + & (p,q)_{H^1(\Omega)} = \langle y_d, q\rangle & \text{for all } q \in Q. \end{array}$$

Actually, this is the system, which was solved in Borzi, Kunisch and Kwak [20] using the collective Gauss-Seidel relaxation for the multigrid method. In the mentioned paper,

two theoretical approaches for multigrid convergence analysis are discussed. The first one uses the Fourier analysis for an estimation of the convergence rates. It turns out that they are comparable to the convergence rates for the scalar Poisson equation. In the second approach a perturbation argument shows the h-independent convergence of the multigrid method under the additional condition, that the coarse grid is fine enough. In the proof the optimality system is split into a decoupled system of two scalar elliptic PDEs and a compact perturbation term. The convergence rate for the whole system can then be estimated by the sum of the convergence rate of the decoupled system, which is already known, and a term which depends on the grid size of the coarsest mesh. That means the convergence rate can be bounded away from one if the grid size of the coarsest mesh is sufficiently small.

On the other hand the optimality system can be written in a symmetric, but indefinite form: Find $(p, y) \in Q \times Y = H^1(\Omega) \times H^1(\Omega)$ such that

$$\begin{array}{lcll} (p,q)_{H^1(\Omega)} & + & (y,q)_{L^2(\Omega)} = \langle y_d, q\rangle & \text{for all } q \in Q, \\ (p,z)_{L^2(\Omega)} & - & \gamma(y,z)_{H^1(\Omega)} = 0 & \text{for all } z \in Y. \end{array} \tag{5.1}$$

This mixed variational problem can again be written as a variational problem on $Q \times Y$: Find $(p, y) \in Q \times Y$ such that

$$\mathcal{B}((p,y),(q,z)) = \langle \mathcal{F}, (q,z)\rangle \qquad \text{for all } (q,z) \in Q \times Y$$

with the bilinear form

$$\mathcal{B}((p,y),(q,z)) = (p,q)_{H^1(\Omega)} + (y,q)_{L^2(\Omega)} + (p,z)_{L^2(\Omega)} - \gamma(y,z)_{H^1(\Omega)}$$

and the linear functional

$$\langle \mathcal{F}, (q,z)\rangle = \langle y_d, q\rangle.$$

Lemma 10. *The bilinear form $\mathcal{B}$ is stable and bounded on $Q \times Y$, i.e., there are positive constants c and C such that*

$$c\|(s,r)\|_{Q\times Y} \le \sup_{0\neq(q,z)\in Q\times Y} \frac{\mathcal{B}((s,r),(q,z))}{\|(q,z)\|_{Q\times Y}} \le C\|(s,r)\|_{Q\times Y}$$

for all $(s, r) \in Q \times Y$, where the norm on $Q \times Y$ is given by

$$\|(q,z)\|^2_{Q\times Y} = \|q\|^2_Q + \|z\|^2_Y = \|q\|^2_{H^1(\Omega)} + \|z\|^2_{H^1(\Omega)}.$$

Proof. Boundedness follows immediately from Cauchy's inequality. To show the stability we have:

$$\begin{aligned} \sup_{0\neq(q,z)\in Q\times Y} \frac{\mathcal{B}((s,r),(q,z))}{\|(q,z)\|_{Q\times Y}} &= \sup_{0\neq(q,z)\in Q\times Y} \frac{(s,q)_{H^1(\Omega)} + (r,q)_{L^2(\Omega)} + (s,z)_{L^2(\Omega)} - \gamma(r,z)_{H^1(\Omega)}}{\|(q,z)\|_{Q\times Y}} \\ &\ge \frac{(s,s)_{H^1(\Omega)} + (r,s)_{L^2(\Omega)} - (s,r)_{L^2(\Omega)} + \gamma(r,r)_{H^1(\Omega)}}{\|(s,r)\|_{Q\times Y}} \\ &\ge c\|(s,r)\|_{Q\times Y} \end{aligned}$$

with $c = \min(\gamma, 1)$. □

We discretize again by continuous and piecewise linear finite elements:

$$Q_k = \{q \in C(\bar{\Omega}) : q|_T \in P_1 \text{ for all } T \in \mathcal{T}_k\},$$
$$Y_k = \{z \in C(\bar{\Omega}) : z|_T \in P_1 \text{ for all } T \in \mathcal{T}_k\}.$$

So we obtain the following discrete variational problem: Find $p_k \in Q_k$ and $y_k \in Y_k$ such that

$$\begin{array}{rcll} (p_k, q)_{H^1(\Omega)} & + & (y_k, q)_{L^2(\Omega)} = \langle y_d, q\rangle & \text{for all } q \in Q_k \\ (p_k, z)_{L^2(\Omega)} & - & \gamma(y_k, z)_{H^1(\Omega)} = 0 & \text{for all } z \in Y_k. \end{array} \tag{5.2}$$

which again can be formulated as a discrete variational problem on $Q_k \times Y_k$: Find $(p_k, y_k) \in Q_k \times Y_k$ such that

$$\mathcal{B}((p_k, y_k), (q_k, z_k)) = \langle \mathcal{F}, (q_k, z_k)\rangle \qquad \text{for all } (q_k, z_k) \in Q_k \times Y_k. \tag{5.3}$$

With the same arguments as above the boundedness and stability of the discrete problem can be shown, where the constants c and C are independent of k:

Lemma 11. *The bilinear form $\mathcal{B}$ is uniformly stable and bounded on $Q_k \times Y_k$, i.e., there are positive constants c and C such that*

$$c\|(s,r)\|_{Q\times Y} \le \sup_{0\neq(q,z)\in Q_k\times Y_k} \frac{\mathcal{B}((s,r),(q,z))}{\|(q,z)\|_{Q\times Y}} \le C\|(s,r)\|_{Q\times Y}$$

for all $(s,r) \in Q_k \times Y_k$ and all k.

Proof. The proof is completely analog to the proof of Lemma 10:

Boundedness follows immediately from Cauchy's inequality. To show the stability we have:

$$\begin{aligned} \sup_{0\neq(q,z)\in Q_k\times Y_k} \frac{\mathcal{B}((s,r),(q,z))}{\|(q,z)\|_{Q\times Y}} &= \sup_{0\neq(q,z)\in Q_k\times Y_k} \frac{(s,q)_{H^1(\Omega)} + (r,q)_{L^2(\Omega)} + (s,z)_{L^2(\Omega)} - \gamma(r,z)_{H^1(\Omega)}}{\|(q,z)\|_{Q\times Y}} \\ &\ge \frac{(s,s)_{H^1(\Omega)} + (r,s)_{L^2(\Omega)} - (s,r)_{L^2(\Omega)} + \gamma(r,r)_{H^1(\Omega)}}{\|(s,r)\|_{Q\times Y}} \\ &\ge c\|(s,r)\|_{Q\times Y} \end{aligned}$$

with $c = \min(\gamma, 1)$.

□

Finally, by introducing the standard nodal basis, we obtain the following saddle point problem in matrix-vector notation:

$$\mathcal{K}_k \begin{pmatrix} \underline{p}_k \\ \underline{q}_k \end{pmatrix} = \begin{pmatrix} \underline{f}_k \\ 0 \end{pmatrix} \quad \text{with} \quad \mathcal{K}_k = \begin{pmatrix} K_k & M_k \\ M_k & -\gamma K_k \end{pmatrix}, \tag{5.4}$$

where M_k denotes the mass matrix representing the $L^2(\Omega)$ scalar product on Y_k, and K_k denotes the stiffness matrix representing the $H^1(\Omega)$ scalar product on Y_k.

5.2 Multigrid Convergence Analysis

The general technique for proving the multigrid convergence is the same as in Section 4.1. We briefly adapt the notation and the formulation of the approximation and smoothing property to the present problem.

One iteration step of the multigrid algorithm for solving the discrete problem (5.2) at level k is given in the following form:

Let $(p_k^{(0)}, y_k^{(0)}) \in Q_k \times Y_k$ be a given approximation of the exact solution $(p_k, y_k) \in Q_k \times Y_k$ to (5.2). Then the iteration proceeds in two stages:

1. Smoothing: For $j = 0, 1, \ldots, m-1$ compute $(p_k^{(j+1)}, y_k^{(j+1)}) \in Q_k \times Y_k$ by an iterative procedure of the form
$$(p_k^{(j+1)}, y_k^{(j+1)}) = \mathcal{S}_k\,(p_k^{(j)}, y_k^{(j)}).$$

2. Coarse grid correction: Set
$$\langle \tilde{\mathcal{F}}, (q,z)\rangle = \langle \mathcal{F}, (q,z)\rangle - \mathcal{B}\left((p_k^{(m)}, y_k^{(m)}), (q,z)\right)$$
for $(q,z) \in Q_{k-1} \times Y_{k-1}$ and let $(\tilde{r}_{k-1}, \tilde{s}_{k-1}) \in Q_{k-1} \times Y_{k-1}$ satisfy
$$\mathcal{B}((\tilde{r}_{k-1}, \tilde{s}_{k-1}), (q,z)) = \langle \tilde{\mathcal{F}}, (q,z)\rangle \quad \text{for all } (q,z) \in Q_{k-1} \times Y_{k-1}. \tag{5.5}$$
If $k = 1$, compute the exact solution of (5.5) and set $(r_{k-1}, s_{k-1}) = (\tilde{r}_{k-1}, \tilde{s}_{k-1})$.
If $k > 1$, compute approximations (r_{k-1}, s_{k-1}) by applying $\mu \geq 2$ iteration steps of the multigrid algorithm applied to (5.5) on level $k-1$ with zero starting values.
Set
$$(p_k^{(m+1)}, y_k^{(m+1)}) = (p_k^{(m)}, y_k^{(m)}) + (r_{k-1}, s_{k-1}).$$

For analyzing the convergence of the multigrid method, we have to introduce a pair of norms in which we can show the approximation property and the smoothing property as already mentioned in the previous chapter: the (L^2-like) mesh-dependent norm $|||(q,z)|||_{0,k}$ on $Q_k \times Y_k$ and a second one, defined on $Q_k \times Y_k$ by
$$|||(r,s)|||_{2,k} = \sup_{0 \neq (q,z) \in Q_k \times Y_k} \frac{|\mathcal{B}((r,s),(q,z))|}{|||(q,z)|||_{0,k}}. \tag{5.6}$$

Now we can formulate the approximation property and the smoothing property. Consider the two-grid algorithm (i.e. exact solution of the coarse grid correction equation (5.5) at level $k-1$). It is assumed that there is a constant c_A which is independent of k such that
$$|||\mathcal{B}|||_{0,k}\, |||(p_k^{(m+1)} - p_k, y_k^{(m+1)} - y_k)|||_{0,k} \leq c_A\, |||(p_k^{(m)} - p_k, y_k^{(m)} - y_k)|||_{2,k}, \tag{5.7}$$
where $(p_k, y_k) \in Q_k \times Y_k$ solves (5.3) and the norm $|||\mathcal{B}|||_{0,k}$ of the bilinear form $\mathcal{B}$ is given by
$$|||\mathcal{B}|||_{0,k} = \sup_{0 \neq (r,s),(q,z) \in Q_k \times Y_k} \frac{|\mathcal{B}((r,s),(q,z))|}{|||(r,s)|||_{0,k}\, |||(q,z)|||_{0,k}}.$$

The second part to complete the proof of the two-grid convergence is the smoothing property: It is assumed that

$$|||(p_k^{(m)} - p_k, y_k^{(m)} - y_k)|||_{2,k} \leq \eta(m)\, |||\mathcal{B}|||_{0,k}\, |||(p_k^{(0)} - p_k, y_k^{(0)} - y_k)|||_{0,k} \tag{5.8}$$

for some function $\eta(m)$ which is independent of k, and

$$\eta(m) \to 0 \quad \text{for } m \to \infty.$$

In the next section we will show the approximation property and afterwards we construct the appropriate patch smoother, such that the smoothing property is fulfilled.

5.3 Approximation Property

Let the (L^2-like) mesh-dependent norm on $Q_k \times Y_k$ be given by

$$|||(q,z)|||_{0,k} = h_k \left(\|\underline{q}\|_{\ell^2} + \|\underline{z}\|_{\ell^2} \right)^{1/2}$$

with $(q,z) \in Q_k \times Y_k$ and their vector representations $(\underline{q}, \underline{z})$.

Again by standard scaling arguments we have for quasi-uniform triangulations and nodal finite elements the following norm equivalence:

$$c|||(q,z)|||_{0,k} \leq \left(\|q\|^2_{L^2(\Omega)} + \|z\|^2_{L^2(\Omega)} \right)^{1/2} \leq C|||(q,z)|||_{0,k}. \tag{5.9}$$

An important regularity result is stated in the following lemma.

Lemma 12. *For all $f, g \in L^2(\Omega)$ the variational problem*

$$\mathcal{B}((r,s),(q,z)) = (f,q)_{L^2(\Omega)} + (g,z)_{L^2(\Omega)} \tag{5.10}$$

for all $(q,z) \in Q \times Y$, has a solution $(r,s) \in H^2(\Omega) \times H^2(\Omega)$. There exists a constant $C > 0$ such that

$$\|r\|_{H^2(\Omega)} + \|s\|_{H^2(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

Proof. From (5.10) with $z = 0$ we obtain the elliptic variational problem for r:

$$(r,q)_{H^1(\Omega)} = (f,q)_{L^2(\Omega)} - (s,q)_{L^2(\Omega)} \qquad \text{for all } q \in H^1(\Omega).$$

The elliptic regularity (Theorem 3) and Lemma 10 imply

$$\|r\|_{H^2(\Omega)} \leq C(\|f\|_{L^2(\Omega)} + \|s\|_{L^2(\Omega)}) \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

From (5.10) with $q = 0$ we obtain the elliptic variational problem for s:

$$\gamma(s,z)_{H^1(\Omega)} = (r,z)_{L^2(\Omega)} - (g,z)_{L^2(\Omega)} \qquad \text{for all } z \in H^1(\Omega).$$

The elliptic regularity (Theorem 3) and Lemma 10 imply

$$\|s\|_{H^2(\Omega)} \leq C(\|r\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}) \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}).$$

□

Addtionally we need the following interpolation result:

Lemma 13. *There exists an interpolation operator*

$$\Pi_k \colon (Q \times Y) \cap (H^2(\Omega) \times H^2(\Omega)) \longrightarrow Q_k \times Y_k$$

with the property that

$$\|(r,s) - \Pi_k(r,s)\|_{Q\times Y} \leq ch_k \left(|r|_{H^2(\Omega)} + |s|_{H^2(\Omega)}\right)$$

for all $(r,s) \in (Q \times Y) \cap (H^2(\Omega) \times H^2(\Omega))$.

Proof. We use for each component the standard Lagrange interpolation operators (with piecewise linear polynomials). The desired approximation results then follow immedatialy from Theorem 2. □

The standard estimate of the discretization error is given in the following lemma:

Lemma 14. *Let* $(r,s) \in Q \times Y$ *satisfy:*

$$\mathcal{B}((r,s),(q,z)) = (f,q)_{L^2(\Omega)} + (g,z)_{L^2(\Omega)} \quad \forall (q,z) \in Q \times Y,$$

and let $(r_k,s_k) \in Q_k \times Y_k$ *satisfy:*

$$\mathcal{B}((r_k,s_k),(q,z)) = (f,q)_{L^2(\Omega)} + (g,z)_{L^2(\Omega)} \quad \forall (q,z) \in Q_k \times Y_k,$$

then the following estimate holds:

$$\|(r-r_k, s-s_k)\|_{Q\times Y} \leq ch_k \left(\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)}\right). \tag{5.11}$$

Proof. Due to the extension of Céa's Lemma to mixed problems, see Theorem 8, we can estimate the discretization error by the approximation error:

$$\|(r-r_k, s-s_k)\|_{Q\times Y} \leq C \inf_{(q_k,z_k)\in Q_k\times Y_k} \|(r-q_k, s-z_k)\|_{Q\times Y}.$$

Using $(q_k,z_k) = \Pi_k(r,s)$ and the interpolation result of Lemma 13 yields

$$\|(r-r_k, s-s_k)\|_{Q\times Y} \leq ch_k \left(|r|_{H^2(\Omega)} + |s|_{H^2(\Omega)}\right).$$

The desired statement follows from the regularity result (Lemma 12). □

For given $(\phi_k,\psi_k) \in Q_k \times Y_k$ we define

$$\langle F,(q,z)\rangle = \mathcal{B}((\phi_k,\psi_k),(q,z)) \quad \forall (q,z) \in Q_k \times Y_k,$$

and

$$\begin{aligned}\|F\|_{L^2(\Omega)} &= \sup_{(q,z)\in Q_k\times Y_k} \frac{\mathcal{B}((\phi_k,\psi_k),(q,z))}{\|(q,z)\|_{L^2(\Omega)}} \\ &\leq c \sup_{(q,z)\in Q_k\times Y_k} \frac{\mathcal{B}((\phi_k,\psi_k),(q,z))}{|||(q,z)|||_{0,k}} = c|||(\phi_k,\psi_k)|||_{2,k}, \end{aligned} \tag{5.12}$$

where we used the norm equivalence (5.9) and definition (5.6).

Let $(\phi_{k-1},\psi_{k-1}) \in Q_{k-1}\times Y_{k-1}$ satisfy

$$\mathcal{B}((\phi_{k-1},\psi_{k-1}),(q,z)) = \langle F,(q,z)\rangle \quad \forall (q,z)\in Q_{k-1}\times Y_{k-1},$$

and let $(\phi,\psi)\in Q\times Y$ satisfy

$$\mathcal{B}((\phi,\psi),(q,z)) = \langle F,(q,z)\rangle \quad \forall (q,z)\in Q\times Y.$$

An immediate consequence is the following orthogonality relation:

$$\mathcal{B}((\phi_k-\phi_{k-1},\psi_k-\psi_{k-1}),(q,z)) = 0 \quad \text{for all } (q,z)\in Q_{k-1}\times Y_{k-1}. \tag{5.13}$$

Lemma 15. *There exists a positive constant C such that*

$$|||(\phi_k-\phi_{k-1},\psi_k-\psi_{k-1})|||_{0,k} \leq Ch_k\|(\phi_k-\phi_{k-1},\psi_k-\psi_{k-1})\|_{Q\times Y}, \tag{5.14}$$

Proof. Let us consider the auxiliary problem

$$\mathcal{B}((r,s),(q,z)) = (\phi_k-\phi_{k-1},q)_{L^2(\Omega)} + (\psi_k-\psi_{k-1},z)_{L^2(\Omega)} \qquad \text{for all } (q,z)\in Q\times Y.$$

The symmetry of the bilinear form and the orthogonality relation (5.13) imply

$$\begin{aligned}\|\phi_k-\phi_{k-1}\|^2_{L^2(\Omega)} + \|\psi_k-\psi_{k-1}\|^2_{L^2(\Omega)} &= \mathcal{B}((r,s),(\phi_k-\phi_{k-1},\psi_k-\psi_{k-1})) \\ &= \mathcal{B}((\phi_k-\phi_{k-1},\psi_k-\psi_{k-1}),(r,s)) \\ &= \mathcal{B}((\phi_k-\phi_{k-1},\psi_k-\psi_{k-1}),(r-r_{k-1},s-s_{k-1}))\end{aligned}$$

for arbitrary $(r_{k-1},s_{k-1})\in Q_{k-1}\times Y_{k-1}$. Together with the boundedness we obtain

$$\|\phi_k-\phi_{k-1}\|^2_{L^2(\Omega)}+\|\psi_k-\psi_{k-1}\|^2_{L^2(\Omega)} \leq C\|(\phi_k-\phi_{k-1},\psi_k-\psi_{k-1})\|_{Q\times Y}\|(r-r_{k-1},s-s_{k-1})\|_{Q\times Y}.$$

Now we choose

$$(r_{k-1},s_{k-1}) = \Pi_{k-1}(r,s).$$

The interpolation result (Lemma 13) together with elliptic regularity (Lemma 12) imply

$$\|\phi_k-\phi_{k-1}\|_{L^2(\Omega)} + \|\psi_k-\psi_{k-1}\|_{L^2(\Omega)} \leq ch_k\|(\phi_k-\phi_{k-1},\psi_k-\psi_{k-1})\|_{Q\times Y}. \tag{5.15}$$

Finally, using the norm equivalence (5.9) completes the proof. □

Lemma 16. *There exists a positive constant C such that*

$$\|(\phi_k - \phi_{k-1}, \psi_k - \psi_{k-1})\|_{Q\times Y} \leq ch_k |||(\phi_k, \psi_k)|||_{2,k}. \tag{5.16}$$

Proof. The statement follows directly from the triangle inequality, the discretization error estimate (5.11) and (5.6):

$$\begin{aligned} \|(\phi_k - \phi_{k-1}, \psi_k - \psi_{k-1})\|_{Q\times Y} &\leq \|(\phi - \phi_k, \psi - \psi_k)\|_{Q\times Y} + \|(\phi - \phi_{k-1}, \psi - \psi_{k-1})\|_{Q\times Y} \\ &\leq ch_k \|F\|_{L^2(\Omega)} \leq ch_k |||(\phi_k, \psi_k)|||_{2,k}. \end{aligned}$$

□

The coarse grid correction step is given by

$$(p_k^{(m+1)}, y_k^{(m+1)}) = (p_k^{(m)}, y_k^{(m)}) + (r_{k-1}, s_{k-1})$$

with

$$\mathcal{B}((r_{k-1}, s_{k-1}), (q, z)) = \langle \tilde{\mathcal{F}}, (q, z)\rangle \quad \text{for all } (q, z) \in Q_{k-1} \times Y_{k-1}$$

and

$$\langle \tilde{\mathcal{F}}, (q, z)\rangle = \langle \mathcal{F}, (q, z)\rangle - \mathcal{B}((p_k^{(m)}, y_k^{(m)}), (q, z)) \quad \text{for all } (q, z) \in Q_{k-1} \times Y_{k-1}.$$

Lemma 17 (Approximation property)**.** *There exists a positive constant c_A such that*

$$|||(p_k^{(m+1)} - p_k, y_k^{(m+1)} - y_k)|||_{0,k} \leq c_A h_k^2 |||(p_k^{(m)} - p_k, y_k^{(m)} - y_k)|||_{2,k}$$

Proof. The coarse grid correction (r_{k-1}, s_{k-1}) satisfies

$$\mathcal{B}((r_{k-1}, s_{k-1}), (q, z)) = \mathcal{B}((p_k - p_k^{(m)}, y_k - y_k^{(m)}), (q, z))$$

for all $(q, z) \in Q_{k-1} \times Y_{k-1}$ and

$$(p_k - p_k^{(m+1)}, y_k - y_k^{(m+1)}) = (p_k - p_k^{(m)}, y_k - y_k^{(m)}) - (r_{k-1}, s_{k-1}).$$

With the notations of Lemma 15 applied to $(\phi_k, \psi_k) = (p_k - p_k^{(m)}, y_k - y_k^{(m)})$ we have $(\phi_{k-1}, \psi_{k-1}) = (r_{k-1}, s_{k-1})$ and $(p_k - p_k^{(m+1)}, y_k - y_k^{(m+1)}) = (\phi_k - \phi_{k-1}, \psi_k - \psi_{k-1})$. Then the approximation property directly follows from Lemma 15 and Lemme 16. □

5.4 Smoothing Property

As a smoother we will again construct an additive Schwarz-type smoothing iteration similar to Section 3.4. Again we have to define appropriate local sub-problems. Let N_k be the number of nodes of the triangulation. For each $i = 1, \ldots, N_k$ representing a node of the triangulation, let $\mathcal{N}_{k,i}$ be the set of all indices consisting of i and the indices of all

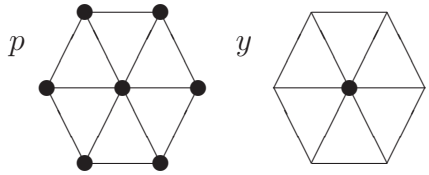


Figure 5.1: local patch

neighboring nodes. Then the associated local patch consists of all unknowns p which are associated to a node with indices from $\mathcal{N}_{k,i}$ and the unknown of y which is associated to the node with index i, see Figure 5.1 for an illustration of a local patch. The corresponding prolongations are the canonical embeddings into $\mathbb{R}^{n_k}$ and $\mathbb{R}^{m_k}$ and are denoted by $\hat{P}_{k,i}$ and $Q_{k,i}$, respectively. We have

$$\sum_{i=1}^{N_k} Q_{k,i} Q_{k,i}^T = I \quad \text{and} \quad \sum_{i=1}^{N_k} \hat{P}_{k,i} \hat{P}_{k,i}^T = \mathcal{D}_k$$

with the $N_k \times N_k$ diagonal matrix $\mathcal{D}_k$ whose diagonal entries $d_{k,j}$ are the local overlap depth at the node with index j. Again

$$1 \leq d_{k,j} \leq d_{\max} \quad \text{for all } j = 1, \ldots, N_k \tag{5.17}$$

with a constant $d_{\max}$ independent of the grid level k.

For the scaled prolongation $P_{k,i}$ given by

$$P_{k,i} = \mathcal{D}_k^{-1/2} \hat{P}_{k,i}$$

the condition (3.10) is fulfilled.

The next step is to choose the matrix $\hat{A}_k$. Due to condition (3.14) we are forced to choose diagonal matrices. Since A_k is positive on the whole space, the natural choice

$$\hat{A}_k = \frac{1}{\sigma} \operatorname{diag} A_k = \frac{1}{\sigma} \operatorname{diag} K_k$$

is suitable. Here σ is chosen small enough, such that

$$\hat{A}_k \geq A_k. \tag{5.18}$$

Since the maximal number of non-zero entries per row in K_k is bounded by a constant, say nnz_p, independently of the grid level k, it is sufficient to choose $\sigma = 1/nnz_p$.

For the local problems we define the local matrices as follows:

$$\begin{aligned} \hat{A}_{k,i} &= \hat{P}_{k,i}^T \hat{A}_k \hat{P}_{k,i}, \\ B_{k,i} &= Q_{k,i}^T B_k \mathcal{D}_k^{1/2} \hat{P}_{k,i}, \\ C_{k,i} &= Q_{k,i}^T C_k Q_{k,i}. \end{aligned}$$

It follows immediately that $\hat{A}_{k,i}$ and $B_{k,i}$ fulfill the commutativity conditions (3.14) and (3.13), respectively.

It remains to choose the local matrix $\hat{S}_{k,i}$. We set

$$\hat{S}_{k,i} = \frac{1}{\tau}\left(C_{k,i} + B_{k,i}\hat{A}_{k,i}^{-1}B_{k,i}^T\right)$$

with τ small enough to ensure

$$\hat{S}_k \geq C_k + B_k\hat{A}_k^{-1}B_k^T. \tag{5.19}$$

Using the definitions of the matrices $\hat{A}_{k,i}$ and $B_{k,i}$, and relation (5.17), we obtain

$$\begin{aligned}
\hat{S}_{k,i} &= \frac{1}{\tau}\left(C_{k,i} + B_{k,i}\hat{A}_{k,i}^{-1}B_{k,i}^T\right) \\
&= \frac{1}{\tau}\left(Q_{k,i}^T C_{k,i}Q_{k,i} + Q_{k,i}^T B_{k,i}\mathcal{D}_k^{1/2}\hat{P}_{k,i}\left(\hat{P}_{k,i}^T\hat{A}_k\hat{P}_{k,i}\right)^{-1}\hat{P}_{k,i}^T\mathcal{D}^{1/2}B^TQ_{k,i}\right) \\
&= \frac{1}{\tau}Q_{k,i}^T\left(C_k + B_k\mathcal{D}_k^{1/2}\hat{A}_k^{-1}\mathcal{D}_k^{1/2}B_k^T\right)Q_{k,i} \\
&\geq \frac{1}{\tau}Q_{k,i}^T\left(C_k + B_k\hat{A}_k^{-1}B_k^T\right)Q_{k,i}.
\end{aligned} \tag{5.20}$$

Since $\hat{S}_k$ is constructed by

$$\hat{S}_k = \left(\sum_{i=1}^{N_k} Q_{k,i}\hat{S}_{k,i}^{-1}Q_{k,i}^T\right)^{-1},$$

we get from (5.20):

$$\hat{S}_k \geq \frac{1}{\tau}\operatorname{diag}\left(C_k + B_k\hat{A}_k^{-1}B_k^T\right). \tag{5.21}$$

That means, it suffices to choose τ such that

$$\frac{1}{\tau}\operatorname{diag}\left(C_k + B_k\hat{A}_k^{-1}B_k^T\right) \geq C_k + B_k\hat{A}_k^{-1}B_k^T.$$

Since the maximal number of non-zero entries per row in $C_k + B_k\hat{A}_k^{-1}B_k^T$ is bounded by a constant, say nnz_y, independent of the grid level k it is sufficient to choose $\tau = 1/nnz_y$.

With this setting, the smoothing iteration can be represented as a global Richardson method

$$\begin{pmatrix}\underline{p}_k^{(j+1)}\\ \underline{y}_k^{(j+1)}\end{pmatrix} = \begin{pmatrix}\underline{p}_k^{(j)}\\ \underline{y}_k^{(j)}\end{pmatrix} + \hat{\mathcal{K}}_k^{-1}\left[\begin{pmatrix}\underline{f}_k\\ 0\end{pmatrix} - \mathcal{K}_k\begin{pmatrix}\underline{p}_k^{(j)}\\ \underline{y}_k^{(j)}\end{pmatrix}\right], \quad \text{with} \quad \hat{\mathcal{K}}_k = \begin{pmatrix}\hat{A}_k & B_k^T\\ B_k & B_k\hat{A}_k^{-1}B_k^T - \hat{S}_k\end{pmatrix} \tag{5.22}$$

and the assumptions of Theorem 10 are satisfied. The last missing part for the smoothing property (4.15) is the estimate

$$\|\mathcal{Q}_k\|_{\mathcal{L}_k} \leq c_R\|\mathcal{K}_k\|_{\mathcal{L}_k} \quad \text{with} \quad \mathcal{Q}_k = \begin{pmatrix}\hat{A}_k - A_k & 0\\ 0 & \hat{S}_k - C_k - B_k\hat{A}_k^{-1}B_k^T\end{pmatrix}, \tag{5.23}$$

where the matrix $\mathcal{L}_k$ representing the mesh-dependent norm $|||\cdot|||_{0,k}$ is given by

$$\mathcal{L}_k = h_k^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

Lemma 18. *With the setting from above there is a constant c_R, independent of the grid level k, such that*

$$\|\mathcal{Q}_k\|_{\mathcal{L}_k} \leq c_R \|\mathcal{K}_k\|_{\mathcal{L}_k} \quad \textit{with} \quad \mathcal{Q}_k = \begin{pmatrix} \hat{A}_k - A_k & 0 \\ 0 & \hat{S}_k - C_k - B_k \hat{A}_k^{-1} B_k^T \end{pmatrix}.$$

Proof. Since

$$0 \leq \begin{pmatrix} \hat{A}_k - A_k & 0 \\ 0 & \hat{S}_k - C_k - B_k \hat{A}_k^{-1} B_k^T \end{pmatrix} \leq \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix}$$

we have

$$\|\mathcal{Q}_k\|_{\mathcal{L}_k} = \left\| \begin{pmatrix} \hat{A}_k - A_k & 0 \\ 0 & \hat{S}_k - C_k - B_k \hat{A}_k^{-1} B_k^T \end{pmatrix} \right\|_{\mathcal{L}_k} \leq \left\| \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix} \right\|_{\mathcal{L}_k}.$$

Using the block diagonal from of $\mathcal{L}_k$ we get

$$\left\| \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix} \right\|_{\mathcal{L}_k} = \frac{1}{h_k^2} \max \left(\frac{1}{\sigma} \| \operatorname{diag} K_k \|_{\ell^2}, \|\hat{S}_k\|_{\ell^2} \right).$$

From (5.17) it follows that

$$\hat{S}_{k,i} = \frac{1}{\tau} Q_{k,i}^T \left(C_k + B_k \mathcal{D}_k^{1/2} \hat{A}_k^{-1} \mathcal{D}_k^{1/2} B_k^T \right) Q_{k,i} \leq \frac{d_{\max}}{\tau} Q_{k,i}^T \left(C_k + B_k \hat{A}_k^{-1} B_k^T \right) Q_{k,i},$$

and from the construction of $\hat{S}_k$ we have

$$\hat{S}_k \leq \frac{d_{\max}}{\tau} \operatorname{diag} \left(C_k + B_k \hat{A}_k^{-1} B_k^T \right).$$

Since $\| \operatorname{diag} M \|_{\ell^2} \leq \|M\|_{\ell^2}$ for any matrix M we obtain

$$\left\| \begin{pmatrix} \hat{A}_k & 0 \\ 0 & \hat{S}_k \end{pmatrix} \right\|_{\mathcal{L}_k} \leq \frac{1}{h_k^2} \max \left(\frac{1}{\sigma} \|K_k\|_{\ell^2}, \frac{d_{\max}}{\tau} \|C_k + B_k \hat{A}_k^{-1} B_k^T\|_{\ell^2} \right).$$

From (5.18) and $K_k \geq M_k$ it follows that

$$\begin{aligned} \|C_k + B_k \hat{A}_k^{-1} B_k^T\|_{\ell^2} &= \left\| \gamma K_k + M_k \left(\frac{1}{\sigma} \operatorname{diag} K_k \right)^{-1} M_k \right\|_{\ell^2} \\ &\leq \gamma \|K_k\|_{\ell^2} + \|M_k\|_{\ell^2} \|K_k^{-1}\|_{\ell^2} \|M_k\|_{\ell^2} \\ &\leq \gamma \|K_k\|_{\ell^2} + \|K_k\|_{\ell^2}. \end{aligned}$$

Here we used the fact that $\|M\|_{\ell^2} \leq \|N\|_{\ell^2}$ for any symmetric matrices M,N with $0 \leq M \leq N$.

Summarizing all the estimates we get

$$\|\mathcal{Q}_k\|_{\mathcal{L}_k} \leq \frac{c_R}{h_k^2}\|K_k\|_{\ell^2}, \quad \text{with } c_R = \max\left(\frac{1}{\sigma}, (1+\gamma)\frac{d_{\max}}{\tau}\right). \tag{5.24}$$

Since the spectral norm of a sub-block is less or equal than the spectral norm of the whole block matrix we have

$$\frac{1}{h_k^2}\|K_k\|_{\ell^2} \leq \|\mathcal{K}_k\|_{\mathcal{L}_k},$$

which, together with (5.24), completes the proof. □

In summary, we have

Theorem 15. *For the additive Schwarz smoother constructed in this section the smoothing property* (5.8) *holds with a smoothing rate* $\eta(m) = \mathcal{O}(1/\sqrt{m})$.

Combining the smoothing property and the approximation property we obtain the two-grid convergence.

Theorem 16 (Convergence of the two-grid method)**.** *There exists a constant* $C > 0$ *such that*

$$|||(p_k^{(m+1)} - p_k, y_k^{(m+1)} - y_k)|||_{0,k} \leq Cm^{-1/2}|||(p_k^{(0)} - p_k, y_k^{(0)} - y_k)|||_{0,k},$$

where (p_k, y_k) *is the solution of the discrete problem* (5.3)*,* $(p_k^{(0)}, y_k^{(0)})$ *is the initial value, and* $(p_k^{(m+1)}, y_k^{(m+1)})$ *is the output of the two-grid iteration. Therefore, the two-grid method is a contraction with contraction number bounded away from one, independent of the grid level, if the number of smoothing steps is sufficiently large.*

Again, the two-grid convergence induces the multigrid convergence of the W-cycle.

Theorem 17 (Convergence of the multigrid method, W-cycle)**.** *There exists a constant* $C > 0$ *such that*

$$|||(p_k^{(m+1)} - p_k, y_k^{(m+1)} - y_k)|||_{0,k} \leq Cm^{-1/2}|||(p_k^{(0)} - p_k, y_k^{(0)} - y_k)|||_{0,k},$$

where (p_k, y_k) *is the solution of the discrete problem* (5.3)*,* $(p_k^{(0)}, y_k^{(0)})$ *is the initial guess, and* $(p_k^{(m+1)}, y_k^{(m+1)})$ *is the output of the multigrid iteration. Therefore the W-cycle multigrid method is a contraction with contraction number bounded away from one, independent of the grid level* k*, if the number of smoothing steps is sufficiently large.*

5.5 An Overrelaxed Method

Let us introduce an additional overrelaxation parameter $\omega > 0$ in the iteration (5.22), which leads to the following method:

$$\begin{pmatrix} p_k^{(j+1)} \\ y_k^{(j+1)} \end{pmatrix} = \begin{pmatrix} p_k^{(j)} \\ y_k^{(j)} \end{pmatrix} + \omega \hat{\mathcal{K}}_k^{-1} \left[\begin{pmatrix} f_k \\ 0 \end{pmatrix} - \mathcal{K}_k \begin{pmatrix} p_k^{(j)} \\ y_k^{(j)} \end{pmatrix} \right]. \tag{5.25}$$

Until now we considered the smoothing method with $\omega = 1$. Theorem 2 of SCHÖBERL AND ZULEHNER [82] states that the relaxed method (5.25) converges for all relaxation parameters $\omega \in (0, 2)$ and the iterates do not blow up in the limiting case $\omega = 2$. Then Reusken's lemma, see REUSKEN [77], HACKBUSCH [60], implies the smoothing property for the original method (with $\omega = 1$).

In ECKER AND ZULEHNER [49] an extension of Reusken's lemma was proven, which can be used to show the smoothing property of the relaxed method (5.25) for all $\omega < 2$. But the smoothing analysis do not give any hint that it could be worthwile to use the relaxed smoothing iteration with $\omega \neq 1$. Nevertheless, we will see in the next section, that we can improve the convergence rates by using a relaxation parameter $\omega > 1$. In the following we will use the local mode analysis to motivate a different choice of the relaxation parameter.

We analyze the smoothing iteration by local Fourier analysis, see Section 3.5. Consider the Fourier space spanned by the functions

$$\phi(\theta, x) = e^{i\theta_1 x_1/h} e^{i\theta_2 x_2/h}, \quad \theta = (\theta_1, \theta_2),\ x = (x_1, x_2).$$

We have to compute the symbol of the smoothing iteration matrix

$$\overline{M}_S = \overline{I - \omega \hat{\mathcal{K}}^{-1} \mathcal{K}},$$

with

$$\mathcal{K} = \begin{pmatrix} K & M \\ M & -\gamma K \end{pmatrix} \quad \text{and} \quad \hat{\mathcal{K}} = \begin{pmatrix} 2\,\mathrm{diag}\, K & M \\ M & M(2\,\mathrm{diag}\, K)^{-1} M - \hat{S} \end{pmatrix},$$

where

$$\hat{S} = 2\,\mathrm{diag}\left(\gamma K + M\mathcal{D}(2\,\mathrm{diag}\, K)^{-1} M\right).$$

Here we used the parameter choice $\sigma = \tau = 1/2$, such that the conditions (5.18) and (5.19) are fulfilled.

Since the discrete operators M and K can be written in stencil form, the Fourier symbols $\overline{M}$ and $\overline{K}$ can be easily computed from the stencil form and are given by

$$\begin{aligned} \overline{M} &= \frac{h^2}{6} \left(\cos\theta_1 + \cos(\theta_1 + \theta_2) + \cos(\theta_1 - \theta_2) + \cos\theta_2 + 4\right), \\ \overline{K} &= \overline{M} + 4 - 2\cos\theta_1 - 2\cos\theta_2. \end{aligned}$$

The Fourier symbols of the remaining discrete operators are just combinations of the two fundamental symbols $\overline{M}$ and $\overline{K}$. The symbols of the discrete operators $\mathcal{K}$ and $\hat{\mathcal{K}}$ are then the following 2×2-matrices:

$$\begin{aligned} \overline{\mathcal{K}} &= \begin{pmatrix} \overline{K} & \overline{M} \\ \overline{M} & -\gamma\overline{K} \end{pmatrix}, \\ \overline{\hat{\mathcal{K}}} &= \begin{pmatrix} 2(4+\frac{2}{3}h^2) & \overline{M} \\ \overline{M} & \overline{\hat{\mathcal{K}}_{22}} \end{pmatrix} \end{aligned}$$

with

$$\overline{\hat{\mathcal{K}}_{22}} = \frac{1}{48}\frac{h^4\left(\cos\theta_1 + \cos(\theta_1+\theta_2) + \cos\theta_2 + \cos(\theta_1-\theta_2) + 4\right)^2}{6+h^2} - \frac{2}{3}\frac{9h^4 + 72\gamma + 24\gamma h^2 + 2\gamma h^4}{6+h^2}.$$

Now any symbolic package can be used to compute the symbol of $I - \omega\hat{\mathcal{K}}^{-1}\mathcal{K}$ and the eigenvalues of this matrix. In the limit case $h \to 0$ we finally obtain for the eigenvalues of $\overline{M}_S$:

$$\lambda_1 = \lambda_2 = 1 - \frac{\omega}{4}\left(2 - \cos\theta_1 - \cos\theta_2\right).$$

The smoothing factor μ_{loc} is then given by

$$\mu_{\text{loc}}(\omega) = \max\left\{\left|1 - \frac{\omega}{4}\right|, |1-\omega|\right\}.$$

In order to obtain the best smoothing properties, the term $\mu_{\text{loc}}(\omega)$ has to be minimized over the admissible range of the relaxation parameter ω:

$$\min_{\omega\in(0,2)} \mu_{\text{loc}}(\omega) = \mu_{\text{loc}}(\omega_{\text{opt}}) = \frac{3}{5}, \quad \text{with } \omega_{\text{opt}} = \frac{8}{5}.$$

For the choice $\omega = 1$ we get $\mu_{\text{loc}}(1) = 0.75$, i.e., we expect a better performance of the multigrid iteration while using the overrelaxed method with $\omega = 1.6$. This will be confirmed by numerical experiments in the next section.

Remark 7. *The smoothing factor estimate is independent of the regularization paramter γ. This is a strong indication that the multigrid method is additionally robust with respect to the regularization parameter γ. Nevertheless, the theory presented here is not able to show this robustness.*

5.6 Numerical Results

We present some numerical results for the unit square $\Omega = (0,1)\times(0,1)$ and homogeneous desired state $y_d = 0$. The coarsest grid consists of two triangles by connecting the nodes $(0,0)$ and $(1,1)$. The regularization parameter γ was set equal to 1. The starting values

for $p_k^{(0)}$ and $y_k^{(0)}$ for the exact solution $p_k = 0$ and $y_k = 0$ were randomly chosen. The discretized problem was solved by a multigrid iteration with a W-cycle and m pre- and m post-smoothing steps. The multigrid iteration was performed until the Euclidean norm of the solution was reduced by a factor $\varepsilon = 10^{-8}$.

Next we have to choose the parameter σ and τ in order to fulfill the conditions (5.18) and (5.19). This means we have to estimate the largest eigenvalues of the following eigenvalue problems:

$$K_k p = \lambda \operatorname{diag}(K_k) p, \tag{5.26}$$

and

$$(\gamma K_k + M_k(\frac{1}{\sigma}\operatorname{diag} K_k)^{-1} M_k) y = \lambda \operatorname{diag}(\gamma K_k + M_k \mathcal{D}_k(\frac{1}{\sigma}\operatorname{diag} K_k)^{-1} M_k) y. \tag{5.27}$$

These eigenvalues are easily estimated by a few steps of a power iteration on a low level. In this case we get $\lambda_1 = \lambda_2 = 0.5$, where λ_1 and λ_2 denotes the largest eigenvalue of the eigenvalue problems (5.26) and (5.27). So we choose

$$\sigma = \tau = \frac{1}{2}.$$

Table 5.1 contains the total number of unknowns $n_k + m_k$, the number of iterations and the (average) convergence rates depending on the level k and the number m of smoothing steps. The typical multigrid convergence behavior, namely the independence of the grid level k and the improvement of the convergence rates with an increasing number of smoothing steps, can be clearly seen.

		smoothing steps							
level	$n_k + m_k$	1+1		2+2		3+3		5+5	
5	2 178	27	0.502	14	0.267	10	0.156	7	0.068
6	8 450	27	0.502	14	0.264	10	0.156	7	0.068
7	33 282	27	0.502	14	0.265	11	0.162	7	0.069
8	132 098	27	0.502	14	0.265	10	0.158	7	0.069
9	526 338	27	0.502	14	0.265	11	0.163	7	0.069

Table 5.1: Convergence rates for the additive Schwarz smoother ($\sigma = 0.5,\ \tau = 0.5$)

In Section 5.5 we discussed the choice of the relaxation parameter ω. Table 5.2 shows the convergence rates for the overrelaxed smoother with $\omega = 1.6$. The convergence rates are significantly better than with the original smoothing iteration with $\omega = 1$.

For comparison we used the point smoother (as proposed in Borzi, Kunisch and Kwak [20]) first as an additive version. For $h \to 0$ the system matrix in (5.4) degenerates to a two-fold Poisson problem, which motivates the choice $\omega = 0.8$ as the relaxation parameter in the Jacobi-type point smoother, since $\omega = 0.8$ is the optimal choice of the

		smoothing steps							
level	$n_k + m_k$	1+1		2+2		3+3		5+5	
5	2 178	16	0.301	9	0.127	7	0.067	5	0.023
6	8 450	16	0.302	9	0.128	7	0.066	5	0.024
7	33 282	16	0.302	10	0.135	7	0.067	5	0.024
8	132 098	16	0.302	10	0.135	7	0.067	5	0.024
9	526 338	16	0.302	10	0.135	7	0.068	5	0.024

Table 5.2: Convergence rates for the overrelaxed Schwarz smoother with $\omega = 1.6$

relaxation parameter for the Jacobi smoother for the scalar Poisson problem. Table 5.3 shows the convergence rates, which are comparable to the ones obtained by the overrelaxed additive Schwarz-type smoother.

		smoothing steps							
level	$n_k + m_k$	1+1		2+2		3+3		5+5	
5	2 178	16	0.301	9	0.127	7	0.067	5	0.023
6	8 450	16	0.302	9	0.128	7	0.066	5	0.024
7	33 282	16	0.302	10	0.135	7	0.067	5	0.024
8	132 098	16	0.302	10	0.135	7	0.067	5	0.024
9	526 338	16	0.302	10	0.135	7	0.068	5	0.024

Table 5.3: Convergence rates using Jacobi-type point smoother, $\omega = 0.8$

Table 5.4 shows the convergence rates with the multiplicative version of the patch smoother. As expected, the rates are significantly better than the rates for the additive smoother.

		smoothing steps							
level	$n_k + m_k$	1+1		2+2		3+3		5+5	
5	2 178	15	0.290	7	0.057	5	0.021	4	0.0082
6	8 450	15	0.288	7	0.055	5	0.020	4	0.0077
7	33 282	15	0.290	7	0.057	5	0.020	4	0.0078
8	132 098	15	0.291	7	0.057	5	0.020	4	0.0080
9	526 338	15	0.291	7	0.057	5	0.020	4	0.0080

Table 5.4: Convergence rates for the multiplicative Schwarz smoother

For comparison Table 5.5 shows the convergence rates for the multiplicative version of the point smoother. The numerical results for the multiplicative version of the patch smoother are slightly better than for the collective Gauss-Seidel case.

		smoothing steps							
level	$n_k + m_k$	1+1		2+2		3+3		5+5	
5	2 178	16	0.299	7	0.059	5	0.022	4	0.0092
6	8 450	16	0.299	7	0.058	5	0.022	4	0.0084
7	33 282	16	0.301	7	0.060	5	0.022	4	0.0086
8	132 098	16	0.301	7	0.060	5	0.022	4	0.0087
9	526 338	16	0.301	7	0.060	5	0.022	4	0.0087

Table 5.5: Convergence rates for the multiplicative point smoother

The numerical experiments shown so far were performed for the regularization parameter $\gamma = 1$. We already mentioned that the theory presented here does not predict convergence rates that are robust in γ. Nevertheless, due to the Fourier analysis presented in Section 5.5, we would expect that the method proposed here is robust in the regularization parameter γ. Table 5.6 shows the expected behavior for different choices of γ.

γ	Iter.	conv. rate
1	16	0.302
10^{-2}	16	0.302
10^{-4}	16	0.302
10^{-6}	16	0.302

Table 5.6: γ-dependence, Patch smoother, $\rho = 1.6$, additive, 1-1 smoothing steps

Finally, Table 5.7 shows the convergence rates for a V-cycle iteration using the overrelaxed additive Schwarz smoother. The convergence is not covered by the theory presented here, but nevertheless, the convergence rates are almost identical as for the W-cycle.

		smoothing steps							
level	$n_k + m_k$	1+1		2+2		3+3		5+5	
5	2 178	16	0.303	9	0.129	7	0.067	5	0.023
6	8 450	16	0.302	9	0.129	7	0.066	5	0.024
7	33 282	16	0.305	10	0.136	7	0.067	5	0.024
8	132 098	16	0.304	10	0.135	7	0.067	5	0.024
9	526 338	16	0.302	10	0.135	7	0.068	5	0.024

Table 5.7: V-cycle convergence rates for the overrelaxed Schwarz smoother with $\omega = 1.6$

In summary, the numerical experiments confirm the theoretical results of a level-independent convergence for the multigrid method with the additive Schwarz smoother. Moreover, tuning the relaxation parameter ω motivated by the local Fourier analysis leads

to a computational performance which is comparable to known results from the literature. The multiplicative smoother leads to better rates, however, a theoretical analysis is still missing. Moreover, the numerical experiments show a robustness in the regularization paramter γ, which can not be predicted by the theoretical results.

Chapter 6

Conclusions

We presented in this thesis a multigrid iteration method for solving discrete mixed problems. The focus lied on solving KKT systems arising from PDE-constrained optimization problems. Using the general approach for constructing a patch smoother, as the most important ingredient of a multigrid method, we were able to give a rigorous convergence analysis for a typical optimal control problem. The construction of the smoother itself carries over to a much wider class of mixed variational problems. The choice of the corresponding local matrices is straightforward and the relaxation parameters can be determined by a simple (and cheap) power iteration for estimating eigenvalues of the corresponding generalized eigenvalue problems. In particular, no structurial information of some Schur complement is needed. So this method can almost be regarded as a black-box approach for solving mixed variational problems.

As for instance in the optimal control problem presented here, there are often some regularization parameters involved. It would be desirable to achieve robustness of the proposed method with respect to these parameters. The theory presented here is not able to prove such a robustness. In Schöberl and Zulehner [83] a robust method is proposed for the special class of optimal control problems with distributed control. The inner products and norms defined there are certainly the suitable candidates. But we were not able to show multigrid convergence independent of this regularization parameter so far. The numerical experiments show the dependence with respect to the regularization parameter in the following sense: There exists a range of values with a reasonable convergence rate. But decreasing the value of the regularization parameter further deteriorate the numerical performance dramatically. The range of values increases with decreasing mesh size.

The story is a bit different in the case of the reduced KKT system of Chapter 5. Here, the numerical experiments certainly indicates, that the convergence rate of the proposed method is independent not only of the mesh parameter, but also of the regularization parameter. Nevertheless, a theoretical analysis for the robustness is still missing. The future work will focus on this direction.

Bibliography

[1] R.A. Adams. *Sobolev Spaces*. Academic Press, New York, 1976.

[2] E. Arian and S. Ta'asan. Multigrid one-shot methods for optimal control problems: Infinite dimensional control. Technical report, NASA Langley Research Center, Hampton VA, 1994.

[3] K. Arrow, L. Hurwicz, and H. Uzawa. *Studies in Linear and Nonlinear Programming*. Stanford University Press, Stanford, CA, 1958.

[4] K.J. Arrow and L. Hurwicz. Gradient method for concave programming. I: Local results. In K.J. Arrow, L. Hurwicz, and H. Uzawa, editors, *Studies in Linear and Nonlinear Programming*, pages 117–126. Stanford University Press, Stanford, CA, 1958.

[5] G.P. Astrakhantsev. An iterative method for solving elliptic net problems. *USSR Comput. Math. Math. Phys.*, 11(2):171–182, 1971.

[6] G.P. Astrakhantsev. Analysis of algorithms of the Arrow-Hurwicz type. *Comput. Math. Phys.*, 41:15–26, 2001.

[7] O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, second edition, 1996.

[8] I. Babuška and A.K. Aziz. Survey lectures on the mathematical foundations of the finite element method. In A.K. Aziz, editor, *The Mathematical Foundation of the Finite Element Method with Applications to Partial Differential Equations*, pages 3–363. Academic Press, New York - London, 1972.

[9] I. Babuška. The finite element method with Lagrangian multipliers. *Numerische Mathematik*, 20:179–192, 1973.

[10] N.S. Bachvalov. On the convergence of a relaxation method with natural constraints on the elliptic operator. *USSR Comput. Math. Math. Phys.*, 6(5):101–135, 1966.

[11] 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.

[12] 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.

[13] A. Battermann and M. Heinkenschloss. Preconditioners for Karush-Kuhn-Tucker systems arising in the optimal control of distributed systems. In W. Desch, F. Kappel, and K. Kunisch, editors, *Optimal Control of Partial Differential Equations*, pages 15–32. Birkäuser Verlag, 1998.

[14] A. Battermann and E. W. Sachs. Block preconditioners for KKT systems arising in PDE-governed optimal control problems. In K.-H. Hoffmann, R. H. W. Hoppe, and W. Schulz, editors, *Fast Solution of Discretized Optimization Problems*, pages 1–18. Birkhäuser Verlag, 2001.

[15] M.P. Bendsøe and O. Sigmund. *Topology Optimization: Theory, Methods and Applications*. Springer, Berlin, 2003.

[16] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[17] G. Biros and O. Ghattas. Parallel Lagrange-Newton-Krylov methods for PDE-constrained optimization. Part I: The Krylov-Schur solver. *SIAM J. Sci. Comput.*, 27(2):687–713, 2005.

[18] A. Borzi. Smoothers for control- and state-constrained optimal control problems. *Comput Visual Sci*, 11:59–66, 2008.

[19] A. Borzi and K. Kunisch. A multigrid scheme for elliptic constrained optimal control problems. *Comp. Opt. Appl.*, 31:309–333, 2005.

[20] A. Borzi, K. Kunisch, and D.Y. Kwak. Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system. *SIAM J. Control Optimization*, 41(5):1477–1497, 2003.

[21] D. Braess. The contraction number of a multigrid method for solving the Poisson equation. *Numerische Mathematik*, 37:387–404, 1981.

[22] D. Braess. The convergence rate of a multigrid method with Gauß-Seidel relaxation for the poisson equation. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods*, number 960 in Lecture Notes in Mathematics, pages 368–386. Springer, Berlin, 1982.

[23] D. Braess. *Finite Elemente*. Springer Berlin, Heidelber, New York, 2nd edition, 1997.

[24] D. Braess. *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, Cambridge, UK, second edition, 2001.

[25] D. Braess and W. Hackbusch. A new convergence proof for the multigrid method including the V-cycle. *SIAM J. Numer. Anal.*, 20:967–975, 1983.

[26] D. Braess and R. Sarazin. An efficient smoother for the Stokes problem. *Appl. Numer. Math.*, 23:3–19, 1997.

[27] J.H. Bramble. *Multigrid Methods*. Longman Scientific & Technical, Longman House, Essex, England, 1993.

[28] J.H. Bramble and J.E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comput.*, 50(181):1–17, 1988.

[29] J.H. Bramble, J.E. Pasciak, and A.T. Vassilev. Analysis of the inexact Uzawa algorithm for saddle point problems. *SIAM J. Numer. Anal.*, 34:1072–1092, 1997.

[30] J.H. Bramble, J.E. Pasciak, and A.T. Vassilev. Uzawa type algorithms for nonsymmetric saddle point problems. *Math. Comp.*, 69:667–689, 2000.

[31] J.H. Bramble, J.E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55:1–22, 1990.

[32] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comput.*, 31:333–390, 1977.

[33] A. Brandt. Rigorous quantitative analysis of multigrid, I: Constant coefficients two-level cycle with l_2-norm. *SIAM J. Numer. Anal.*, 31:1695–1730, 1994.

[34] S. Brenner. Convergence of the multigrid V-cycle algorithm for second-order boundary value problems without full elliptic regularity. *Math. Comp.*, 71(238):507–525, 2002.

[35] S. C. Brenner. Multigrid methods for parameter dependent problems. *RAIRO, Modelisation Math. Anal. Numer.*, 30(3):265–297, 1996.

[36] S.C. Brenner. A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity. *SIAM J. Numer. Anal.*, 30(1):116–135, 1993.

[37] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, 2002.

[38] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *R.A.I.R.O.*, 8:129–151, 1974.

[39] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, 1991.

[40] M. Burger and W. Mühlhuber. Iterative regularization of parameter identification problems by sequential quadratic programming methods. *Inverse Problems*, 18(4):943–69, 2002.

[41] M. Burger and W. Mühlhuber. Numerical approximation of an SQP-type method for parameter identification. *SIAM Journal on Numerical Analysis*, 40(5):1775–1797, 2002.

[42] X. Chen. On preconditioned Uzawa methods and SOR methods for saddle point problems. *J. Comput. Appl. Math.*, 100:207–224, 1998.

[43] X.L. Chen and J. Zou. An inexact Uzawa-type iterative method for solving saddle point problems. *Int. J. Computer Math.*, 80:55–64, 2003.

[44] P. Ciarlet. *Mathematical Elasticity. Volume I: Three Dimensional Elasticity*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988.

[45] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.

[46] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.*, pages 77–84, 1975.

[47] M. Crouzeix and P.A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *RAIRO Anal. Numér.*, 7(R-3):33–76, 1973.

[48] T. Dreyer, B. Maar, and V. Schulz. Multigrid optimization in applications. *J. Comput. Appl. Math.*, 120(1-2):67–84, 2000.

[49] A. Ecker and W. Zulehner. On the smoothing property of multi-grid methods in the non-symmetric case. *Numerical Linear Algebra with Applications*, 3(2):161–172, 1996.

[50] H.C. Elman and G.H. Golub. Inexact and preconditioned Uzawa algorithms for saddle point problems. *SIAM J. Numer. Anal.*, 31:1645–1661, 1994.

[51] M. Fortin and R. Glowinski. *Augmented Lagrangian Methods. Applications to the Numerical Solution of Boundary Value Problems*, volume 15 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1983. Translated from the French by B. Hunt and D.C. Spicer.

[52] R.W. Freund and N.M. Nachtigal. QMR: A quasi-minimal residual method for non-Hermitian linear systems. *Numerische Mathematik*, 60(3):315–339, 1991.

[53] V. Girault and P.A. Raviart. *Finite Element Methods for Navier-Stokes Equations*. Springer Berlin, Heidelberg, 1986.

[54] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer Series in Computational Physics. Springer-Verlag, New York, 1984.

[55] G.H. Golub and C. Greif. On solving block-structured indefinite linear systems. *SIAM J. Sci. Comput.*, 24:2076–2092, 2003.

[56] G.H. Golub, C. Greif, and J.M. Varah. An algebraic analysis of a block diagonal preconditioner for saddle point problems. *SIAM J. Matrix Anal. Appl.*, 27:779–792, 2006.

[57] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston-London-Melbourne, 1985.

[58] W. Hackbusch. Fast solution of elliptic optimal control problems. *J. Optimization Theory Appl.*, 31:565–581, 1980.

[59] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer-Verlag, Heidelberg, 1985.

[60] W. Hackbusch. *Iterative Solutions of Large Sparse Systems of Equations*. Springer Verlag, New York, 1994.

[61] W. Hackbusch and U. Trottenberg, editors. *Multigrid Methods*. Number 960 in Lecture Notes in Mathematics. Springer, Berlin, 1982.

[62] S.B. Hazra and V. Schulz. Simultaneous pseudo-timestepping for PDE-model based optimization problems. *BIT*, 44(3):457–472, 2004.

[63] M.R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards*, 49:409–436, 1952.

[64] R.H.W. Hoppe, C. Linsenmann, and S.I. Petrova. Primal-dual Newton methods in structural optimization. *Comp. Visual. Sci.*, 9:71–87, 2006.

[65] R.H.W. Hoppe and S.I. Petrova. Primal-dual Newton interior point methods in shape and topology optimization. *Numer. Linear Algebra Appl.*, 11:413–429, 2004.

[66] M. Jung and U. Langer. *Methode der finiten Elemente für Ingenieure*. B.G. Teubner, Stuttgart-Leipzig-Wiesbaden, 2001.

[67] U. Langer and W. Queck. On the convergence factor of Uzawa's algorithm. *J. Comput. Appl. Math.*, 15:191–202, 1986.

[68] U. Langer and W. Queck. *Preconditioned Uzawa-type iterative methods for solving mixed finite element equations*, volume 3 of *Wissenschaftliche Schriftenreihe*. Technische Universität Karl-Marx-Stadt, Karl-Marx-Stadt, 1987.

[69] J. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Berlin-Heidelberg-New York: Springer-Verlag, 1971.

[70] W. Mühlhuber. *Efficient Solvers for Optimal Design Problems with PDE Constraints*. PhD thesis, Johannes Kepler University, Linz, Austria, 2002.

[71] P. Oswald. *Multilevel Finite Element Approximation. Theory and Applications*. Teubner Verlag, Stuttgart, 1994.

[72] Chr.C. Paige and M.A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis*, 12:617–629, 1975.

[73] O. Pironneau. *Optimal Shape Design for Elliptic Systems*. Springer Series in Computational Physics. Springer, New York, 1984.

[74] O. Pironneau. *Finite Element Methods for Fluids*. John Wiley & Sons, Chichester, 1989.

[75] W. Queck. The convergence factor of preconditioned algorithms of the Arrow-Hurwicz type. *SIAM J. Numer. Anal.*, 26:1016–1030, 1989.

[76] S. Reitzinger. *Algebraic Multigrid Methods for Large Scale Finite Element Equations*. Schriften der Johannes Kepler Universität Linz. Universitätsverlag Rudolf Trauner, Linz, Austria, 2001.

[77] A. Reusken. A new lemma in multigrid convergence theory. Report RANA 91-07, Eindhoven, 1991.

[78] T. Rusten and R. Winther. A preconditioned iterative method for saddle-point problems. *SIAM J. Matrix Anal. Appl.*, 13:887–904, 1992.

[79] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.

[80] Y. Saad and M.H. Schultz. GMRES: a generalized minimal residual algorithm for non-symmetric linear systems. *SIAM Journal on Scientific Computing*, 7(3):856–869, 1986.

[81] Y. Saad and H.A. van der Vorst. Iterative solution of linear systems in the 20th century. *J. Comput. Appl. Math.*, 123:1–33, 2000.

[82] J. Schöberl and W. Zulehner. On Schwarz-type smoothers for saddle point problems. *Numerische Mathematik*, 95:377–399, 2003.

[83] J. Schöberl and W. Zulehner. Symmetric Indefinite Preconditioners for Saddle Point Problems with Applications to PDE-Constrained Optimization Problems. *SIAM J. Matrix Anal. Appl.*, 29:752–773, 2007.

[84] V. Schulz and G. Wittum. Transforming smoothers for PDE constrained optimization problems. *Comput. Visual Sci*, 2008. online available.

[85] R. Simon and W. Zulehner. On Schwarz-type smoothers for saddle point problems with applications to PDE-constrained optimization problems. Technical Report 01, Special Research Program SFB F013, JK University of Linz, January 2007.

[86] K. Stüben. An introduction to algebraic multigrid. In *Multigrid*, pages 413–533. Academic Press, London, 2001.

[87] K. Stüben and U. Trottenberg. Multigrid methods: Fundamental algorithms, model problem analysis and applications. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods*, number 960 in Lecture Notes in Mathematics, pages 1–176. Springer, Berlin, 1982.

[88] S. Ta'asan. "One-shot" methods for optimal control of distributed parameter systems I: The finite dimensional case. ICASE-Report 91-2, NASA Langley Research Center, Hampton VA, 1991.

[89] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen. Theorie, Verfahren und Anwendungen*. Vieweg, Wiesbaden, 2005.

[90] U. Trottenberg, C. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, London, 2001.

[91] H. Uzawa. Iterative methods for concave programming. In K.J. Arrow, L. Hurwicz, and H. Uzawa, editors, *Studies in Linear and Nonlinear Programming*, pages 154–165. Stanford University Press, Stanford, CA, 1958.

[92] S. Vanka. Block-implicit multigrid solution of Navier-Stokes equations in primitive variables. *J. Comput. Phys.*, 65:138–158, 1986.

[93] P. Vassilevski and R. Lazarov. Preconditioning mixed finite element saddle-point elliptic problems. *Numer. Linear Algebra Appl.*, 3:1–20, 1996.

[94] R. Verfürth. A multilevel algorithm for mixed problems. *SIAM J. Numer. Anal.*, 21:264–271, 1984.

[95] P. Wesseling. *An Introduction to Multigrid Methods*. Wiley, Chichester, 1992.

[96] P. Wesseling. *Principles of Computational Fluid Dynamics*, volume 29 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2001.

[97] G. Wittum. Multi-grid methods for Stokes and Navier-Stokes equations. Transforming smoothers: Algorithms and numerical results. *Numer. Math.*, 54(5):543–563, 1988.

[98] G. Wittum. On the convergence of multi-grid methods with transforming smoothers. *Numer. Math.*, 57(1):15–38, 1990.

[99] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992.

[100] K. Yosida. *Functional Analysis*. Springer Berlin, Heidelberg, New York, 6th edition, 1980.

[101] W. Zulehner. A class of smoothers for saddle point problems. *Computing*, 65:227–246, 2000.

[102] W. Zulehner. Analysis of iterative methods for saddle point problems: A unified approach. *Math. Comput.*, 71(238):479–505, 2002.

Eidesstattliche Erklärung

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

Diese Dissertation wurde bisher weder im In- noch im Ausland als Prüfungsarbeit vorgelegt.

Linz, Juni 2008 Dipl.-Math.techn. René Simon

Curriculum Vitae

Name: René Simon

Date of birth: April 13, 1976

Place of birth: Berlin, Germany.

Nationality: German.

Affiliation: Special Research Program (SFB) F013
Numerical and Symbolic Scientific Computing
Johannes Kepler University Linz
Altenbergerstraße 69, A-4040 Linz, Austria
`http://www.sfb013.uni-linz.ac.at`

Education:

September 1982 - July 1986	Elementary school.
September 1986 - July 1996	Grammar school.
April 1998 - July 2004	studies of *Technical Mathematics* at Technical University Berlin.
July 2004	Diploma in Optimal Control.
September 2004	Started as a PhD student at SFB F013, Project F1309: *Multilevel Solvers for Large Scale Discretized Optimization Problems.*

Employment:

September 2004 - March 2008	Research Assistant at the SFB F013.
Since April 2008	Research Assistant at University Duisburg-Essen